AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology:
Human Perception and Performance

ISSN: 0096-1523

2025, Vol. 51, No. 4, 492–506
https://doi.org/10.1037/xhp0001287

Template-Based and Saliency-Driven Attentional Control Converge to Coactivate on a Common, Spatially Organized Priority Map

Zexuan Niu, J. Toby Mordkoff, and Andrew Hollingworth
Department of Psychological and Brain Sciences, The University of Iowa

Visual attention can be controlled both by a match to known target attributes (template-based guidance) and by physical salience (saliency-driven guidance). However, it remains unclear how these mechanisms interact to determine attentional priority. Here, we contrasted two accounts of this interaction. Under a *coactive* mechanism, template-based and saliency-driven guidance are simultaneously integrated in a common priority signal. Under a *noncoactive* mechanism, the two sources of control do not converge on a common priority signal, either because they are separated architecturally (separate-activations model) or temporally (sequential model). In a redundancy-gain paradigm, search targets were defined either as a match to a shape cue (template-based), the presence of a singleton-colored item (saliency-driven), or both (redundant). We assessed whether the response time distribution in the redundant condition contained a substantial proportion of trials that were faster than could have been generated by the faster of the two individual guidance processes operating independently in parallel, that is, violation of the race model inequality (RMI). This effect can be generated only by a coactive mechanism. The results showed robust violations of the RMI when both features appeared at the same location, consistent with a coactive model. In addition, violations of the RMI were eliminated when redundant features were displayed at different locations, indicating that guidance signals combine on a spatially organized priority map.

Public Significance Statement
In everyday life, humans must efficiently direct their attention to relevant items, such as attending to a hammer when wishing to drive in a nail or attending to a looming object to quickly avoid it. Thus, understanding the mechanisms by which human attention is controlled is central to understanding how people can successfully navigate their environment and implement goal-directed behavior. Multiple factors influence where attention is oriented. Here, we examined how multiple sources of attention guidance are integrated. The results indicate that multiple sources of guidance converge on a common representation of behavioral relevance and that this combined representation is organized by spatial location.

Keywords: attention, visual search, perceptual saliency, cognitive control

Where we direct attention within a complex visual scene strongly influences what we consciously perceive (e.g., Most et al., 2005), what we identify (e.g., Lachter et al., 2004), the actions we perform (e.g., Johansson et al., 2001; Land & Hayhoe, 2001), and the memories we form (e.g., Hollingworth & Henderson, 2002). With such fundamental consequences of attention for our internal models of the world, it has long been recognized that understanding human perception, action, and memory depends on understanding how attention is controlled. What determines where attention is directed within a scene?

A key distinction in theories of attention has been between template-based control and saliency-driven control (Wolfe, 1994; Wolfe et al., 1989). Template-based control reflects the prioritization of items that match an internal representation of known feature values, such as selectively attending to blue items when informed that the target will be blue (e.g., Egeth et al., 1984) or selectively attending to small, cylindrical items when the target of search is a pen (e.g., Yang & Zelinsky, 2009). Saliency-driven control reflects the prioritization of items that are physically distinct, such as a uniquely colored item among homogeneously colored items or a moving item among stationary items. Guidance by physical salience is observed both when the salient item is the target of search (e.g., Bravo & Nakayama, 1992; Treisman & Gelade, 1980) and when it is a distractor (e.g., Franconeri et al., 2004; Theeuwes, 1992). In both cases, the stimulus does not match a known template value,

This article was published Online First February 17, 2025.
Iring Koch served as action editor.
Zexuan Niu https://orcid.org/0009-0002-2522-3510
Part of the research was presented at the 2022 Psychonomic Society Conference and the 2023 Vision Sciences Society Conference. The data and analysis code are available at https://osf.io/a9x3n/?view_only=62984c826c2f4f088ea8537bdfb8ec0b.

Zexuan Niu served as lead for data curation. Zexuan Niu, J. Toby Mordkoff, and Andrew Hollingworth contributed equally to conceptualization, methodology, writing–original draft, writing–review and editing, and formal analysis. Zexuan Niu and Andrew Hollingworth contributed equally to investigation and visualization.

Correspondence concerning this article should be addressed to Zexuan Niu, Department of Psychological and Brain Sciences, The University of Iowa, Psychological and Brain Sciences Building, 340 Iowa Avenue, Iowa City, IA 52242, United States. Email: zexuan-niu@uiowa.edu

and guidance is solely dependent on physical saliency.[1] Although both forms of guidance can influence the locus of attention, an important, unresolved question in the study of attentional control is how template-based and saliency-driven mechanisms of guidance interact to determine attentional priority.

For the purpose of understanding this interaction, we assume that the allocation of attention is driven by activity in one or more spatially organized priority maps. A priority map is a representation of relative attentional salience/priority, where stimuli at multiple locations generate signals that represent their strength as candidates for the recruitment of visual attention and gaze, typically in a competitive architecture. There are potentially multiple neural priority maps that span the visual hierarchy. Early in the visual processing stream, multiple salience maps can be maintained for different perceptual dimensions (Thayer & Sprague, 2023). As priority representations get more proximal to motor systems, such as the oculomotor system, it is likely that integration of different forms of guidance is necessitated by the need to select a single target location as the goal of the motor response (e.g., Thompson et al., 2005). Such higher-level maps, integrating information from multiple sensory dimensions, are found in posterior parietal cortex, frontal eye fields, and superior colliculus (for reviews, see Bisley & Goldberg, 2010; Fecteau & Munoz, 2006; Li, 2002; Thompson & Bichot, 2005).

Here, we consider two broad mechanisms of the possible interaction between template-based and saliency-driven guidance for determining relative activity on attentional priority maps. In the first, which we term a *coactive* mechanism, template-based and saliency-driven guidance are simultaneously integrated in a common priority signal to guide attention. In the second, which we term a *noncoactive* mechanism, although both sources of information have the capability to drive attention, they do not necessarily converge to generate an integrated priority signal. Note that coactivation, or absence of coactivation, could occur at different levels of the perceptual hierarchy, depending on the type of priority map that is functional in driving attention in a given task. For example, a task that requires simple detection might be driven by activity in relatively early, dimension-specific salience maps (Treisman & Gelade, 1980), but a task that requires executing a saccade to a target may require higher-level priority maps more proximal to the oculomotor system.

A coactive mechanism has two main requirements, illustrated in Figure 1A. First, a common priority map must receive and integrate biasing signals from the two sources of guidance. Second, the arrival of these biasing signals on the priority map must overlap substantially in time (Miller, 1986). Noncoactive mechanisms could fail to meet one or both conditions. Failing to meet just the first condition (Figure 1B) would imply that, even though two sources of guidance can potentially operate simultaneously, they are independent mechanisms of control and thus do not converge on a common priority signal. Specifically, whichever guidance system reaches threshold activation first will solely determine the locus of attention. We term this a *separate-activations* model. Failing to meet just the second condition (Figure 1C) would imply that, although the architecture of the system may support integration on a common priority map, the two biasing signals operate at different times within a perceptual episode and thus do not have sufficient temporal overlap to support coactivation. We term this a *sequential* model. A further theoretical possibility is that a noncoactive mechanism meets neither condition (Figure 1D), which we term a *separate-sequential* model. We will not consider this final possibility further, as its predictions are redundant with those of the separate-activations and sequential models.

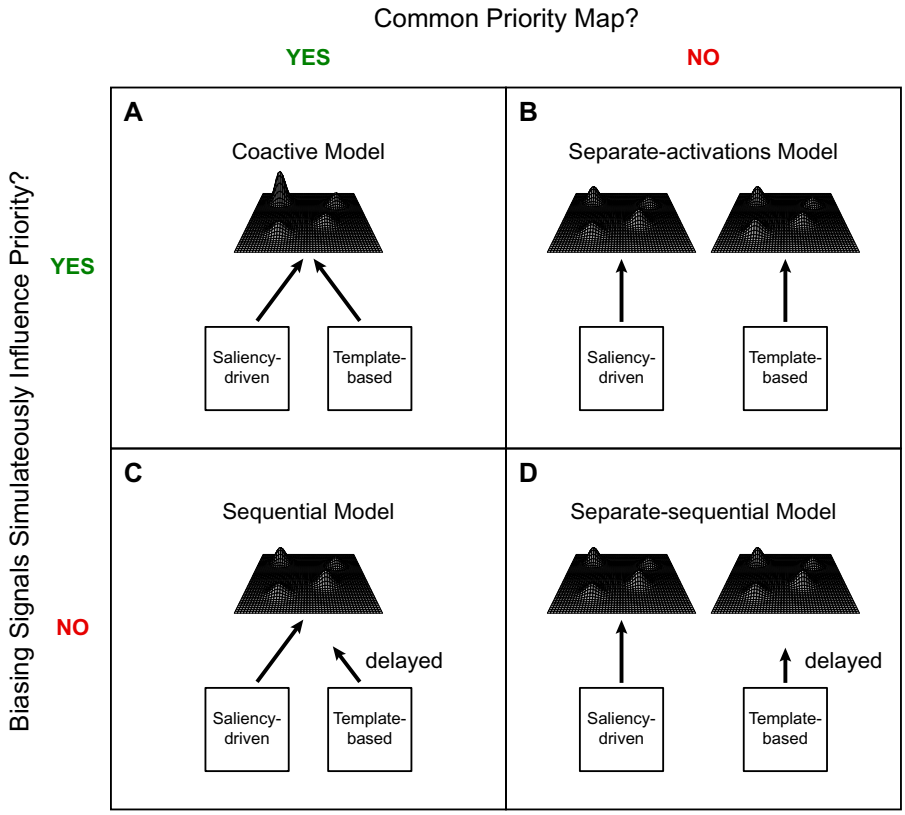
Coactive models are the most common in theories of visual attention. A prominent example of this architecture is the Guided Search model (Wolfe, 1994, 2021). Here, activity in feature maps for different perceptual dimensions is based on both local physical differences (saliency-driven) and knowledge of task-relevant values (template-based). Specifically, task-related template representations filter the activity in separate, dimension-specific feature maps to increase the relative strength of matching values. Activity in local feature maps is then summed on a common priority map, with attention directed sequentially to the locations of items with the highest priority. This basic assumption—that template values filter perceptual representations, integrating saliency-driven and template-based factors—is shared by several other prominent models of attention control (Desimone & Duncan, 1995; Hamker, 2005; Liesefeld et al., 2018; Navalpakkam & Itti, 2005; Serences et al., 2005). Note that there are alternative forms of coactive architecture. For example, saliency-driven and template-based signals could be processed independently until they converge on the priority map itself (e.g., Torralba et al., 2006). Although coactive models have been successful in explaining a range of empirical phenomena, there has yet to be a direct test of the core claim that saliency-driven and template-based signals are combined in an integrated priority signal. That is, there has yet to be a test that could distinguish coactive guidance from possible forms of noncoactive guidance.

Moreover, several lines of evidence are more consistent with a noncoactive mechanism than with a coactive mechanism. A recent study by van Heusden et al. (2022; see also van Zoest et al., 2004) supported a sequential model, in which saliency-driven guidance precedes, and thus does not coactivate with, template-based guidance. Targets in their visual search task were defined both by physical salience and by match to a template attribute. The van Heusden et al. (2022) results indicated that the two periods of guidance were nonoverlapping: saccades driven by physical salience tended to occur during the first 250 ms following stimulus onset, whereas saccades driven by match to the template tended to occur after 250 ms. If the two sources of guidance do not function simultaneously, they

[1] A clear definition of terms is necessary to ensure that the referenced representations and processes are specified precisely (Liesefeld et al., 2024). We consider guidance processes as falling into the category of *saliency-driven* when the differences between representations of local objects, leading to differences in attentional prioritization, are inherent in the physical properties of the stimulus itself. As mentioned above, this would include cases where a physically distinct item is not task-relevant (as in the literature on attention capture) and cases where a physically distinct item is relevant as the target of search (as in the literature on "pop-out" search). That is, we do not consider task relevance as a critical factor for defining saliency-driven guidance. For *template-based* guidance, differential priority is not necessarily inherent in the stimulus properties and is, instead, induced by filtering based on an internal representation of relevant stimulus values. We use the term *template* to describe any internal representation of target feature values that modulates priority in a manner consistent with the current task set. As may be evident from this discussion, our terminology maps quite closely to the classic distinction between *bottom-up* and *top-down* guidance of attention (specifically, as defined in Wolfe, 1994). Note that we avoid invoking a distinction between *stimulus-driven* and *goal-directed* guidance, as this dichotomy fails to distinguish between goal-relevant orienting that is implemented based on physical salience (as in "pop-out" search) and goal-relevant orienting that is induced by the application of template-based modulation (as in the search for a nonsalient, template-matching item).

Figure 1

Illustration of Coactive and Noncoactive Mechanisms of Saliency-Driven and Template-Based Guidance



Note. A coactive model (Panel A) must meet two criteria: (a) the two sources of bias converge on a common priority map signal and (b) the two sources of bias overlap temporally. A model that does not meet the first criterion posits independent systems of attentional control (separate-activations model, Panel B). A model that does not meet the second criterion posits that the sequential applicational of the two sources of control functionally precludes coactivation (sequential model, Panel C). A model that meets neither criterion is illustrated in Panel D. See the online article for the color version of this figure.

cannot generate an integrated priority signal. Interestingly, another recent study has made the reverse argument: that template-based processes precede saliency-driven processes (Becker et al., 2023). These authors observed that early guidance constituted a template-matching process, whereas saliency-driven processes tended to influence later identification processes, after attention had already been directed to a candidate target.

Sequential models have gained the most prominence in the literature on attention capture. Theeuwes (1992, 2010) has claimed that there is an initial period of attention guidance, following stimulus onset, in which the locus of attention is determined solely by physical salience. Only after this initial period of saliency-driven guidance can goal-directed biases, such as template-based processes, be applied. In the context of attention capture, attention is initially drawn to regions of high physical salience, independently of task relevance (Jonides & Yantis, 1988; Theeuwes, 1993). Although this model was developed to explain attention capture, the claims apply broadly to any situation in which saliency-driven and template-based guidance are both available. Thus, the Theeuwes model is an example of a noncoactive account, at least during the period immediately following stimulus onset.

In addition to noncoactive models of the sequential variety, separate-activations models find support in the literature. Again, in this model, the two sources of guidance may be available simultaneously but depend on independent systems of attentional control. Pinto et al. (2013) tested whether searching for a template-matching item and for physical saliency depend on overlapping mechanisms. They had participants perform both types of tasks. Performance on the two tasks was uncorrelated at a participant level, leading the authors to argue that the two sources were independent mechanisms. Separate systems for template-based and saliency-driven control are also consistent with claims that the two depend on different cortical networks (Ciaramelli et al., 2010; Corbetta & Shulman, 2002; Hahn et al., 2006), although initial separation need not preclude the possibility that the two sources of guidance are integrated at a later stage.

A separate-activations architecture is also possible in the specific case that both saliency-driven and template-based processes could be implemented by modulating activity in separate, dimension-specific saliency maps. This might occur if the effect of the template-based process was to increase the perceptual salience of matching items (Desimone & Duncan, 1995; Liesefeld et al., 2018; Mazer &

Gallant, 2003), and the template-based and saliency-driven processes were on different perceptual dimensions, recruiting separate saliency maps in relatively early visual regions. In this situation, the two dimension-specific salience signals could race to uniquely determine priority in higher-level maps controlling the allocation of attention.[2] Note that the present method was designed to be sensitive to this possibility. The basic task involved redundant template- and saliency-driven signals on different perceptual dimensions.

One key reason it has been difficult to distinguish between coactive and noncoactive mechanisms is that they often generate similar predictions on traditional behavioral indexes, such as mean response time (RT). Consider a visual search task in which the target either matches the search template value (template-based condition), is physically salient (saliency-driven condition), or both matches the search template value and is physically salient (redundant condition). All models considered so far predict lower mean RT in the redundant condition compared with the average of the two single-target conditions: The coactive model because the two redundant sources will sum on a common priority map, leading to more efficient guidance than by either source alone; the sequential model because the redundant trials will always allow guidance by the faster of the two sources; and the separate-activations model because the two independent processes will race toward an orienting solution, providing statistical facilitation of RT (Raab, 1962) compared with guidance by any single source alone (i.e., a "horse race" advantage).

In the present study, we sought to distinguish between the predictions of these models. The basic method is illustrated in Figure 2A. In the visual search task, participants first saw a shape cue, followed by a circular array of colored shapes. They were instructed that the target was a match to the shape cue and/or a color singleton. Specifically, there were three target-present conditions, each requiring the same response. In the template-based condition, the cued shape was present in the array, and all objects had the same color. In the saliency-driven condition, the cued shape was not present in the array, and one item had a color different from the others. We refer to these two conditions, collectively, as single-target trials. Finally, in the redundant condition, the cued shape was present, and this item was also a color singleton. Target-absent trials were included, in which the cued shape was not present and all items had the same color.

As discussed above, all models under consideration predict a redundancy gain: faster mean RT on redundant trials compared with single-target trials. However, well-developed techniques exist for testing the distributional predictions of coactive and noncoactive mechanisms in this type of design (Miller, 1982; Mordkoff & Danek, 2011; Mordkoff & Miller, 1993). Broadly put (details are provided in the Method), we tested for the presence of redundant-target trials with RTs that were faster than could have been generated by the faster of the two single-target guidance processes operating in parallel. Specifically, the separate-activations model holds that the two guidance processes each constitute an independent opportunity to observe a response by a particular time. Under this model, the sum of the probabilities observed in the single-target conditions places an upper bound on the response probabilities that should be observed in the redundant-target condition. This bound is termed the *race model inequality* (RMI; Miller, 1982; Mordkoff & Yantis, 1991). The same bound applies to all noncoactive mechanisms, including the sequential model.

If, for the faster responses in the distribution, the response probability in the redundant-target condition consistently exceeds the sum of the single-target probabilities, this *violation* of the RMI demonstrates that there was a substantial proportion of trials with RTs faster than could have been generated by the faster of two independent guidance processes operating in parallel. This type of result indicates that the two processes were integrated in generating the response, thereby supporting the coactive model. In contrast, if the probability of a response in the redundant-target condition never exceeds the sum of the single-target probabilities, this would be consistent with noncoactive mechanisms, such as the sequential model or separate-activations model (although the absence of RMI violations cannot be taken as strong positive evidence in favor of such models).

Experiment 1

In Experiment 1, we tested whether redundant template-based and saliency-driven guidance produces violations of the RMI. That is, we tested whether these two sources of guidance are integrated simultaneously on a common priority map. We conducted two versions of the experiment: (a) an initial, online version and (b) an in-person version to confirm that the results would replicate under more controlled conditions.

Method

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The data and analysis code are available at https://osf.io/a9x3n/?view_only=62984c826c2f4f088ea8537bdfb8ec0b. Data were analyzed using R (Version 4.3.1) and JASP (Version 0.17.3). This study's design and its analysis were not preregistered.

Participants

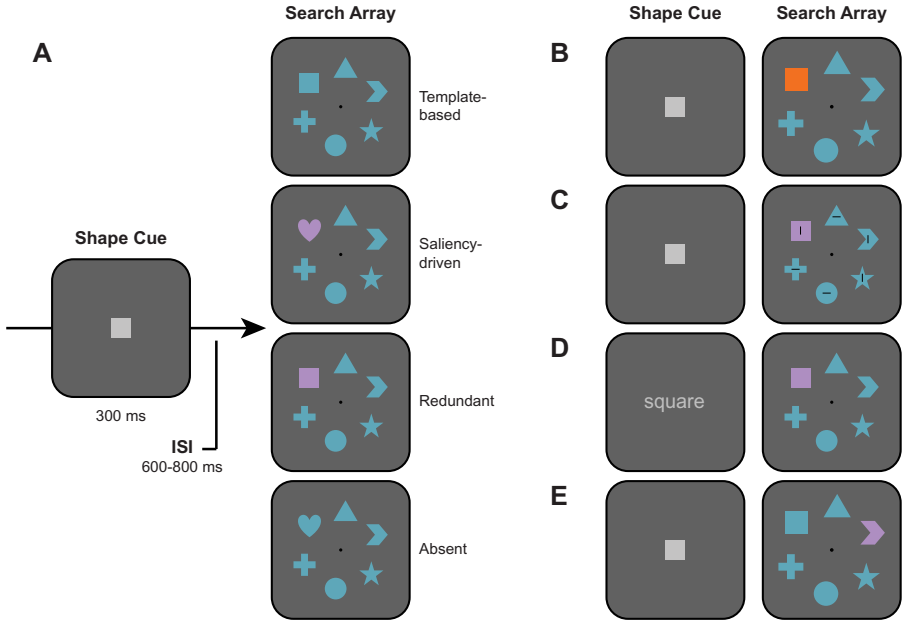
We aimed to include 24 participants for the in-person study and 32 for the online study. A given participant completed only one experiment. All procedures were approved by the University of Iowa Institutional Review Board. The participants were recruited from the University of Iowa SONA subject pool, were between the ages of 18 and 30, and were naive to the hypotheses under investigation. They received course credit for their participation.

The choice of N was conducted as follows. First, we based our power analysis on the simple contrast testing for a violation of the RMI (described in detail below), which in previous studies has tended to produce a smaller effect size than the other effect of interest here, the overall redundancy gain (e.g., Bahle et al., 2020). When violations of the RMI are both predicted and observed, one or more of the contrasts typically produces an effect size of at least $d_z = 0.8$ (Bahle et al., 2020; Mordkoff & Yantis, 1991). A d_z of 0.8 corresponds to f^2 of 0.64,

[2] Krummenacher et al. (2001, 2002) and Zehetleitner et al. (2009) have examined this question by asking, similarly to the approach in the present study, whether saliency signals from different perceptual dimensions coactivate in the guidance of attention. They observed evidence in favor of coactivation, indicating that dimension-specific salience signals do not implement a separate-activations model but are, instead, integrated into a common representation of priority. However, recent attempts to replicate these results have not been successful (Bahle et al., 2024; Liesefeld et al., 2017). Thus, we consider it an open possibility that activity in separate, dimension-specific salience maps race to uniquely determine attentional priority.

Figure 2

Illustration of the Method and Experimental Design



Note. (A) In Experiments 1A and 1B, participants first saw a shape cue (drawn randomly on each trial from seven possible shapes), followed by a search array consisting of six different shapes. Participants were instructed to respond "present" if the array contained the cued shape and/or if the array contained a color singleton. The shapes were drawn in a base color, selected randomly on each trial from 360 evenly spaced values around a CIELAB color wheel. There were three target-present conditions. In the template-based condition, the cued shape was present in the array, and all shapes were drawn in the base color. In the saliency-driven condition, one shape was drawn in a discrepant color ($\pm 40°$ from the base color in color space), creating a color singleton; the array did not contain the cued shape. Finally, in the redundant condition, the cued shape was present, and this item was also the color singleton. Target-absent trials contained neither the cued shape nor a color singleton, requiring an "absent" response. Panels B–E illustrate the redundant condition in each of Experiments 2–5. (B) In Experiment 2, color singletons differed from the base color by 180° in color space. (C) In Experiment 3, each object contained an oriented line (horizontal/vertical), and participants reported the orientation of the line in the target item. (D) In Experiment 4, shapes were cued with a label rather than with an image of the shape. (E) In Experiment 5, the redundant condition contained a cue-matching shape and a color singleton, but these attributes were associated with different objects at different locations within the array. ISI = interstimulus interval. See the online article for the color version of this figure.

requiring 14 participants to achieve 80% power for $\alpha = .05$ (Cohen, 1988). For the in-person version, we conservatively chose an N of 24, which provides power sufficient to detect an effect of approximately $d_z = 0.6$. For the online study, since online data could potentially increase error variance, we chose an N of 32, which provides power sufficient to detect an effect of approximately $d_z = 0.5$.

We replaced participants with a false-alarm rate in the target-absent condition greater than 20% or a miss rate greater than 20% in any of the target-present conditions. Five participants were replaced in the in-person version of the study, and 11 were replaced in the online version of the study. Since multiple slots were posted simultaneously for the online version, it was possible to exceed our recruiting goal. Ultimately, 24 participants met the inclusion criteria for the in-person study (18 female and six male), and 33 met the inclusion criteria for the online study (26 female, five male, and two not reporting). For the included participants in Experiment 1A, the mean miss rate was 2.80%, and the mean false alarm rate was 6.78%. For the included participants in Experiment 1B, the mean miss rate was 3.73%, and the mean false alarm rate was 10.15%.

Apparatus

For the in-person version of the study, the stimuli were presented on a liquid-crystal display monitor (resolution: 1,920 × 1,080 pixels) with a refresh rate of 144 Hz at a viewing distance of 60 cm, maintained by a forehead rest. Manual responses were collected via a keyboard. The experiment was implemented in OpenSesame software (Mathôt et al., 2012). The online version was the same, except that participants completed the experiment using their personal computers. The experiment was converted to Javascript and hosted on a JATOS server maintained by the University of Iowa. We report stimulus size both in visual angle (for the in-person study) and in pixels (for the online study).

Stimuli

All stimuli were presented against a gray (R:96, G:96, B:96) background with a black central fixation ring subtending 0.49° × 0.49° visual angle (20 × 20 pixels). Seven possible shape values (square, star, triangle, circle, heart, chevron, and plus sign) were included in

the current study. Colors were chosen from a CIELAB color space ($L = 65$, $a = 20$, $b = 30$, radius = 60, D65 illuminant).

For the cue display, one shape was chosen randomly on each trial from the seven possible shape values. Random selection of the shape cue from trial to trial ensured that participants maintained an active template representation (Carlisle et al., 2011). Specifically, search for static targets tends to be less efficient than search for dynamically changing targets (Berggren et al., 2020), which could have placed template-based guidance at a relative disadvantage compared with saliency-driven guidance if the shape cue had been constant. The use of randomly changing shape cues also avoided the possibility of long-term learning effects. Extended exposure to static target and/or distractor values may lead to long-term re-tuning of relatively early sensory responses (Adam & Serences, 2021), which could have potentially clouded the template-based/saliency-driven distinction.

The shape in the cue display was presented in gray (R:211, G:211, B:211). The search array consisted of six colored objects (3.17° × 3.17°; 120 × 120 pixels on average) evenly distributed around a virtual circle with a radius of 5.55° (210 × 210 pixels). The first array object location was chosen randomly; the remaining five objects were each offset by 60° around the virtual circle. Each array object had a unique shape, randomly assigned to the six locations. The objects were drawn in a base color, chosen randomly on each trial from 360 possible colors, evenly spaced around the color wheel.

For template-based trials (cue-matching target), one randomly chosen array item matched the cued shape. All of the objects, including the target, were drawn in the base color. For saliency-driven trials (singleton-color target), one array item was assigned a color value offset by ±40° from the base color. The array did not contain a cue-matching shape. Note that the random selection of color on each trial ensured that participants could not predict, and thus could not develop a search template for, the value of the singleton color; the target was defined by color discrepancy alone. For redundant trials, the color-singleton item also matched the shape cue. Finally, on target-absent trials, all shapes were drawn in the base color, and none matched the shape cue.

Procedure

The trial event sequence is illustrated in Figure 2A. Each trial began with a presentation of a central fixation cross for 500 ms, followed by the cue display for 300 ms, a variable ISI selected randomly between 600 and 800 ms, and finally the search array, which remained visible until manual keyboard response. Participants were instructed that a target could be a shape-matching object, an object with a discrepant color, or both. They pressed "Q" to indicate target present and "P" to indicate target absent and were instructed to respond as quickly and accurately as possible. On incorrect trials, feedback was provided ("Incorrect!" in Red, Arial font, 1,000 ms duration). After a correct response or incorrect feedback, there was a 500-ms blank inter-trial interval before the fixation cross for the next trial appeared.

For the in-person study, upon arriving for the experimental session and providing informed consent, participants were given oral and written instructions. For the online study, only written instructions were provided. Participants then completed a practice block of 24 trials. In the main experiment, they completed 30 experimental blocks of 24 trials each. The 24 trials in each block were divided evenly among four conditions: template-based target, saliency-driven target, redundant target, and target absent. Therefore, the probabilities that a display would include a template-matching or saliency-driven target were independent. Trials from the four conditions were randomly intermixed within each block. Participants were required to take a break between blocks, and they were able to rest for as long as desired. The entire experiment lasted approximately 60 min.

Data Analysis

The key test was whether the RT distributions in the redundant condition contained a substantial proportion of trials that were faster than could have been generated by the faster of the two individual guidance processes operating independently in parallel: that is, violation of the RMI. The RMI, which is a simple extension of Boole's Inequality, is defined as follows (Miller, 1982):

$$p(\mathrm{RT} < t \mid T\mathrm{t} \text{ and } T\mathrm{s}) \leq p(\mathrm{RT} < t \mid T\mathrm{t}) + p(\mathrm{RT} < t \mid T\mathrm{s}), \quad (1)$$

where the left side of the equation is the cumulative distribution function (CDF) for the redundant-target trials, and the right side is the sum of the two single-target CDFs. Note that *T*t refers to template-based trials, *T*s refers to saliency-driven trials, and *T*t and *T*s refer to redundant trials.

To implement the RMI test, we divided the CDF in the redundant condition into quantiles for the fifth through 95th percentiles, in 5% steps. Separately for each subject, the values of RT that corresponded to the fifth through 95th percentiles of the redundant condition were calculated. These were the values of t at which the RMI was tested for this subject. In the second step, adherence to the RMI was evaluated at each of the values of t by subtracting the sum of the two cumulative probabilities in the single-target conditions from the cumulative probability in the redundant condition. Only a positive value is a violation of the RMI, but the specific signed value is retained for hypothesis testing. In the last step, at each of the quantiles, the mean violation (across subjects) of the RMI was tested against zero using a one-sided, one-tailed, t test (i.e., we tested only for positive mean violations). Note that even when the underlying system involves coactivation, violations of the RMI are only expected in the faster end of the distribution (i.e., the lower percentiles), because the left side of the RMI has a maximum value of one, whereas the right side continues to two.

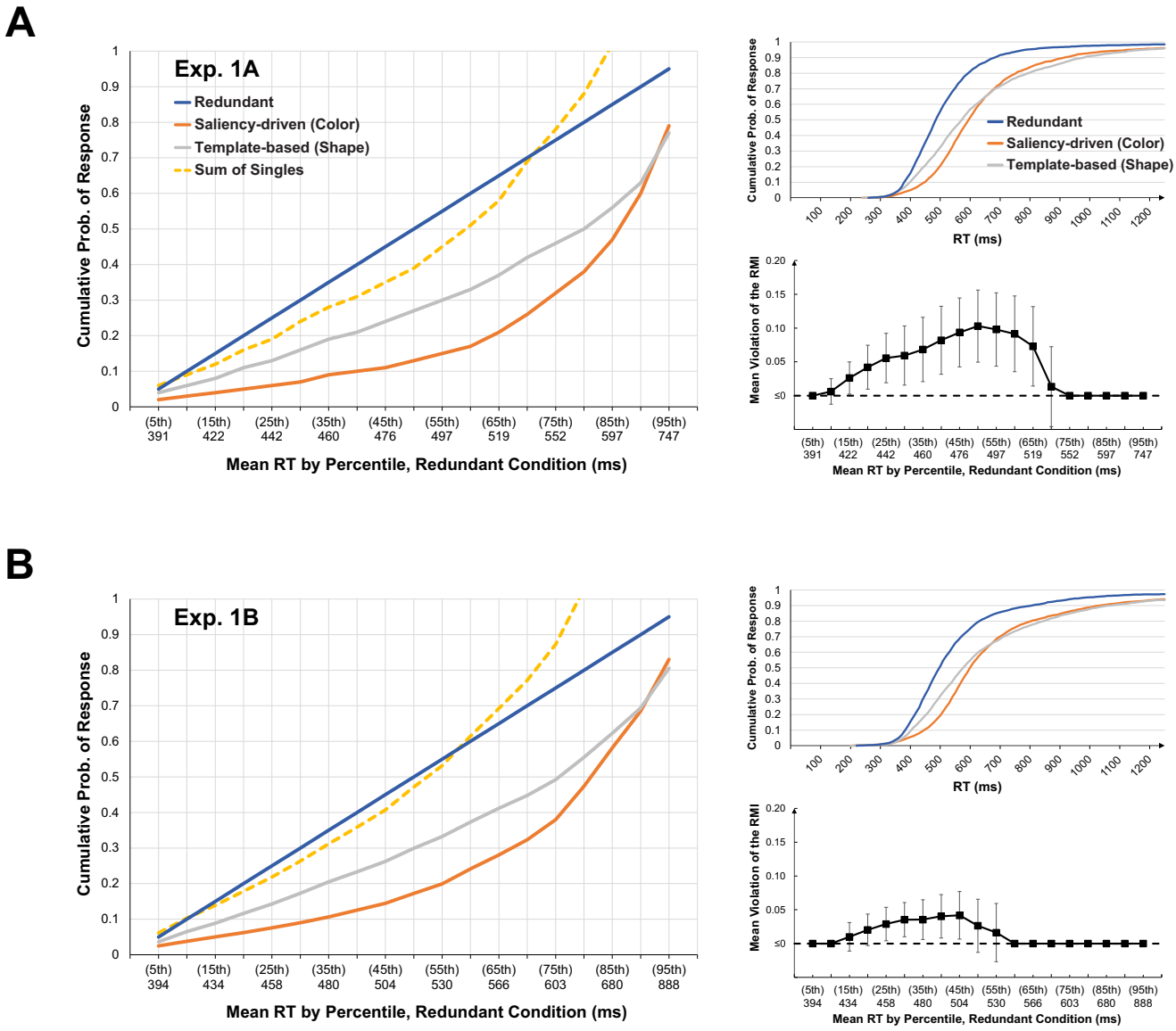
For the RMI and mean RT analyses, the following trials were removed (Bahle et al., 2020): (a) the first 10 blocks of the experiment to eliminate practice effects,[3] (b) the first two trials from the remaining blocks, which served as warm-up trials, (c) incorrect trials, (d) trials after incorrect responses, which can reflect posterror recovery effects (Danielmeier & Ullsperger, 2011), and (e) trials with RTs less than 200 ms (too rapid to have reflected target detection) or greater than 6,000 ms (reflecting lapses in task engagement). RT trimming led to the elimination of 0.10% of the remaining data in Experiment 1A and 0.49% of the data in Experiment 1B.

[3] In almost all laboratory attention tasks, performance improves substantially with practice, and RTs decline quite rapidly during the early stages of the experiment. This can compromise an RMI analysis, which is based on the distributions of absolute RT values. Thus, we first allow RTs to largely stabilize over the first 10 blocks of the experiment before we collect the data that enters into the RMI analysis (see also Bahle et al., 2020). The reasoning behind this approach is discussed more extensively in Mordkoff and Miller (1993).

Results

The results of Experiments 1A and 1B are displayed in Figure 3. We implemented three main analyses. First, we examined mean RT in the two single-target conditions to compare the efficiency of guidance from template-based and saliency-driven sources. Second, we examined redundancy gains in mean RT, comparing RT in the redundant condition with RT in the single-target conditions. Note that the sequential model makes an additional prediction in this context. Although this model predicts a redundancy gain relative to the combined data from

Figure 3
Results From Experiments 1A (A) and 1B (B)



Note. In each panel, the main graph shows the cumulative response distributions, referenced to the distribution in the redundant condition. Specifically, the blue (dark gray) line on the diagonal is the cumulative probability of a response in the redundant condition at each percentile of the RT distribution in the redundant condition. The orange (meduim gray) and light gray lines are the cumulative probability of response in each of the single-target conditions at the RT for each percentile of the redundant condition. The yellow (light gray) dashed line is the sum of the two single-target probabilities. Violations of the RMI are indicated when the yellow (light gray) line falls below the blue (dark gray): that is, the probability of a response in the redundant condition is higher than the sum of the probabilities in the two single-target conditions, consistent with coactivation. The top-right graph in each panel shows the raw CDFs in each of the three target-present conditions, averaged across participants. The bottom-right graph in each panel shows a violation plot illustrating the probability difference between the redundant condition and the sum of the two single-target conditions. Only positive differences (i.e., violations of the RMI) are displayed, as the test of the RMI is unidirectional. Error bars are 95% confidence intervals. Exp. = Experiment; Prob. = probability; RT = response time; RMI = race model inequality; CDF = cumulative distribution function. See the online article for the color version of this figure.

the two single-target conditions, it predicts no redundancy gain relative to the *faster* of the two single-target conditions for each participant. That is, if one source of guidance is always implemented before the other, mean RT for the faster source of guidance alone should establish a floor for performance in the redundant condition. Finally, we report the critical analysis probing violations of the RMI.

Mean RT in the Single-Target Conditions

For the in-person version, mean RT in the template-based condition was 641 ms, and mean RT in the saliency-driven condition was 650 ms. These were not reliably different, $t(23) = 0.80$, $p = .434$, $d_z = 0.16$. For the online version, mean RT in the template-based condition was 684 ms, and mean RT in the saliency-driven condition was 698 ms. They were not reliably different, $t(32) = 0.78$, $p = .443$, $d_z = 0.14$. The similar distributions of RT in the two conditions contradict the basic assumption of any sequential model. Specifically, this is inconsistent with the assumption that there is a period following stimulus onset in which only stimulus-driven guidance is functional (Theeuwes, 2010; van Heusden et al., 2022). Note that in Experiment 2, we maximized the salience of the singleton target to provide a stronger test of this model.

Redundancy Gains in Mean RT

First, mean RT on the redundant-target trials was compared with mean RT on single-target trials, collapsing across shape-target and color-singleton target. Reliable redundancy gains were observed in both the in-person study and the online study: in-person redundancy gain = 129 ms, $t(23) = 13.03$, $p < .001$, $d_z = 2.66$, and online redundancy gain = 128 ms, $t(32) = 13.53$, $p < .001$, $d_z = 2.36$. Second, we compared mean RT on redundant-target trials with mean RT for the faster of the two single-target conditions (determined on a participant-by-participant basis). Reliable redundancy gains were still observed in both experiments: in-person redundancy gain = 106 ms, $t(23) = 10.37$, $p < .001$, $d_z = 2.12$, and online redundancy gain = 96 ms, $t(32) = 10.97$, $p < .001$, $d_z = 1.91$.

Violations of the RMI

In both the in-person and online version, we observed reliable violations of the RMI (Figure 3). In the in-person version, reliable violations ($p < .05$) were observed from the 15th to the 65th percentile (Figure 3A), and d_z for these reliable violations ranged from 0.46 to 0.82. In the online version, reliable violations were observed from the 25th to 45th percentile (Figure 3B), and d_z for these reliable violations ranged from 0.39 to 0.47. Note that it is common that reliable violations are not observed in the very fastest quantiles, as sensitivity can be reduced by rapid, anticipatory responses (Miller, 1982).

We note that the range and magnitude of the observed violations differed between Experiment 1A and 1B despite nearly identical designs. This highlights the difficulty in comparing quantitative differences in violations across experiments. For example, violations of the RMI are only observed when the processing of the two targets overlap in time (see Miller, 1986), so any changes to the task, targets, or participant sample that cause changes in the amount of overlap will cause changes in the number and magnitude of violations. Similarly, even a few fast-guess responses can conceal evidence of coactivation at the lowest percentiles (see Eriksen, 1988), which can also be a source of differences in RMI violations across experiments.

Discussion

In Experiment 1, we observed robust violations of the RMI, indicating that saliency-driven and template-based processes coactivate in the guidance of attention. The results support the hypothesis that these two sources of guidance converge on a common priority signal. In contrast, our RMI results did not support either the sequential model or the separate-activations model. Two other aspects of the results also failed to support the sequential model. First, we observed a redundancy gain in mean RT relative to the *faster* of the two individual guidance processes, contrary to the prediction of the sequential model. Second, the time courses of guidance from template-based and saliency-driven processes were highly overlapping. This is consistent with other results indicating that systems responsible for template-based control, such as visual working memory, can influence attention during the first wave of sensory processing following stimulus onset (Hollingworth et al., 2013a, 2013b): template-based guidance need be no slower than saliency-driven guidance. In addition, target detection and processing have been previously observed to be no less efficient in a purely template-driven search task—detecting a circle among heterogenous shapes—than in a saliency-driven search task—detecting a circle among diamonds (Bacon & Egeth, 1994).

Experiments 2–4

We conducted three additional experiments to further probe how template-based and saliency-driven processes are integrated.

Experiment 2

The color singletons in Experiment 1 differed from the base color on each trial by 40°. This degree of physical salience was clearly sufficient to guide attention, because participants performed at near-ceiling levels for saliency-driven trials (i.e., when there was a salient item but no template-matching item). However, to maximize the salience signal, and thus provide a stronger test of the hypothesis that saliency-driven guidance precedes template-based guidance (Theeuwes, 2010; van Heusden et al., 2022), color singletons in Experiment 2 differed from the base color by 180° (Figure 2B). With the 180° color difference, along with the use of an array with a sufficiently large number of items (six), the singleton color should have had a level of salience similar to that of the salient distractor in a typical attention capture paradigm. Note, however, that the standard of saliency should not necessarily be set by attention capture in the present paradigm, since the salient item was the target, not a distractor.

Experiment 3

The basic task in Experiment 1 was simple detection, which theoretically does not require a focal shift of attention to the target location in order to respond. In Experiment 3, we tested whether evidence of coactivation would generalize to a task that explicitly required a shift of attention to the target location. Experiment 3 replicated Experiment 1, but with a secondary discrimination task instead of a detection task. Each object contained an oriented line,

and participants reported the orientation of the line in the target object, requiring them to shift attention to the target before the response (Figure 2C).[4]

Experiment 4

We assume that the target-detection response in the template-based and redundant conditions of Experiment 1 was based on a match to an active template representation of the cued shape. An alternative account is that responses were driven, or at least influenced by, perceptual priming from the cue image to the search array. That is, mere exposure to the cue image may have interacted with perceptual processing of the search array to increase the relative priority of the matching shape, which would not necessarily constitute a template-based process. To eliminate this alternative, in Experiment 4 we replicated Experiment 1, but the target shape was cued by a written label rather than by the shape stimulus itself (Figure 2D). The labels were visually arbitrary with respect to the shapes (the word "triangle" does not look like a triangle) and thus could not generate perceptual-level priming of the target shape.

Method

Participants

Experiments 2–4 were run online. We observed an effect size of approximately $d_z = 0.5$ in Experiment 1B, and thus we also recruited 32 new participants in each of Experiments 2–4. Participants in Experiment 2 were recruited from Prolific and were compensated at a rate of $12 per hour. Participants in Experiments 3 and 4 were recruited in the same manner as in Experiment 1B. Participants were replaced based on the same criteria in Experiment 1: seven in Experiment 2, nine in Experiment 3, and nine in Experiment 4. Thirty-three participants (18 female and 15 male) met inclusion criteria in Experiment 2, 33 (19 female and 14 male) in Experiment 3, and 32 (24 female, seven male, and one not reporting) in Experiment 4. For the included participants in Experiment 2, the mean miss rate was 3.08%, and the mean false alarm rate was 8.10%. For the included participants in Experiment 3, the mean error rate on the discrimination task was 5.49%. For the included participants in Experiment 4, the mean miss rate was 3.04%, and the mean false alarm rate was 8.54%.

Stimuli, Procedure, and Data Analysis

Each experiment included a modification of the Experiment 1 method.

Experiment 2: 180° Color Difference. Experiment 2 was identical to Experiment 1, except color singletons differed by 180° in color space from the base color.

Experiment 3: Line Discrimination. An internal black line (3 × 30 pixels) was displayed within each item, with line orientation (horizontal, vertical) chosen randomly. Participants reported the orientation of the line in the target item. Given this forced-choice task, target-absent trials were eliminated. Each block of 24 trials included eight trials in each of the three target-present conditions (template-based target, saliency-driven target, and redundant target).

Experiment 4: Label Cue. Experiment 4 was identical to Experiment 1, except the shape cue was presented in the form of a written label. The labels were: "square," "star," "triangle," "circle," "heart," "arrow," and "plus sign." The label "arrow" was used to describe the chevron shape, as "chevron" might not have been a familiar term to all participants. In the instructions at the beginning of the experiment, participants were shown the shapes and the corresponding labels.

Trials were excluded from analysis based on the same criteria as in Experiment 1. RT trimming led to the elimination of 0.17% of the remaining data in Experiment 2, 0.34% in Experiment 3, and 0.14% in Experiment 4.

Results

Experiment 2: 180° Color Difference

The results of Experiment 2 are displayed in Figure 4A. Mean RT in the template-based condition was 708 ms, and mean RT in the saliency-driven condition was 681 ms. These were not reliably different, $t(32) = 1.71$, $p = .097$, $d_z = 0.30$. A reliable redundancy gain of 129 ms was observed, $t(32) = 17.93$, $p < .001$, $d_z = 3.12$. A redundancy gain of 104 ms was still observed relative to the faster of the two single-target conditions, $t(32) = 12.93$, $p < .001$, $d_z = 2.25$. Finally, reliable violations of the RMI were observed from the 10th to the 50th percentile, and d_z for these reliable violations ranged from 0.50 to 0.91.

Experiment 3: Line Discrimination

The results of Experiment 3 are displayed in Figure 4B. Mean RT in the template-based condition was 1,096 ms, and mean RT in the saliency-driven condition was 1,042 ms. These were reliably different, $t(32) = 2.08$, $p = .046$, $d_z = 0.36$. A reliable redundancy gain of 196 ms was observed, $t(32) = 19.00$, $p < .001$, $d_z = 3.31$. A redundancy gain of 139 ms was still observed relative to the faster of the two single-target conditions, $t(32) = 12.35$, $p < .001$, $d_z = 2.15$. Finally, reliable violations of the RMI were observed from the fifth to the 35th percentile, and d_z for these reliable violations ranged from 0.39 to 0.90.

Experiment 4: Label Cue

The results of Experiment 4 are displayed in Figure 4C. Mean RT in the template-based condition was 811 ms, and mean RT in the saliency-driven condition was 697 ms. These were reliably different, $t(31) = 4.96$, $p < .001$, $d_z = 0.88$. A reliable redundancy gain of 172 ms was observed, $t(31) = 19.75$, $p < .001$, $d_z = 3.49$. A redundancy gain of 110 ms was still observed relative to the faster of the two single-target conditions, $t(31) = 12.37$, $p < .001$, $d_z = 2.19$. Finally, reliable violations of the RMI were observed from the 15th to the 60th percentile, and d_z for these violations ranged from 0.45 to 1.08.

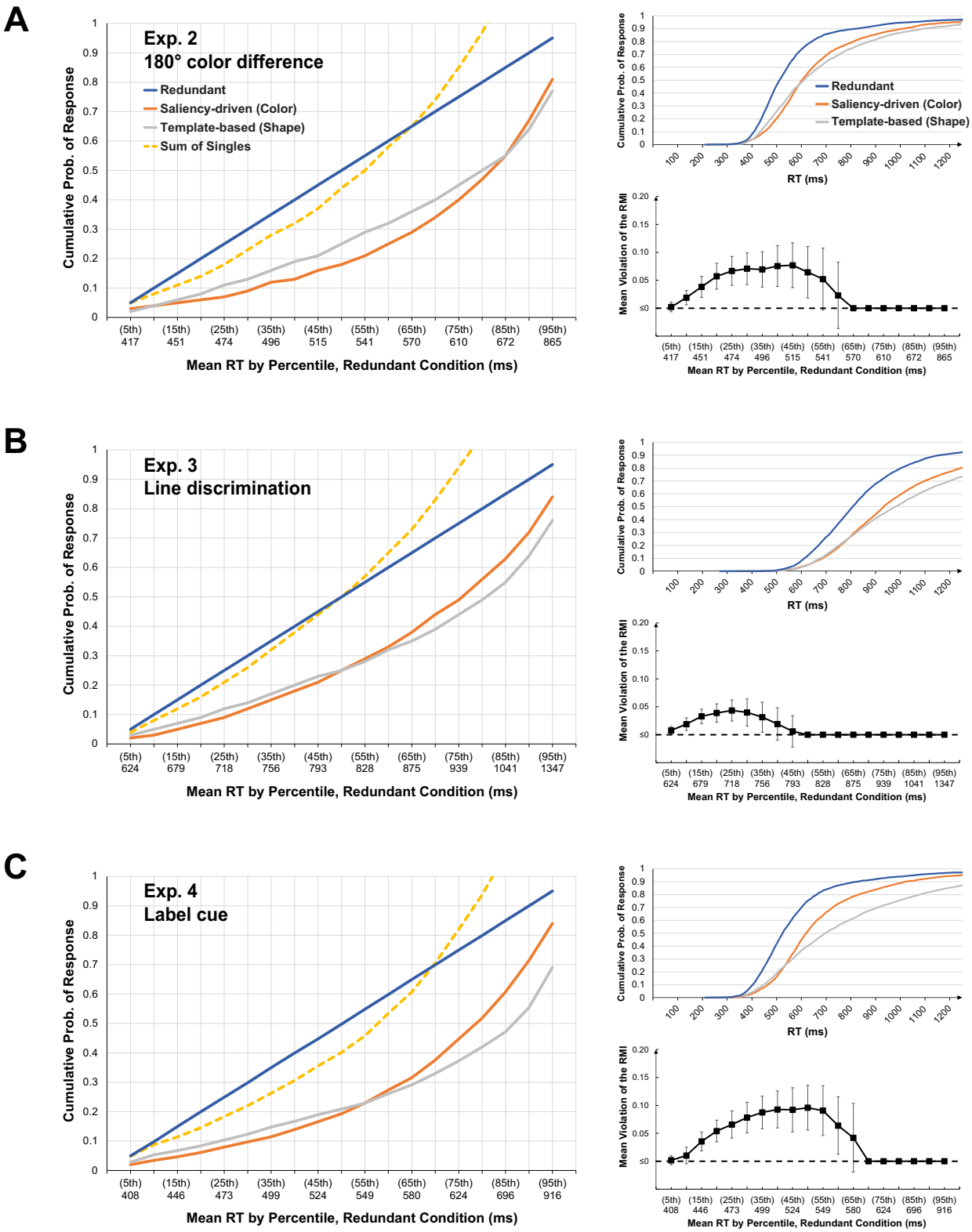
Discussion

Experiment 2 replicated the results of Experiment 1 when maximizing the color difference between the singleton color and the

[4] We used the secondary discrimination method only in Experiment 3, and not in the other experiments, because secondary discrimination adds line-discrimination and response-selection operations to the task, potentially introducing variability in RT after the proposed coactive process. Adding sources of variability after the coactive stage can severely limit the power of the method to observe violations of the RMI (Mordkoff & Yantis, 1991), making the secondary discrimination task suboptimal as a general method for probing coactivation.

Figure 4

Results From Experiments 2 (A), 3 (B), and 4 (C)



Note. Error bars are 95% confidence intervals. Exp. = Experiment; Prob. = probability; RT = response time; RMI = race model inequality. See the online article for the color version of this figure.

base color, providing a strong test of a sequential model in which saliency-driven processing precedes template-based processing. Despite this, the response distributions for template-based and saliency-driven trials remained highly overlapping, and robust violations of the RMI were observed. The similar results with color singletons that were 40° (Experiment 1) and 180° (Experiment 2) from the base color is consistent with recent evidence that attention guidance from salient singletons does not vary substantially across the range of color contrasts used here (Stilwell et al., 2023).

In Experiment 3, robust violations of the RMI were observed in a discrimination version of the paradigm that required a focal shift of attention to the target item. Coactivation could not have occurred during line discrimination or response generation, as the line orientation and response were orthogonal to the target-defining attributes. Thus, coactivation can be inferred to have occurred during the process of target detection and orienting.

Experiment 4 eliminated an alternative account in which the evidence of coactivation in Experiments 1–3 was driven by low-level priming from the shape cue to the search array. The written labels used in Experiment 4 could not have generated perceptual priming of the target shape, and yet robust violations of the RMI were still observed.

Note that we did find a mean RT difference between the two single-target conditions in Experiments 3 and 4. However, these differences were unlikely to have been caused by the sequential application of the two forms of guidance. As is evident in Figure 4B and 4C, the accumulation of positive response probabilities in the two single-target conditions started at essentially the same time. The difference was not in onset time, but rather in the rate of accumulation. In Experiment 4, this might have been caused by the fact that written labels provide less precise template representations compared with visually presented shape cues (Vickery et al., 2005; Wolfe et al., 2004).

Experiment 5

Theories proposing a coactive mechanism usually assume that the priority map functional in controlling attention is organized spatially, with signals from different sources of guidance influencing activity at particular locations (Hamker, 2005; Navalpakkam & Itti, 2005; Wolfe, 1994, 2021). Under this assumption, coactivation should occur only when the redundant features matching the template-based and saliency-driven processes are present at the same location. Experiment 5 replicated Experiment 1, but instead of having redundant features appear in the same object, the template-based and saliency-driven features were present within different objects at different locations (Figure 2E). We predicted that the evidence of coactivation would be eliminated.

Method

Participants

The effect sizes across the four online studies conducted so far ranged from approximately $d_z = 0.5$ to $d_z = 1.1$. Conservatively, we used the smallest effect size as the basis for Experiment 5, again recruiting a target N of 32 new participants. Participants were recruited in the same manner as in Experiment 1B, and the data were collected online. Twenty-one participants were replaced based on the same criteria as in Experiment 1. The larger proportion of replaced participants was due to an unusually large number who performed at near-chance levels. Experiment 5 was run very near the end of the academic semester, when many students are rushing to complete their research-participation requirement and the proportion of participants who fail to engage in experimental tasks is typically higher. Thirty-five participants (23 female, 11 male, and one not reporting) met the inclusion criteria. For the included participants in Experiment 5, the mean miss rate was 4.14%, and the mean false alarm rate was 9.00%.

Stimuli, Procedure, and Data Analysis

Experiment 5 was identical to Experiment 1, except for the redundant condition. The target shape and singleton color were associated with two different, randomly chosen objects in the array. The task instructions were unchanged: detect the presence of any shape-matching object or any discrepant color.

Trials were excluded from analysis based on the same criteria as in Experiment 1. RT trimming led to the elimination of 0.19% of the remaining data.

Results

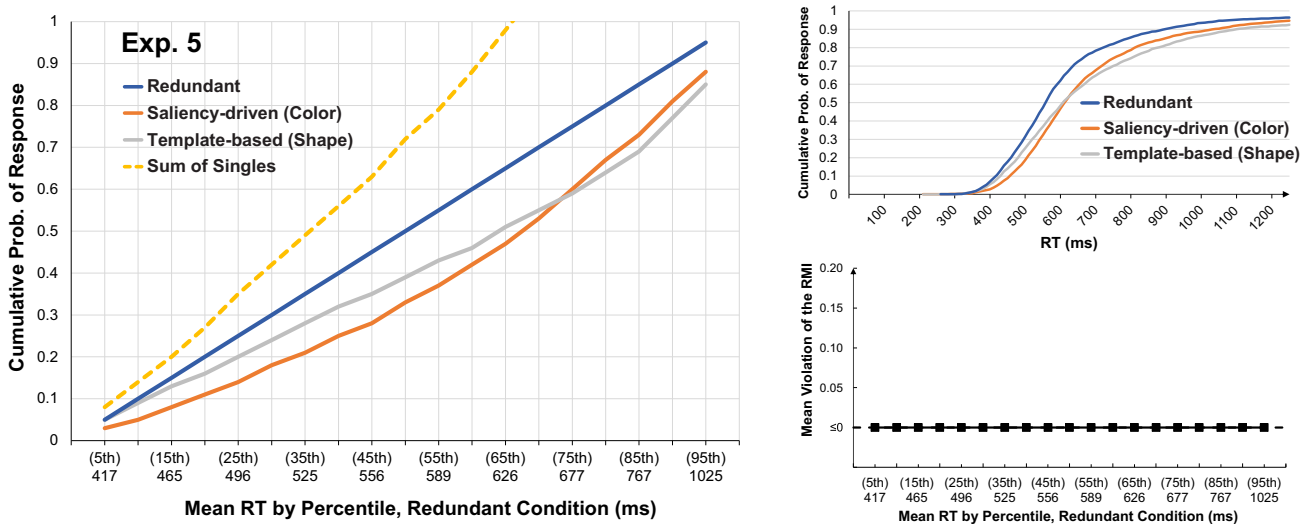
The results of Experiment 5 are displayed in Figure 5. Mean RT in the template-based condition was 712 ms, and mean RT in the saliency-driven condition was 699 ms. These were not reliably different, $t(34) = 0.67$, $p = .505$, $d_z = 0.11$. A reliable redundancy gain of 81 ms was observed, $t(34) = 9.12$, $p < .001$, $d_z = 1.54$. A redundancy gain of 42 ms was still observed relative to the faster of the two single-target conditions, $t(34) = 4.69$, $p < .001$, $d_z = 0.79$. Despite these redundancy gains, there were no violations of the RMI.

In addition, we compared the magnitude of the basic redundancy gain in Experiment 5 (spatially separated redundant signals) with the magnitude of the redundancy gain in Experiment 1B (spatially integrated redundant signals). Experiment 1B was equated with Experiment 5 on all design and participant sample properties except for the spatial separation of targets. The redundancy gain on mean RT was reliably larger in Experiment 1B (128 ms) than in Experiment 5 (81 ms), $t(66) = 3.57$, $p < .001$, $d = 0.87$, providing converging evidence that the guidance benefits generated by redundant template-based and saliency-driven processes are mediated by spatial location.

Discussion

In Experiment 5, no violations of the RMI were observed when the redundant target features were associated with different items at different locations. That is, coactivation was eliminated under conditions where a mechanistic account of signal aggregation (location-based) predicted it would be eliminated. Experiment 5 also converges with the results of Experiment 3 to isolate the source of coactivation. The mere presence of two target-defining signals is not sufficient for coactivation; no evidence of coactivation is observed when the redundant signals fail to converge on the same location within the combined priority map (see also Mordkoff & Danek, 2011). In sum, signals at different locations appear to implement a separate-activations mechanism: They race toward an orienting solution, producing a modest redundancy gain based on the statistical facilitation of RT (Raab, 1962) but without producing violations of the RMI. In contrast, redundant signals at the same location coactivate, producing robust violations of the RMI.

Figure 5
Results From Experiment 5



Note. Exp. = Experiment; Prob. = probability; RT = response time; RMI = race model inequality. See the online article for the color version of this figure.

General Discussion

The present results provide strong support for models of attention control holding that saliency-driven and template-based sources of guidance converge, simultaneously, to jointly influence activity on a common, spatially organized priority map. For target stimuli that both (a) matched a search template feature value and (b) were physically salient, the speed of target detection and orienting was faster than could have been explained by the faster of the two individual guidance processes operating in parallel (i.e., violation of the RMI). Violations of the RMI eliminate all noncoactive models, including models holding that the architecture consists of independent template-based and saliency-driven orienting mechanisms racing to a single solution (separate-activations model) and models holding that the sequential application of the two sources of guidance functionally precludes coactivation (sequential model). Moreover, the absence of RMI violations when redundant target properties were associated with different locations indicates that signal aggregation occurs on a spatially organized priority map.

How might the two sources of guidance be combined to generate an integrated priority signal? One possibility is that search templates dynamically filter new sensory processing to enhance the signal associated with template-matching stimuli (Desimone & Duncan, 1995; Wolfe, 1994, 2021). In this view, integration occurs at a perceptual level, before perceptual signals are passed to a higher-level priority map guiding attention. This is consistent with broad evidence indicating that specific template values stored in visual working memory can modulate perceptual processing (Pessoa et al., 2003; Serences et al., 2009; Shulman et al., 1997; Teng & Kravitz, 2019). Alternatively, template-based and saliency-driven biases may proceed relatively independently and be subserved by separate neural pathways (Corbetta & Shulman, 2002), with only the outputs of the two systems integrated on the priority map itself. The range of possible integration mechanisms is expanded when one considers spatial biases. Template-based spatial biases (e.g., knowledge of probable target location) could interact with relatively early sensory processing (e.g., Moran & Desimone, 1985), but spatial biases could also influence activity on a priority map directly, given that the map is spatially organized. This preshaping of priority map activity (e.g., Duncan et al., 2023) could then be integrated with differential salience arriving from perceptual processing of the scene. Because higher-level neural priority maps do not necessarily code surface feature values, this type of direct priority map modulation would not necessarily be available for surface feature dimensions, as tested here.

Thus far, we have been interpreting violations of the RMI as evidence of coactivation, such that the results of Experiments 1–4 provide support for summation on a common priority map. However, the Interactive Race Model of simple target detection (Mordkoff & Yantis, 1991), which does not involve coactivation in the usual sense of the term, is capable of producing data that violate the RMI under certain conditions. Specifically, a separate-activations model can sometimes violate the RMI when the occurrence of targets is positively correlated, such that the presence of one target is statistical evidence of the presence of the other (see also Mordkoff & Egeth, 1993). These have been referred to as "biased contingencies." We do not believe that this alternative explanation applies to the current results for two reasons. First, in Experiments 1, 2, 4, and 5, the presence of targets was uncorrelated. For example, the overall probability of a color target and the conditional probability of a color target given the presence of a shape target were equal (both being 0.50) in each of these experiments. The same is true for the shape targets. Second, in all five experiments, the specific values of color and shape that were used as the targets for a given trial were selected randomly. Therefore, even a more sophisticated version of the Interactive Race Model that could track the conditional probabilities for each feature value would have no consistent relationships to learn.

Finally, we found no evidence in the present study to suggest that there was a period following stimulus onset during which only saliency information was functional in guiding attention, contrary to the most prominent version of the sequential model (Theeuwes, 2010;

van Heusden et al., 2022). Granted, we tested one type of physical salience and one type of template match. However, the present stimuli and task were modeled on previous studies that have examined attention capture (e.g., Egeth et al., 1984), and thus the current conditions were not atypical of those in which claims of sequential guidance have been made. More specifically, detection of color singletons, particularly when the color difference was 180° in Experiment 2, should have been highly efficient (Treisman & Gelade, 1980). Despite saliency-driven guidance being given ample opportunity to precede template-based guidance, robust violations of the RMI were observed, RT distributions for the two sources of guidance were strongly overlapping, and redundancy gains were observed relative to the faster of the two individual guidance processes, all of which indicates that there was no initial period of attention guidance driven solely by physical saliency.

Could a modified version of the sequential model account for our results by relaxing the assumption that the first guidance process is completed before the second begins? That is, could one observe the present results under a model in which saliency-driven and template-based processes have staggered initiation times but overlap in the time course of implementation? There are both empirical and theoretical reasons to reject such an alternative. Empirically, there was no evidence in the present experiments to indicate different initiation times for the two sources of guidance. In all five experiments, the accumulation of positive probabilities for the two single-target conditions began at almost precisely the same time (see raw CDFs in Figures 3 and 4). This was true even in Experiment 4 (label cue version, Figure 4C), where there was a substantial mean RT advantage for the saliency-driven condition compared with the template-based condition. This mean advantage was not caused by a difference in initiation time, but rather by a difference in the rate of accumulation of responses, and early violations of the RMI in Experiment 4 still indicated that there was no initial period of guidance governed solely by physical saliency. Theoretically, it is possible that two staggered guidance processes could produce violations of the RMI if the first was not sufficient to generate an orienting decision before the second began. However, such a modification, in addition to being inconsistent with the present empirical results, would no longer be consistent with the claims of published models, which hold that saliency-driven processes are sufficient to direct attention before template-based processes are applied.

The results of the present experiments contrast with other studies providing evidence of sequential guidance under some experimental circumstances (van Heusden et al., 2022; van Zoest et al., 2004). What might account for the difference? The studies finding sequential guidance have tended to use displays in which the optimal strategy was likely to first filter the display based on physical salience, reducing the effective set size, with template information applied only to this subset of salient items. For example, in van Heusden et al. (2022), displays consisted of a 9 × 9 grid of 81 oriented stimuli. Seventy-nine items had the same orientation, and two salient items had discrepant orientations (one tilted left and one tilted right). The task was to fixate the left-tilted item. In such displays, a sensible strategy would be to first use physical salience to isolate the two potential target candidates and only then apply template-based guidance to this subset of items to discriminate the target from the salient distractor. In contrast, in the present task, participants had an incentive to implement template-based guidance as early as possible and simultaneously with saliency-driven guidance, as some target-present trials were defined solely by the presence of a template-matching item. Another possible difference between the two studies is the selection task: detection versus oculomotor orienting. It is possible that template-based guidance is specifically delayed, relative to saliency-driven guidance, for the selection of saccade targets. However, we think this is unlikely given evidence that the content of visual working memory (which implements template-based modulation) can influence target selection for saccades generated in as little as 120 ms after stimulus onset (Hollingworth et al., 2013a, 2013b). In sum, although there may be certain tasks in which guidance becomes effectively sequential, the present data demonstrate that this is not a general property of attentional control. Template-based guidance can be applied simultaneously with the registration of physical salience, with the two sources of guidance combined in an integrated priority signal.

References

Adam, K. C. S., & Serences, J. T. (2021). History modulates early sensory processing of salient distractors. *Journal of Neuroscience*, *41*(38), 8007–8022. https://doi.org/10.1523/JNEUROSCI.3099-20.2021

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, *55*(5), 485–496. https://doi.org/10.3758/bf03205306

Bahle, B., Mordkoff, J. T., Niu, Z., & Hollingworth, A. (2024). *Attentional control from saliency signals on multiple perceptual dimensions* [Unpublished manuscript]. Department of Psychological and Brain Sciences, University of Iowa.

Bahle, B., Thayer, D. D., Mordkoff, J. T., & Hollingworth, A. (2020). The architecture of working memory: Features from multiple remembered objects produce parallel, coactive guidance of attention in visual search. *Journal of Experimental Psychology: General*, *149*(5), 967–983. https://doi.org/10.1037/xge0000694

Becker, S. I., Grubert, A., Horstmann, G., & Ansorge, U. (2023). Which processes dominate visual search: Bottom-up feature contrast, top-down tuning or trial history? *Cognition*, *236*(7), Article 105420. https://doi.org/10.1016/j.cognition.2023.105420

Berggren, N., Nako, R., & Eimer, M. (2020). Out with the old: New target templates impair the guidance of visual search by preexisting task goals. *Journal of Experimental Psychology: General*, *149*(6), 1156–1168. https://doi.org/10.1037/xge0000697

Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*, *33*(1), 1–21. https://doi.org/10.1146/annurev-neuro-060909-152823

Bravo, M. J., & Nakayama, K. (1992). The role of attention in different visual-search tasks. *Perception & Psychophysics*, *51*(5), 465–472. https://doi.org/10.3758/BF03211642

Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional templates in visual working memory. *Journal of Neuroscience*, *31*(25), 9315–9322. https://doi.org/10.1523/JNEUROSCI.1097-11.2011

Ciaramelli, E., Grady, C., Levine, B., Ween, J., & Moscovitch, M. (2010). Top-down and bottom-up attention to memory are dissociated in posterior parietal cortex: Neuroimaging and neuropsychological evidence. *The Journal of Neuroscience*, *30*(14), 4943–4956. https://doi.org/10.1523/JNEUROSCI.1209-09.2010

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Earlbaum Associates.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology*, *2*, Article 233. https://doi.org/10.3389/fpsyg.2011.00233

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Duncan, D. H., van Moorselaar, D., & Theeuwes, J. (2023). Pinging the brain to reveal the hidden attentional priority map using encephalography. *Nature Communications*, *14*(1), Article 4749. https://doi.org/10.1038/s41467-023-40405-8

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(1), 32–39. https://doi.org/10.1037/0096-1523.10.1.32

Eriksen, C. W. (1988). A source of error in attempts to distinguish coactivation from separate activation in the perception of redundant targets. *Perception & Psychophysics*, *44*(2), 191–193. https://doi.org/10.3758/bf03208712

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences*, *10*(8), 382–390. https://doi.org/10.1016/j.tics.2006.06.011

Franconeri, S. L., Simons, D. J., & Junge, J. A. (2004). Searching for stimulus-driven shifts of attention. *Psychonomic Bulletin & Review*, *11*(5), 876–881. https://doi.org/10.3758/BF03196715

Hahn, B., Ross, T. J., & Stein, E. A. (2006). Neuroanatomical dissociation between bottom-up and top-down processes of visuospatial selective attention. *Neuroimage*, *32*(2), 842–853. https://doi.org/10.1016/j.neuroimage.2006.04.177

Hamker, F. H. (2005). The reentry hypothesis: The putative interaction of the frontal eye field, ventrolateral prefrontal cortex, and areas V4, IT for attention and eye movement. *Cerebral Cortex*, *15*(4), 431–447. https://doi.org/10.1093/cercor/bhh146

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(1), 113–136. https://doi.org/10.1037//0096-1523.28.1.113

Hollingworth, A., Matsukura, M., & Luck, S. J. (2013a). Visual working memory modulates low-level saccade target selection: Evidence from rapidly generated saccades in the global effect paradigm. *Journal of Vision*, *13*(13), Article 4. https://doi.org/10.1167/13.13.4

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

Johansson, R. S., Westling, G. R., Backstrom, A., & Flanagan, J. R. (2001). Eye-hand coordination in object manipulation. *Journal of Neuroscience*, *21*(17), 6917–6932. https://doi.org/10.1523/JNEUROSCI.21-17-06917.2001

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, *43*(4), 346–354. https://doi.org/10.3758/BF03208805

Krummenacher, J., Müller, H. J., & Heller, D. (2001). Visual search for dimensionally redundant pop-out targets: Evidence for parallel-coactive processing of dimensions. *Perception & Psychophysics*, *63*(5), 901–917. https://doi.org/10.3758/bf03194446

Krummenacher, J., Müller, H. J., & Heller, D. (2002). Visual search for dimensionally redundant pop-out targets: Parallel-coactive processing of dimensions is location specific. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(6), 1303–1322. https://doi.org/10.1037//0096-1523.28.6.1303

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): Still no identification without attention. *Psychological Review*, *111*(4), 880–913. https://doi.org/10.1037/0033-295X.111.4.880

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*(25–26), 3559–3565. https://doi.org/10.1016/S0042-6989(01)00102-X

Li, Z. P. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, *6*(1), 9–16. https://doi.org/10.1016/S1364-6613(00)01817-9

Liesefeld, H. R., Lamy, D., Gaspelin, N., Geng, J. J., Kerzel, D., Schall, J. D., Allen, H. A., Anderson, B. A., Boettcher, S., Busch, N. A., Carlisle, N. B., Colonius, H., Draschkow, D., Egeth, H., Leber, A. B., Müller, H. J., Röer, J. P., Schubö, A., Slagter, H. A., … Wolfe, J. (2024). Terms of debate: Consensus definitions to guide the scientific discourse on visual distraction. *Attention, Perception, & Psychophysics*, *86*(5), 1445–1472. https://doi.org/10.3758/s13414-023-02820-3

Liesefeld, H. R., Liesefeld, A. M., Müller, H. J., & Rangelov, D. (2017). Saliency maps for finding changes in visual scenes? *Attention, Perception, & Psychophysics*, *79*(7), 2190–2201. https://doi.org/10.3758/s13414-017-1383-9

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

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

Mazer, J. A., & Gallant, J. L. (2003). Goal-related activity in V4 during free viewing visual search: Evidence for a ventral stream visual salience map. *Neuron*, *40*(6), 1241–1250. https://doi.org/10.1016/S0896-6273(03)00764-5

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*(2), 247–279. https://doi.org/10.1016/0010-0285(82)90010-X

Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics*, *40*(5), 331–343. https://doi.org/10.3758/bf03203025

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*(4715), 782–784. https://doi.org/10.1126/science.4023713

Mordkoff, J. T., & Danek, R. H. (2011). Dividing attention between color and shape revisited: Redundant targets coactivate only when parts of the same perceptual object. *Attention, Perception, & Psychophysics*, *73*(1), 103–112. https://doi.org/10.3758/s13414-010-0025-2

Mordkoff, J. T., & Egeth, H. E. (1993). Response time and accuracy revisited: Converging support for the interactive race model. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(5), 981–991. https://doi.org/10.1037/0096-1523.19.5.981

Mordkoff, J. T., & Miller, J. (1993). Redundancy gains and coactivation with two different targets: The problem of target preferences and the effects of display frequency. *Perception & Psychophysics*, *53*(5), 527–535. https://doi.org/10.3758/BF03205201

Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 520–538. https://doi.org/10.1037/0096-1523.17.2.520

Most, S. B., Scholl, B. J., Clifford, E. R., & Simons, D. J. (2005). What you see is what you set: Sustained inattentional blindness and the capture of awareness. *Psychological Review*, *112*(1), 217–242. https://doi.org/10.1037/0033-295X.112.1.217

Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research*, *45*(2), 205–231. https://doi.org/10.1016/j.visres.2004.07.042

Pessoa, L., Kastner, S., & Ungerleider, L. G. (2003). Neuroimaging studies of attention: From modulation of sensory processing to top-down control. *The Journal of Neuroscience*, *23*(10), 3990–3998. https://doi.org/10.1523/JNEUROSCI.23-10-03990.2003

Pinto, Y., van der Leij, A. R., Sligte, I. G., Lamme, V. A., & Scholte, H. S. (2013). Bottom-up and top-down attention are independent. *Journal of Vision*, *13*(3), Article 16. https://doi.org/10.1167/13.3.16

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, *24*(5 Series II), 574–590. https://doi.org/10.1111/j.2164-0947.1962.tb01433.x

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, *20*(2), 207–214. https://doi.org/10.1111/j.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

Shulman, G. L., Corbetta, M., Buckner, R. L., Raichle, M. E., Fiez, J. A., Miezin, F. M., & Petersen, S. E. (1997). Top-down modulation of early sensory cortex. *Cerebral Cortex*, *7*(3), 193–206. https://doi.org/10.1093/cercor/7.3.193

Stilwell, B. T., Adams, O. J., Egeth, H. E., & Gaspelin, N. (2023). The role of salience in the suppression of distracting stimuli. *Psychonomic Bulletin & Review*, *30*(6), 2262–2271. https://doi.org/10.3758/s13423-023-02302-5

Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. *Nature Human Behaviour*, *3*(8), 827–836. https://doi.org/10.1038/s41562-019-0640-4

Thayer, D. D., & Sprague, T. C. (2023). Feature-specific salience maps in human cortex. *Journal of Neuroscience*, *43*(50), 8785–8800. https://doi.org/10.1523/JNEUROSCI.1104-23.2023

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606. https://doi.org/10.3758/BF03211656

Theeuwes, J. (1993). Visual selective attention: A theoretical analysis. *Acta Psychologica*, *83*(2), 93–154. https://doi.org/10.1016/0001-6918(93)90042-P

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006

Thompson, K. G., & Bichot, N. P. (2005). A visual salience map in the primate frontal eye field. *Progress in Brain Research*, *147*, 249–262. https://doi.org/10.1016/s0079-6123(04)47019-8

Thompson, K. G., Bichot, N. P., & Sato, T. R. (2005). Frontal eye field activity before visual search errors reveals the integration of bottom-up and top-down salience. *Journal of Neurophysiology*, *93*(1), 337–351. https://doi.org/10.1152/jn.00330.2004

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*(4), 766–786. https://doi.org/10.1037/0033-295X.113.4.766

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

van Heusden, E., van Zoest, W., Donk, M., & Olivers, C. N. L. (2022). An attentional limbo: Saccades become momentarily non-selective in between saliency-driven and relevance-driven selection. *Psychonomic Bulletin & Review*, *29*(4), 1327–1337. https://doi.org/10.3758/s13423-022-02091-3

van Zoest, W., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 746–759. https://doi.org/10.1037/0096-1523.30.4.749

Vickery, T. J., King, L. W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, *5*(1), Article 8. https://doi.org/10.1167/5.1.8

Wolfe, J. M. (1994). Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. https://doi.org/10.3758/BF03200774

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*(4), 1060–1092. https://doi.org/10.3758/s13423-020-01859-9

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419–433. https://doi.org/10.1037/0096-1523.15.3.419

Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*, *44*(12), 1411–1426. https://doi.org/10.1016/j.visres.2003.11.024

Yang, H., & Zelinsky, G. J. (2009). Visual search is guided to categorically-defined targets. *Vision Research*, *49*(16), 2095–2103. https://doi.org/10.1016/j.visres.2009.05.017

Zehetleitner, M., Krummenacher, J., & Müller, H. J. (2009). The detection of feature singletons defined in two dimensions is based on salience summation, rather than on serial exhaustive or interactive race architectures. *Attention, Perception, & Psychophysics*, *71*(8), 1739–1759. https://doi.org/10.3758/app.71.8.1739

Received July 24, 2024
Revision received November 13, 2024
Accepted November 19, 2024 ■