Trends in
Cognitive Sciences

CellPress
OPEN ACCESS

Review

Good-enough attentional guidance

Xinger Yu,[1,2] Zhiheng Zhou,[1] Stefanie I. Becker,[3] Sage E.P. Boettcher,[4] and Joy J. Geng[1,2,*]

Theories of attention posit that attentional guidance operates on information held in a target template within memory. The template is often thought to contain veridical target features, akin to a photograph, and to guide attention to objects that match the exact target features. However, recent evidence suggests that attentional guidance is highly flexible and often guided by non-veridical features, a subset of features, or only associated features. We integrate these findings and propose that attentional guidance maximizes search efficiency based on a 'good-enough' principle to rapidly localize candidate target objects. Candidates are then serially interrogated to make target-match decisions using more precise information. We suggest that good-enough guidance optimizes the speed–accuracy–effort trade-offs inherent in each stage of visual search.

Highlights

Natural behavior frequently involves the need to look for task-relevant information. The function of attention is to facilitate this process by enhancing sensory signals from objects that match our goals. Objects that are selected are more likely to enter into conscious awareness, be remembered, and be acted upon.

We propose that the information used to enhance sensory signals is not based on a veridical replica of the target in memory but instead on a good-enough principle. Good-enough information is ideal for attentional guidance because it is sufficient to localize likely potential candidate targets rapidly, allowing more precise subsequent decision processes to determine accuracy.

Good-enough guidance is inherently context-dependent and can involve the use of non-veridical target features, a subset of target features, or even only features or objects that are associated with the target.

Attentional guidance during visual search

We constantly look for people and objects, whether it be a friend in a crowd, our kids after school, keys in a cluttered room, or a list of ingredients at the grocery store. The ability to find things efficiently depends on attentional mechanisms that selectively guide attention toward potential targets (**attentional guidance**; see Glossary) and then make a **target-match decision**. The common act of looking for something has been studied in the laboratory using the **visual search** paradigm [1]. Visual search paradigms nicely capture the recursive demands of attention as they unfold over time: the maintenance of target information in memory, the use of target information to guide attention and eye movement towards a candidate target object, and the evaluation of the identity of the candidate object as a target-match based on higher-acuity **foveal vision** (Figure 1). If the object is not the desired target, the attend/look–identify cycle begins again. If it is the target, search is completed. Visual search paradigms are a test-bed for understanding what information is used to guide attention, and what information is subsequently used to identify whether the object matches the target [2–7].

Models of attention use the construct of the attentional or **target template** [8] to characterize the information held in working or long-term memory about the target. Although the target template is a basic concept in all models of attention, there are still open questions regarding its definition (Box 1). However, it is agreed that the features in the target template are used to adjust sensory gain and set processing priority to enhance signals from objects that match the target (Box 2). The object with the highest attentional priority wins the competition for attention and is examined in more detail by covertly attending to the object or overtly generating a saccadic eye movement (**saccade**) to determine if the identity matches the target. This attend/look–identify cycle continues until the target is found [4]. As such, since its inception more than 30 years ago, the target template has played a crucial role in models of visual search. However, our understanding remains incomplete. Nevertheless, models mostly characterize the template as a complete and veridical representation of the target. These models of visual search posit that the likelihood of attention shifting to an object depends on how well that object matches the template [1,8–10]. Consistent with this, target templates with higher precision result in faster visual search times [2,11–14].

[1]Center for Mind and Brain, University of California Davis, Davis, CA, USA
[2]Department of Psychology, University of California Davis, Davis, CA, USA
[3]School of Psychology, University of Queensland, Brisbane, QLD, Australia
[4]Department of Experimental Psychology, Oxford University, Oxford, UK

*Correspondence:
jgeng@ucdavis.edu (J.J. Geng).

Thus, even when only a subset of feature information is used for guidance, one would expect the features to be veridical in the target template to maximize search efficiency. However, work over the past decade has begun to suggest that the information used to guide attention is not always based on all veridical features of the target but is based on target information that can be rapidly processed and reliably distinguished from the distractor context [15–17]. This work builds on findings that some features are more effective in guiding attention than others [3,18–20] by showing that, even when using 'undoubted guiding attributes', the features used may be imprecise and approximate, only a partial match to the target, or even only probabilistically associated with the target. The results underpin the importance of differentiating between the representation of object information in the target template in memory and the use of that information to guide attention through modulations of sensory gain [9,21–26]. Although the capacity of the target template may be unlimited [9], the precision and capacity of guidance are more constrained owing to a host of factors including limitations in the number of active sensory representations, acuity limitations in **peripheral vision**, variable visual factors such as occlusion, feature uncertainty, differing viewpoints, and extra-target contextual factors such as distractors, clutter, and lighting or visibility [27–31].

More importantly, these results are consistent with a theory of good-enough guidance and suggest that the principles that dictate attentional guidance are not based on a complete, veridical, target template in memory but are based on a subset of information that rapidly guides attention toward high-probability candidate target objects but without the need to be entirely accurate. This is not to say that guidance cannot be precise but that there are often good reasons for it to not be precise. This concept is reminiscent of other theories in the field of perception, including the study of eye movements and the fields of language comprehension and production [32–34]. We suggest that the use of good-enough template information serves the goal of attentional guidance – to rapidly localize probable target objects that can then be serially inspected to make a more precise identity-based target-match decision (Figure 1). Such a mechanism for guidance satisfies the balance between being selective and being efficient. Using good-enough information localizes probable targets without over-emphasizing accuracy given the low cost of inaccurate shifts of attention and/or eye movements.

The idea of good-enough guidance should be distinguished from inherent limitations in the visual system that determines what information is 'as good as it gets'. Most notably, limits in the spatial and color acuity in peripheral vision will always serve as an upper limit in the precision of the

Glossary

Attentional guidance: the process of shifting covert or overt attention to selectively process information based on priorities set by the task context.

Foveal vision: visual processing in a small region around the gaze position that has maximal acuity and contrast sensitivity. Synonymous with central vision.

Peripheral vision: visual processing outside foveal vision, encompassing a larger field of view with lower visual acuity and contrast sensitivity. Visual acuity in peripheral vision varies by visual eccentricity and the visual context, such as crowding.

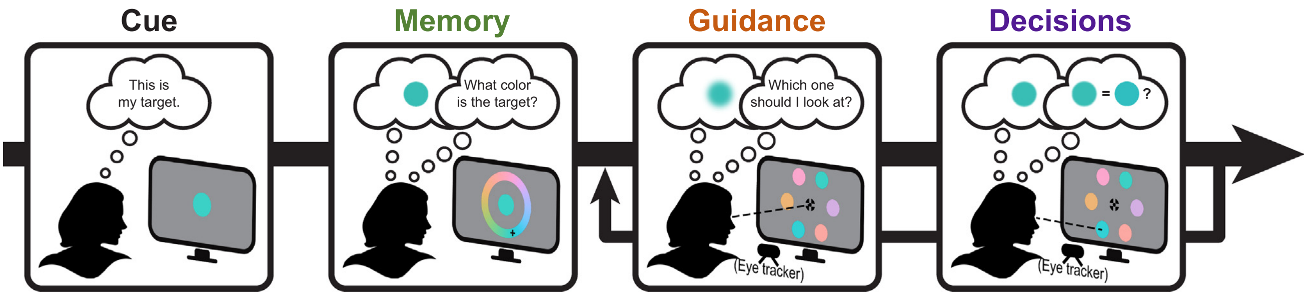
Saccade: a rapid eye movement that shifts the center of gaze from one location on the visual field to another.

Speed–accuracy trade-off: a phenomenon relating the speed of a response to its accuracy. The more time an individual takes to make a response, the greater the accuracy of the response.

Target-match decisions: the evaluation of the identity of an object as a target-match or a target-mismatch.

Target template: the representation of target information maintained in working or long-term memory during visual search; synonymous with attentional template, search template, mental template.

Visual search: the process of looking for a particular or multiple objects (i.e., targets) among other objects (i.e., distractors).



Trends in Cognitive Sciences

Figure 1. Illustration of a paradigm that measures three stages of visual search. The appearance of the cue display sets the target template in memory; periodic memory probe displays (occurring on only a subset of trials) measure the target features in memory. Attentional guidance begins at the onset of the search display based on the prioritization of good-enough target features to select a candidate object. The target-match decision is made on the selected item. If the selected item is identified as a non-target, the attend/look–identify cycle begins again. If the selected item is identified as the target, then the trial is terminated. Note that the thought bubbles are caricatures of the cognitive processes that the participant is hypothesized to engage in; they are not necessarily conscious, verbalizable, thoughts. The guidance and decision cycle could involve eye movements, as illustrated, but could also unfold covertly.

Box 1. Target templates

The template in models of attention is widely understood to be the internal representation of the target object during visual search that is used to facilitate the selection of task-relevant sensory information. For this reason, it is often referred to as the 'target template', although 'attentional template' and 'mental template' are synonymous terms. Despite the important role that templates hold in models of attention, the exact nature of target templates remains poorly understood. The template has been characterized as being held in working memory, but there is reason to believe that long-term memory is involved and may even be the more stable source of target information over time. For example, target information is transferred from working to long-term memory with learning over short time-periods [125], and long-term memory does not suffer from capacity limits, which explains how dozens of targets or a wider network of target-related information can be used simultaneously during visual search [126]. A crucial question to answer concerns how long-term and working memory representations interact during visual search (e.g., during hybrid search in which the number of relevant targets exceeds working memory capacity). For example, perhaps information held in working memory is more flexible and already reflects the most useful subset of information from the more complete target template that can be further weighted by good-enough principles for attentional guidance [27,62]. There is some evidence that working memory representations 'divide' labor by providing top-down guiding information. For example, different regions of lateral frontal cortex are essential for the selection of features and space [25,26,127]. Thus, working memory representations may be *ad hoc* representations of the target that interact with more complete longer-term representations held elsewhere. This is consistent with the Guided Search 6.0 architecture in which the target template is a representation in long-term memory that is used to set the parameters for guidance and decision processes based on more temporary working memory representations [9].

A related issue regarding the characterization of target templates is the question of what can be considered to be a template. Guided Search 6.0 introduces the idea of two templates: a 'target template' in long-term memory and a second 'guiding template' in working memory. This suggests the existence of two separate representations in long-term and working memory that serve two separate functions. However, it is unclear whether these two representations are fully independent or are essentially yoked within a cognitive and neural network that also includes additional target representations (e.g., in visual cortex) (Box 2). For example, it may be that target information held for longer than a few seconds automatically becomes encoded in long-term memory [125]; conversely, features from the target template in long-term memory determine what information about a target is available to compute guidance and identity decisions. Functionally, the characterizations of 'two templates' versus 'one template whose information is read out for many uses' are equivalent: both allow different versions of target representations to exist at the same time, most likely in different brain regions. In fact, target information can be found throughout the brain [128], and it is therefore important to identify what information is used for what computations [129] rather than debating the number of representations that can be measured. A related debate has arisen in the working memory literature over how/where working memory is encoded [130–133]. Given this, debating 'one or multiple' templates may be a red herring in the development of attentional models – both descriptions can map onto the same system and mechanisms.

Good-enough guidance is agnostic regarding the debate about 'one or more templates', but builds on Guided Search 6.0 by suggesting that the information used for guidance and identity decisions may differ from each other in content and precision so as to satisfy different computational goals, but they are both derived from the target template held in memory. Thus, the upper limit of guidance will always be defined in relation to the precision of the target template in memory [61]. This view is consistent with findings that target representations are pervasive throughout the brain but differ in functional significance [128,129].

information available to guide attention [29,35–38]. However, we argue that what is 'good-enough' is frequently coarser or has fewer feature dimensions than the strict limits of vision [29]. Instead, what is good-enough for guidance is mainly determined by the simplest discriminative information necessary to select a probable target. The exact **speed–accuracy trade-off** will also be dependent on the current willingness of the observer to exert sufficient effort to guide attention to the best candidate target versus merely one of the best.

Good-enough guidance uses fuzzy feature information

Good-enough guidance predicts that the precision of attentional guidance will be lower (or equal, but never higher) than what is available in the target template in memory. For example, recent studies show that attention is guided to a broader range of non-target distractors than are mistaken for actual targets (Figure 2, rows 1–2). Target-similar distractors that have a high likelihood of being covertly attended are rarely mistaken for the actual target, suggesting that guidance selects candidate targets with features that are easily rejected as non-targets [16,39].

Although guidance will sometimes select the wrong object because its features cannot be discriminated from the target based on limits in perceptual acuity in the visual periphery [35–37,40,41], peripheral vision appears to be sufficient to discriminate simple features and even more complex object and scene information [29,37,42,43]. Furthermore, the precision of attentional guidance and first saccades to the target improves when there are incentives for emphasizing accuracy over speed [44–48]. For example, the accuracy of eye movements to the target increases when observers delay the saccade and increase the time to inspect search environments using peripheral vision [47,49,50]. Likewise, later saccades are more selective for target-similar stimuli, suggesting that the precision of guidance improves as information is accumulated over time within a single display during search [51,52]. Measurements of guidance should therefore become closer to that of memory when there is more time to accumulate evidence about item features (e.g., when accuracy should be prioritized over speed). However such deliberate saccade planning appears to be cognitively effortful [45,46,48,53–58]. This suggests that fuzzy guidance reflects not only visual limits but also trade-offs between rapidly localizing candidate targets (e.g., with a saccade) and localizing the correct target within the task context.

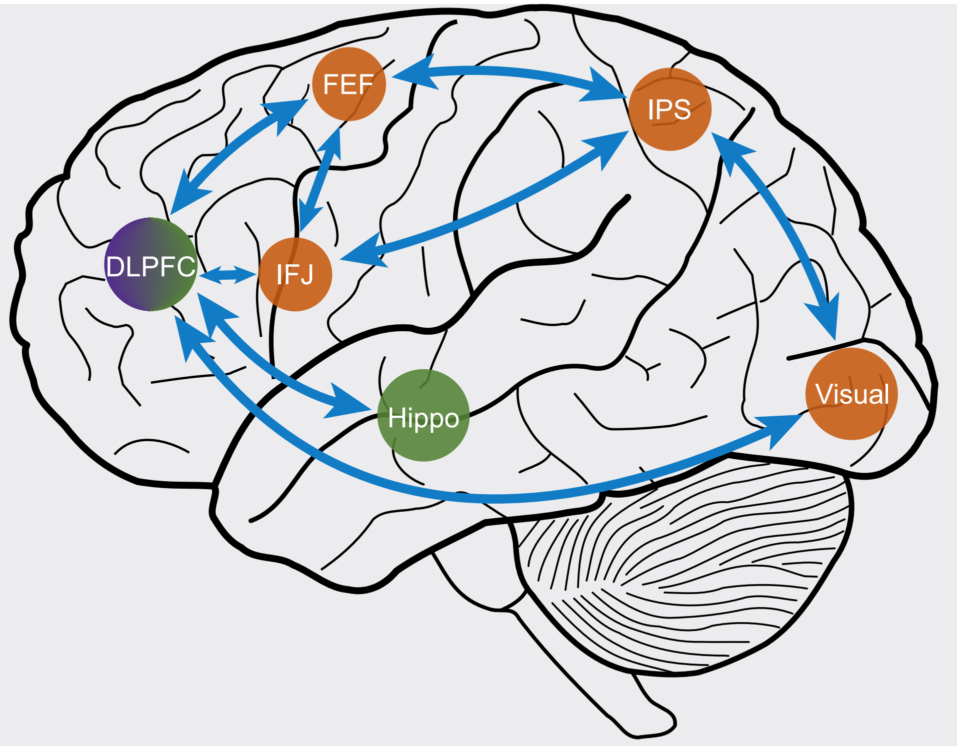
These data suggest that good-enough guidance not only occurs because of fixed limits in visual processing but also because guidance is optimized to be fast, even when inaccurate: errors typically result in low costs because post-selection decision processes can quickly identify them and trigger a new saccade to a different potential target [55,59,60]. Interestingly, although the measured precision of target decisions is greater than guidance, memory representations

Box 2. Brain networks for attentional guidance

Attentional guidance relies on a network of regions that have frequently been described in terms of 'sources' and 'sites'. Source regions are domain-general control regions that encode information about what is expected to be relevant based on the target or task; site regions have specific sensory response profiles that are enhanced or suppressed based on the current task context. The specific network involved depends in part on the type of information attended. For example, during spatial attention, source signals from the frontal eye-fields (FEF) and the intraparietal sulcus (IPS) modulate visual processing in spatially corresponding retinotopic locations [134,135]. The inferior frontal junction (IFJ; also referred to as the ventral prearcuate regions, VPA, in non-human primates) encodes target features (e.g., color, shapes, faces, houses) that modulate sensory neurons with corresponding receptive field properties [25,26]. FEF/LIP and IFJ/VPA appear to encode what is relevant at the current moment of time, but other regions including the dorsolateral prefrontal cortex (DLPFC) and hippocampus (Hippo) appear to encode the prospective memory of what will be relevant in the future and over time [136–139] at different levels of category abstractness [140]. This suggests that attentional control networks operate at different timescales to prioritize prospective information in anticipation of what is to come, sensory processing at the current moment of time, and retrospective information to update long-term statistical representations of the target (Figure I).

The effect of modulation at sites of sensory processing occurs as early as the lateral geniculate nucleus (LGN) and primary visual cortex (V1), and extends throughout the visual hierarchy [141–143]. Virtually all neurons in the visual system ranging from V1 simple cells to high-level cortical regions for nonspecific objects, faces, and scenes are robustly modulated by attentional settings [27,144–146]. Many mechanisms have been proposed to describe how sensory processing is modulated at sites, including shifts in contrast gain, response gain, and shifting or sharpening of neuronal turning [147]. What is common across these mechanisms is that task-relevant visual information is enhanced, and irrelevant information is suppressed. Although the characterization of sensory regions as sites might seem to suggest that they are the passive recipients of feedback modulation, this would be an error because there is plenty of evidence that local dynamics in sensory cortex contribute to the maintenance and updating of attentional priority [148,149]. It remains an open question how the dynamics of long-term and working memory maintenance interact with those of sensory modulation over time to maintain and perform task goals.

We hypothesize that good-enough guidance is based on this existing architecture for attentional control and can be predicted by a calculation of what template information will rapidly localize candidate targets and segment out non-targets in the current context. Moreover, the information used for guidance should always be constrained by the organization of feature representations within the visual system. That is, what is good-enough will always be an object type or feature that can be modulated by attention (color, orientation, spatial frequency, motion, objects [150–154]).

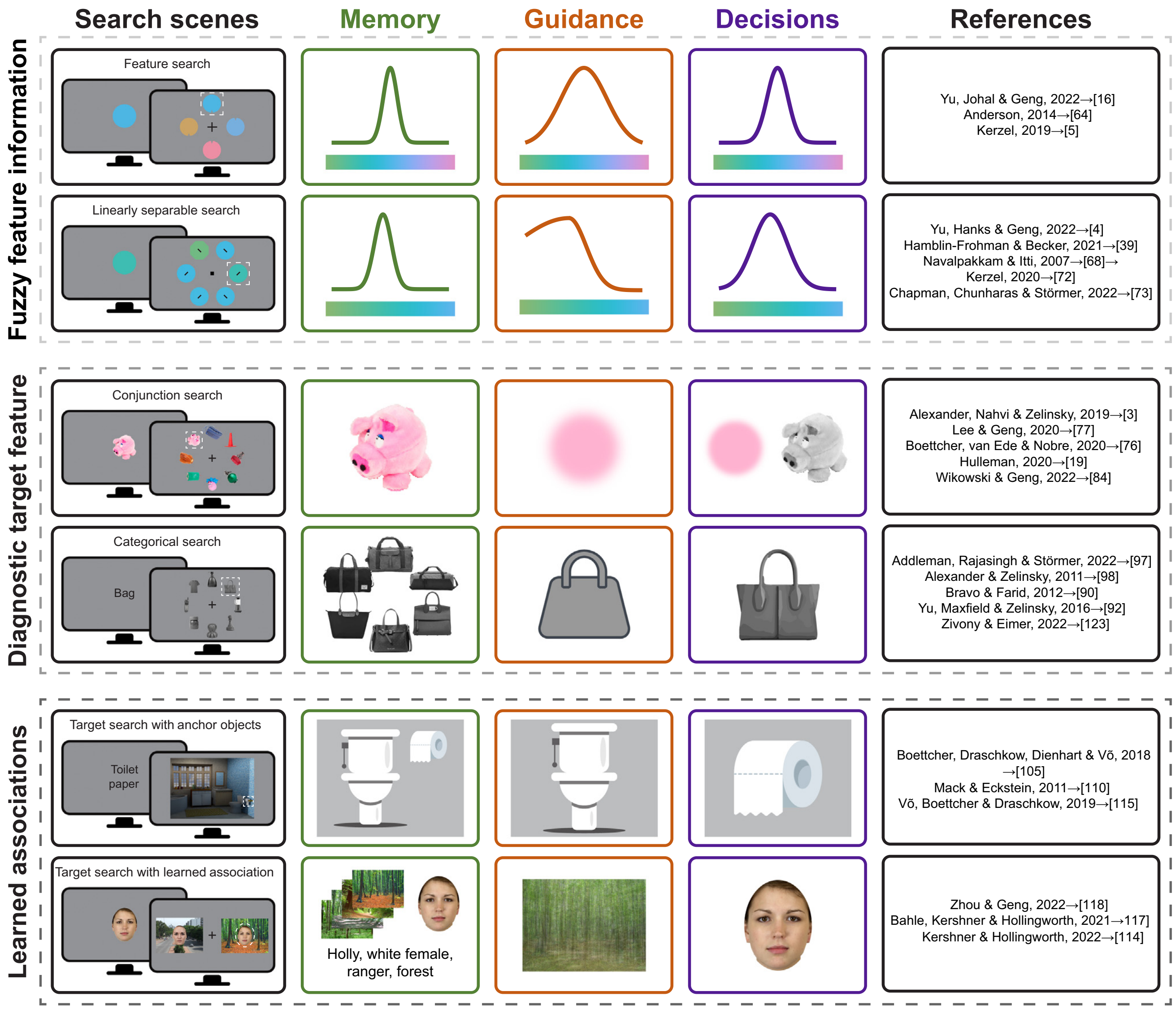


Trends in Cognitive Sciences

Figure I. Illustration of hypothesized brain networks contributing to memory, guidance, and decisions during visual search. Memory in green, guidance in orange, decisions in purple. Abbreviations: DLPFC, dorsolateral frontal cortex; FEF, frontal eye fields; Hippo, hippocampus; IFJ, inferior frontal junction; IPS, intraparietal sulcus.

appear even more precise. This leads to the hypothesis that the target template in memory is the most precise and complete representation of the target, whereas guidance and target-match decisions are based on only a subset of the target template. The target template in memory may itself vary in precision [61,62], but we hypothesize that the precision of memory will be the upper limit of what is possible for guidance, even though guidance will typically appear less precise. The precision of the information used for guidance and decisions will be determined by what is necessary to accomplish the task demands at the current moment in time rather than what is available or possible [5,12,16,63,64]. To distinguish the precision of information used at each stage of processing (memory, guidance, decisions), it is necessary to measure each separately within the same task.

The trade-off in precision between guidance and decisions can also help to resolve a debate that has arisen in studies of off-veridical target representations. Off-veridical target representations refer to the finding that target features are shifted, or distorted, away from distractors that are predictably linearly separable [65–68]. For example, looking for an orange target among yellower distractors results in attentional capture by objects that are 'redder' than the true target. Similar effects have been reported for targets defined by shape, size, line orientation, direction of motion, and even for facial expression search [4,60,69–73]. The size of target shifting within visual search contexts is greater than that of simultaneous stimulus contrast alone [73,74]. Moreover, when shifting offers no advantage for target selection, it is vastly reduced [75]. These results suggest that the shifting of the target representation is a task-adaptive mechanism to increase the



Trends in Cognitive Sciences

Figure 2. Hypothesized template information used for memory, guidance, and decisions during visual search. The column on the far left illustrates an example experimental task; targets are outlined by a white box, but this would not be present in an actual experimental display. The memory column (green) illustrates the information held in long-term or working memory of the target template. The guidance column (orange) illustrates the hypothesized good-enough features that are used for attentional guidance during each task context. The decision column (purple) illustrates information used for target-match decisions. The column on the far right cites studies from which the examples are constructed [3–5,16,19,39,64,68,72,73,76,77,84,90,92,97,98,105,110,111,114,117,118,123,155]. The face image is reproduced, with permission, from [124].

representational distinctiveness of targets from distractors during target discrimination tasks, including visual search.

The debate that emerged was over the exact form of the shift in target representation. Some argued that the target distortion was relational, such that an orange target would result in selection of the 'reddest' item in any display [71]. Others argued that the distortion reflected only a shift in target tuning towards more optimal 'reddish' feature values [68]. The good-enough framework for guidance helps to resolve this debate by suggesting that both models may be correct, with good-enough guidance using a broader relational code but target-match decisions using a more precise optimal code [4,39]. From this perspective, target shifting occurs in the target template to increase the target-to-distractor distinctiveness, but the implementation differs in precision for guidance and for match decisions. Measurements of guidance will appear to be relational because the best potential targets are expected to have the most exaggerated color within the display (i.e., the 'reddest'), and therefore the precise color value is not essential. Using a 'fuzzy' relational code will typically be successful in localizing a good-enough candidate target and a more precise optimal code can be used to check whether it is indeed the target after selection. Interestingly, target shifting is significantly reduced when another target feature dimension is highly dissimilar from distractors, and can be used to more rapidly discriminate candidate targets [75]. This suggests that relational guidance is a strategic form of good-enough guidance.

Good-enough guidance uses the most diagnostic target feature

Another way in which good-enough guidance can be seen is through the flexible prioritization of different visual features within a single target template. One might assume that the best way to find a target is to use all of its known features to guide attention precisely to the exact target, but this does not seem to be the case. Moreover, it is notable that search targets in our everyday lives may be defined on the basis of abstract category templates at varying levels of specificity, rather than by the specific exemplars often used in laboratory experiments (e.g., I may be searching within a store for any gift appropriate for my niece, any stuffed toy, or even any teddy-bear she might like, rather than this a specific teddy-bear).

How does guidance operate under these situations? This can play out in two ways: (i) by prioritizing one feature of a multidimensional target, or (ii) by prioritizing category-level information rather than specific features (Figure 2, rows 3–4).

In the case of a multidimensional target (e.g., a red circle), good-enough guidance occurs when one feature ('red') is prioritized over the other ('circle') during the initial selection of candidate targets. Once attention and/or an eye movement is guided to a candidate object, additional feature dimensions may contribute to the match decision to improve identification accuracy. However, what determines which feature has the highest priority? This likely depends on expectations of the sensory context, including distractor features, as well as on natural biases within the visual system. For example, when distractors are more similar to the target in one dimension, the less similar feature dimension is prioritized [76–80]. When target features are themselves uncertain, the most reliable dimension is prioritized [81–84]. Similar principles of flexible weighting have been identified for attentional guidance towards information with higher expected precision between modalities [85] and visual-motor codes [86]. However, when all else is equal, some visual features appear to be relied upon more than others to select candidate objects. For example, color commonly dominates guidance, perhaps because color processing is rapid and frequently carries high perceptual discriminability [3,18–20,39,82,87].

In addition to prioritizing one of many features, category-level information is also often used when it is sufficient to localize candidate targets. Real-world objects such as apples, cars, soccer balls, water cups, and our friends can be defined across multiple precise feature dimensions (e.g., color, shape, orientation, texture, size, motion). Nevertheless, instead of using highly precise information, a search for these objects often relies on higher-order category-level representations. For example, search for targets cued by a word is found more quickly when they contain typical visual features for their category (e.g., when searching for a 'bird', a canary will be found faster than an ostrich [83]). Moreover, this category effect is modulated by scene information such that target templates created from word cues are automatically scene-consistent; for example, the word 'bottle' produces a target template of a baby bottle when preceded by the scene of a baby's room, or of a beer bottle when the search is preceded by an image of a bar [88]. Generally, category representations may be preferable because they are more stable and serve as consistent predictive priors for any member of the category and under different viewing conditions [62,89–97]. Thus, attentional guidance may default to good-enough category features even when more specific item information is available. This may be true when there is either uncertainty about the exact form of a target or when distractors from the same category are also present [84,90,97,98].

Although it might seem counterintuitive to only use a subset of information for guidance, doing so makes sense from a good-enough perspective because reducing the number of dimensions makes guidance more efficient. For example, there is evidence that multiple memory representations can be maintained with high precision (e.g., the identity of two target colors), but selecting multiple sensory representations in parallel (e.g., picking out two targets with different colors at once) slows information processing [30,92]. Constraints in the number of sensory representations that can be used for guidance may stem from competition between the simultaneous networks that are necessary to maintain interactions between perceptual, mnemonic, and motor regions. Because attentional guidance relies on network interactions between memory and motor representations in frontal cortex and perceptual representation in sensory cortex, maintaining multiple such networks may cause interference or produce efficiency costs [27,30,55,99,100]. Thus, using fewer dimensions for guidance and selection satisfies the need to find probable candidate targets without creating potential interference between features. However, there does not seem to be a fixed number of features that are always optimal to be held efficiently; and the number of dimensions that can be held with minimal effort might vary by individual and trial [101,102].

Good-enough guidance uses learned associations

Target templates are primarily characterized as containing only information about the target, but there is mounting evidence that non-target information that predicts the target is also used to guide attention. A well-known example comes from studies of scene perception in which structural and semantic information contribute to the localization and identification of target objects [103–109] (Figure 2, row 5). For example, 'anchor' objects are non-target objects that hold strong spatial predictions regarding other objects within an environment [110,111]. Võ and colleagues hypothesize that anchor objects are used in real-world environments to start a cascade of hierarchical searches until the target object is located. Finding a piece of chalk within a classroom might begin with attentional guidance by the chalkboard (large, green, rectangular). After the chalkboard is located, then guidance may be updated to the features of the actual target (small, white, cylindrical). Although the chalkboard is not the target, it provides good-enough information to guide attention to a smaller spatial region with a high probability of containing the chalk. Anchor objects are an example of good-enough attentional guidance within the real world and may help, in part, to explain why attentional guidance in naturalistic settings exceeds expectations based on traditional laboratory tasks [112].

Associated information is only considered to be a part of the target template when it is used for active attentional guidance. By contrast, semantic associations that are passively activated by a target but are not useful for guiding attention should not be considered to be part of the template. The distinction lies in whether the associated objects are used as a proxy for the target that guides attention or are only passively primed because of semantic relatedness. We hypothesize that proxy objects momentarily replace the target in guiding attention because they contain features that are easier to identify or to locate than the target itself, and are therefore good-enough predictors of the target that reduce uncertainty about where the target may be located [113]. Proxy information can include arbitrary context-dependent associations, physical regularities, or nonpermanent features of an object [87,114–116].

In addition to anchor objects, the use of proxies for guidance can also be found in short-term learned associations. For example, exposure to backpacks that have a particular color (e.g., black) results in faster detection and earlier fixations on black backpacks compared to other-colored backpacks during subsequent visual search [117]. This suggests that the color 'black' was used as a proxy for locating the target backpack even though color was not a defining feature of the target during visual search (Figure 2, row 6). Similarly, teaching observers an association between the face of a specific individual (e.g., 'Emma') and a scene category ('beaches') results in faster localization and identification of the target face when it appears with its associated scene, but only when search for the target face is difficult; when face discrimination is easy, the scene advantage disappears [118]. This suggests that proxy information is used to set sensory gain and guide attention when target search is difficult, but not when the features of the search target are easily distinguished from distractors.

The theory of good-enough attentional guidance predicts that proxy information will only be used during visual search if the following conditions are true: (i) the proxy predicts the target location, and (ii) the expected discriminability of the proxy within the current environmental context is greater than that of the target. A proxy object may be more discriminable than the target because it is visually salient (e.g., owing to size, color, shape, motion, etc.) or because it possesses perceptual features with a natural processing advantage [3,19]. If there are multiple potential proxies, the one that is the most discriminable and has the highest predictive value, given the current willingness of the observer to exert effort, will win the competition to guide attention first. Proxy information will therefore only be good-enough when it is predictive of the location or identity of the target, and can be detected more quickly than features that define the identity of the target. When targets are easily distinguishable from distractors, proxies should not be used to guide attention because they no longer confer an advantage in the speed–accuracy tradeoff: target location based on the proxy will not be faster and it will sometimes be wrong.

Concluding remarks

The work reviewed provides evidence for good-enough attentional guidance. The basic premise is that attentional guidance is often based on the simplest, good-enough, information that will produce a high-quality estimate of where a candidate target object is located at the current moment of time. Very often, this information will not be a veridical characteristic of the specific target. We hypothesize that good-enough guidance is efficient because it is fast and is followed by an eye (and/or head or body) movement that allows higher-acuity foveal vision to be used to make a more accurate target-match decision. This strategy operates well when errors have a low cost and can be rapidly corrected with an eye, head, or body movement toward the next potential target. However, if costs are higher, for example because there are limits in time or opportunities to find the target, we expect guidance to be determined by more precise information from one or more features. This will result in greater accuracy in selecting the target, but there will be costs associated with the accumulation of more precise evidence before each saccadic eye-movement, which should be reflected in longer saccade latencies or pre-saccadic dwell times [47,48,51]. Thus, we argue that, under typical conditions, attentional guidance operates by default on a good-enough principle because it is more efficient than waiting for more precise information to be accumulated before attention and eye movements are shifted; however, the calculus of what is good-enough will depend on the context and the amount of effort an individual is able or willing to exert [45,53–55,57,104]. Guidance will be based on more precise information

Outstanding questions

Is the use of good-enough information influenced by differences in the speed of processing between features or feature weighting?

What are the temporal dynamics of good-enough information transfer from the target template in memory to mechanisms of sensory gain used during attentional guidance?

How quickly can the information within the target template be learned in a new environmental context and used to shape good-enough guidance?

Do we have conscious access to our good-enough guidance settings?

Are there individual differences in the ability to optimize good-enough guidance, and does this change over the lifespan?

What brain networks support the interactions between templates in memory, features for guidance, and features for target-match decisions?

How does effort relate to the use of good-enough information during attentional guidance and decisions?

under conditions where selection accuracy is paramount and the good-enough principle leads to too many errors.

It may be that other perceptual or motor limitations contribute to the use of good-enough information for guidance. For example, the drop-off in visual acuity from central to peripheral vision may serve as an upper limit to the precision of information that can be used to guide attention [29,35–37]. This might increase reliance on good-enough features that are more easily perceived. Likewise, there may be a natural motor rhythm in the theta range that drives spatial attention and saccadic eye-movements [119]. However, we propose that the choice of what information is used for good-enough guidance is largely determined by the expected attentional efficiency in locating the target within the current sensory context (see Outstanding questions). Prior expectations of attentional efficiency will depend on multiple sources of information at different timescales, including selection history within the local context, feedback, long-term expectations, and explicit rules [120]. This principle does not preclude the use of veridical or complex information to guide attention – it only proposes that more veridical or complex information can be used and will only be used when there is no other information to drive attention more rapidly with sufficient accuracy, or when there are other motivational factors present that require greater accuracy. A formal model of the trade-off between good-enough information and veridical information within a given context remains to be developed, but evidence for this idea can be found in speed–accuracy trade-offs formalized in drift rates and criterion thresholds in accumulation models [4,39,121,122].

The argument that good-enough template information is used to guide attention is rooted in the notion that the target template in memory contains more information than is being currently used to guide attention. A subset of information from the target template that is good-enough for the current purpose is used for guidance, but a different subset may be used in the next moment to make identity decisions. If the object is not the target, the subset used to guide attention may stay the same if there are other potential targets to explore, or it may be updated to a new subset [51]. Inherent in this idea is that guidance and identity decisions operate in a continuous loop during visual search until the target is found; experience gained in each iteration serves as feedback that may be used to update the information from the template that is used for guidance. A constant in this cycle is the information held in memory within the target template that contains all the information we might have about the target [9]. The memory representation is therefore expected to be the most precise and most complete, whereas the guidance information is potentially the most imprecise, and decision information is somewhere between the two [16].

In summary, we propose that good-enough attentional guidance satisfies the question of 'what is attentional guidance during visual search for?' in an optimal manner: guidance is for localizing candidate targets that can then be rapidly evaluated as target-matches or non-matches. This definition does not require guidance to be very accurate, only sufficiently accurate to reduce the number of items to be investigated without too much cognitive effort so as to maximize the speed–accuracy–effort trade-off necessary to accomplish visual search.

Acknowledgments

This work was supported by National Institutes of Health (NIH) grant R01MH113855-01 to J.J.G. and Australian Research Council (ARC) DP210103430 to S.I.B.

Declaration of interests

The authors declare no conflicts of interest.

References

1. Wolfe, J.M. (2020) Visual search: how do we find what we are looking for? *Annu. Rev. Vis. Sci.* 6, 539–562
2. Hout, M.C. and Goldinger, S.D. (2015) Target templates: the precision of mental representations affects attentional guidance and decision-making in visual search. *Atten. Percept. Psychophys.* 77, 128–149
3. Alexander, R.G. *et al.* (2019) Specifying the precision of guiding features for visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 45, 1248–1264
4. Yu, X. *et al.* (2022) Attentional guidance and match decisions rely on different template information during visual search. *Psychol. Sci.* 33, 105–120
5. Kerzel, D. (2019) The precision of attentional selection is far worse than the precision of the underlying memory representation. *Cognition* 186, 20–31
6. Eimer, M. (2014) The neural basis of attentional control in visual search. *Trends Cogn. Sci.* 18, 526–535
7. Malcolm, G.L. and Henderson, J.M. (2010) Combining top-down processes to guide eye movements during real-world scene search. *J. Vis.* 10, 4
8. Duncan, J. and Humphreys, G.K. (1989) Visual search and stimulus similarity. *Psychol. Rev.* 96, 433–458
9. Wolfe, J.M. (2021) Guided Search 6.0: an updated model of visual search. *Psychon. Bull. Rev.* 28, 1060–1092
10. Treisman, A.M. and Gelade, G. (1980) A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136
11. Nako, R. *et al.* (2014) Item and category-based attentional control during search for real-world objects: can you find the pants among the pans? *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1283–1288
12. Rajsic, J. and Woodman, G.F. (2020) Do we remember templates better so that we can reject distractors better? *Atten. Percept. Psychophys.* 82, 269–279
13. Malcolm, G.L. and Henderson, J.M. (2009) The effects of target template specificity on visual search in real-world scenes: evidence from eye movements. *J. Vis.* 9, 8
14. Vickery, T.J. and Jiang, Y.V. (2009) Inferior parietal lobule supports decision making under uncertainty in humans. *Cereb. Cortex* 19, 916–925
15. Geng, J.J. and Witkowski, P. (2019) Template-to-distractor distinctiveness regulates visual search efficiency. *Curr. Opin. Psychol.* 29, 119–125
16. Yu, X. *et al.* (2022) Visual search guidance uses coarser template information than target-match decisions. *Atten. Percept. Psychophys.* 84, 1432–1445
17. Becker, S.I. *et al.* (2021) The attentional blink: a relational accountof attentional engagement. *Psychon. Bull. Rev.* 28, 219–227
18. Wolfe, J.M. and Horowitz, T.S. (2017) Five factors that guide attention in visual search. *Nat. Hum. Behav.* 1, 0058
19. Hulleman, J. (2020) Quantitative and qualitative differences in the top-down guiding attributes of visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 46, 942–964
20. Huang, L. (2015) Color is processed less efficiently than orientation in change detection but more efficiently in visual search. *Psychol. Sci.* 26, 646–652
21. Becker, S.I. *et al.* (2019) Target templates in singleton search vs. feature-based search modes. *Vis. Cogn.* 27, 502–517
22. Vo, V.A. *et al.* (2017) Spatial tuning shifts increase the discriminability and fidelity of population codes in visual cortex. *J. Neurosci.* 37, 3386–3401
23. Serences, J.T. and Boynton, G.M. (2007) Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron* 55, 301–312
24. Helfrich, R.F. *et al.* (2018) Neural mechanisms of sustained attention are rhythmic. *Neuron* 99, 854–865
25. Baldauf, D. and Desimone, R. (2014) Neural mechanisms of object-based attention. *Science* 344, 424–427
26. Bichot, N.P. *et al.* (2019) The role of prefrontal cortex in the control of feature attention in area V4. *Nat. Commun.* 10, 5727
27. Manohar, S.G. *et al.* (2019) Neural mechanisms of attending to items in working memory. *Neurosci. Biobehav. Rev.* 101, 1–12
28. Huynh Cong, S. and Kerzel, D. (2022) The allocation of working memory resources determines the efficiency of attentional templates in single- and dual-target search. *J. Exp. Psychol. Gen.* 151, 2977–2989
29. Vater, C. *et al.* (2022) Peripheral vision in real-world tasks: a systematic review. *Psychon. Bull. Rev.* 29, 1531–1557
30. Ort, E. *et al.* (2019) Humans can efficiently look for but not select multiple visual objects. *Elife* 8, e49130
31. Whitney, D. and Levi, D.M. (2011) Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends Cogn. Sci.* 15, 160–168
32. Koranda, M.J. *et al.* (2022) Good-enough production: selecting easier words instead of more accurate ones. *Psychol. Sci.* 33, 1440–1451
33. Ferreira, F. *et al.* (2002) Good-enough representations in language comprehension. *Curr. Dir. Psychol. Sci.* 11, 11–15
34. Ballard, D.H. *et al.* (1992) Hand-eye coordination during sequential tasks. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 337, 331–338
35. Hulleman, J. and Olivers, C.N.L. (2017) The impending demise of the item in visual search. *Behav. Brain Sci.* 40, e132
36. Rosenholtz, R. *et al.* (2012) A summary statistic representation in peripheral vision explains visual search. *J. Vis.* 12, 14
37. Lleras, A. *et al.* (2022) Incorporating the properties of peripheral vision into theories of visual search. *Nat. Rev. Psychol.* 1, 590–604
38. Cohen, M.A. *et al.* (2020) The limits of color awareness during active, real-world vision. *Proc. Natl. Acad. Sci. U. S. A.* 117, 13821–13827
39. Hamblin-Frohman, Z. and Becker, S.I. (2021) The attentional template in high and low similarity search: optimal tuning or tuning to relations? *Cognition* 212, 104732
40. Rosenholtz, R. *et al.* (2007) Measuring visual clutter. *J. Vis.* 7, 11–22
41. Manassi, M. and Whitney, D. (2018) Multi-level crowding and the paradox of object recognition in clutter. *Curr. Biol.* 28, R127–R133
42. Hansen, T. *et al.* (2009) Color perception in the intermediate periphery of the visual field. *J. Vis.* 9, 26
43. Pereira, E.J. and Castelhano, M.S. (2014) Peripheral guidance in scenes: The interaction of scene context and object content. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 2056–2072
44. Wagner, I. *et al.* (2023) Humans trade off search costs and accuracy in a combined visual search and perceptual task. *Atten. Percept. Psychophys.* 85, 23–40
45. Seideman, J.A. *et al.* (2018) Saccade metrics reflect decision-making dynamics during urgent choices. *Nat. Commun.* 9, 2907
46. Pearson, D. *et al.* (2016) Value-modulated oculomotor capture by task-irrelevant stimuli is a consequence of early competition on the saccade map. *Atten. Percept. Psychophys.* 78, 2226–2240
47. Eckstein, M.P. *et al.* (2001) Quantifying the performance limits of human saccadic targeting during visual search. *Perception* 30, 1389–1401
48. van Heusden, E. *et al.* (2022) An attentional limbo: saccades become momentarily non-selective in between saliency-driven and relevance-driven selection. *Psychon. Bull. Rev.* 29, 1327–1337
49. Ghahghaei, S. and Verghese, P. (2015) Efficient saccade planning requires time and clear choices. *Vis. Res.* 113, 125–136
50. Schütz, A.C. *et al.* (2012) Dynamic integration of information about salience and value for saccadic eye movements. *Proc. Natl. Acad. Sci. U. S. A.* 109, 7547–7552
51. Hamblin-Frohman, Z. *et al.* (2023) The progression of visual search in multiple item displays: first relational, then feature-specific. *ArXiv* Published online January 9, 2023. http://arxiv.org/abs/2301.03157
52. Zhang, Z. and Carlisle, N.B. (2022) Assessing recoding accounts of negative attentional templates using behavior and eye tracking. *J. Exp. Psychol. Learn. Mem. Cogn.*
53. Irons, J.L. and Leber, A.B. (2018) Characterizing individual variation in the strategic use of attentional control. *J. Exp. Psychol. Hum. Percept. Perform.* 44, 1637–1654

54. Droll, J.A. and Hayhoe, M.M. (2007) Trade-offs between gaze and working memory use. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1352–1365
55. Petitet, P. *et al.* (2021) The computational cost of active information sampling before decision-making under uncertainty. *Nat. Hum. Behav.* 5, 935–946
56. Irons, J.L. and Leber, A.B. (2016) Choosing attentional control settings in a dynamically changing environment. *Atten. Percept. Psychophys.* 78, 2031–2048
57. Tunnermann, J. *et al.* (2021) How feature context alters attentional template switching. *J. Exp. Psychol. Hum. Percept. Perform.* 47, 1431–1444
58. Li, A. *et al.* (2023) How do people find pairs? *PsyArXiv* Published online January 3, 2023. https://doi.org/2010.31234/osf.io/31568q31239
59. Harris, C.M. and Wolpert, D.M. (2006) The main sequence of saccades optimizes speed–accuracy trade-off. *Biol. Cybern.* 95, 21–29
60. York, A. and Becker, S.I. (2020) Top-down modulation of gaze capture: feature similarity, optimal tuning, or tuning to relative features? *J. Vis.* 20, 6
61. Williams, J.R. *et al.* (2022) Guidance of attention by working memory is a matter of representational fidelity. *J. Exp. Psychol. Hum. Percept. Perform.* 48, 202–231
62. Buschman, T.J. and Miller, E.K. (2022) Working memory is complex and dynamic, like your thoughts. *J. Cogn. Neurosci.* 35, 17–23
63. Lau, J.S. *et al.* (2021) Target templates in low target-distractor discriminability visual search have higher resolution, but the advantage they provide is short-lived. *Atten. Percept. Psychophys.* 83, 1435–1454
64. Anderson, B.A. (2014) On the precision of goal-directed attentional selection. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1755–1762
65. Bauer, B. *et al.* (1996) Visual search for colour target that are or are not linearly separable from distractors. *Vis. Res.* 36, 1439–1465
66. Hodsoll, J. and Humphreys, G.W. (2001) Driving attention with the top down: the relative contribution of target templates to the linear separability effect in the size dimension. *Percept. Psychophys.* 63, 918–926
67. D'Zmura, M. (1991) Color in visual search. *Vis. Res.* 31, 951–966
68. Navalpakkam, V. and Itti, L. (2007) Search goal tunes visual features optimally. *Neuron* 53, 605–617
69. Won, B.-Y. *et al.* (2020) Flexible target templates improve visual search accuracy for faces depicting emotion. *Atten. Percept. Psychophys.* 82, 2909–2923
70. Scolari, M. *et al.* (2012) Optimal deployment of attentional gain during fine discriminations. *J. Neurosci.* 32, 7723–7733
71. Becker, S.I. *et al.* (2010) The role of relational information in contingent capture. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1460–1476
72. Kerzel, D. (2020) Direct evidence for the optimal tuning of attention. *J. Exp. Psychol. Hum. Percept. Perform.* 46, 716–728
73. Chapman, A.F. *et al.* (2022) Optimal tuning of feature-based attention warps the perception of visual features. *PsyArXiv* Published online June 15, 2022. https://doi.org/2010.31234/osf.io/hbfup
74. Pouget, A. and Bavelier, D. (2007) Paying attention to neurons with discriminating taste. *Neuron* 53, 473–475
75. Yu, X. *et al.* (2022) Shifting target templates away from linearly separable distractor features is task-adapative. *PsyArXiv* Published online December 28, 2022. https://doi.org/2010.31234/osf.io/ahq31236d
76. Boettcher, S.E.P. *et al.* (2020) Functional biases in attentional templates from associative memory. *J. Vis.* 20, 7
77. Lee, J. and Geng, J.J. (2020) Flexible weighting of target features based on distractor context. *Atten. Percept. Psychophys.* 82, 739–751
78. Liesefeld, H.R. and Muller, H.J. (2019) Distractor handling via dimension weighting. *Curr. Opin. Psychol.* 29, 160–167
79. Lerebourg, M. *et al.* (2022) Expected distractor context biases the attentional template for target shapes. *BioRxiv* Published online October 25, 2022. https://doi.org/2010.1101/2022.2010.2018.512686
80. Barras, C. and Kerzel, D. (2017) Target-nontarget similarity decreases search efficiency and increases stimulus-driven control in visual search. *Atten. Percept. Psychophys.* 79, 2037–2043
81. Witkowski, P. and Geng, J.J. (2019) Learned feature variance is encoded in the target template and drives visual search. *Vis. Cogn.* 27, 487–501
82. Hansmann-Roth, S. *et al.* (2019) Representing color and orientation ensembles: Can observers learn multiple feature distributions? *J. Vis.* 19, 2
83. Robbins, A. and Hout, M.C. (2020) Typicality guides attention during categorical search, but not universally so. *Q. J. Exp. Psychol.* 73, 1977–1999
84. Witkowski, P.P. and Geng, J.J. (2022) Attentional priority is determined by predicted feature distributions. *J. Exp. Psychol. Hum. Percept. Perform.* 48, 1201–1212
85. Knill, D.C. and Pouget, A. (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–719
86. Boettcher, S.E. *et al.* (2021) Output planning at the input stage in visual working memory. *Sci. Adv.* 7, eabe8212
87. Bravo, M.J. and Farid, H. (2016) Observers change their target template based on expected context. *Atten. Percept. Psychophys.* 78, 829–837
88. Robbins, A. and Hout, M.C. (2020) Scene priming provides clues about target appearance that improve attentional guidance during categorical search. *J. Exp. Psychol. Hum. Percept. Perform.* 46, 220–230
89. Wolfe, J.M. *et al.* (1992) The role of categorization in visual search for orientation. *J. Exp. Psychol. Hum. Percept. Perform.* 18, 34–49
90. Bravo, M.J. and Farid, H. (2012) Task demands determine the specificity of the search template. *Atten. Percept. Psychophys.* 74, 124–131
91. Hout, M.C. *et al.* (2017) Categorical templates are more useful when features are consistent: Evidence from eye movements during search for societally important vehicles. *Atten. Percept. Psychophys.* 79, 1578–1592
92. Yu, C.P. *et al.* (2016) Searching for category-consistent features: a computational approach to understanding visual category representation. *Psychol. Sci.* 27, 870–884
93. Cunningham, C.A. and Wolfe, J.M. (2014) The role of object categories in hybrid visual and memory search. *J. Exp. Psychol. Gen.* 143, 1585–1599
94. Wyble, B. *et al.* (2013) Contingent attentional capture by conceptually relevant images. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 861–871
95. Zivony, A. and Eimer, M. (2022) Expectation-based blindness: predictions about object categories gate awareness of focally attended objects. *Psychon. Bull. Rev.* 29, 1879–1889
96. Yeh, L.C. and Peelen, M.V. (2022) The time course of categorical and perceptual similarity effects in visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 48, 1069–1082
97. Addleman, D.A. *et al.* (2022) Attention to object categories: selection history determines the breadth of attentional tuning during real-world object search. *PsyArXiv* Published online September 30, 2022. https://doi.org/2010.31234/osf.io/s31238nyz
98. Alexander, R.G. and Zelinsky, G.J. (2011) Visual similarity effects in categorical search. *J. Vis.* 11, 9
99. Thyer, W. *et al.* (2022) Storage in visual working memory recruits a content-independent pointer system. *Psychol. Sci.* 33, 1680–1694
100. Lewis-Peacock, J.A. *et al.* (2015) Neural evidence for the flexible control of mental representations. *Cereb. Cortex* 25, 3303–3313
101. Adam, K.C.S. and deBettencourt, M.T. (2019) Fluctuations of attention and working memory. *J. Cogn.* 2, 33
102. Oberauer, K. (2019) Working memory and attention – a conceptual analysis and review. *J. Cogn.* 2, 36
103. Biederman, I. (1972) Perceiving real-world scenes. *Science* 177, 77–80
104. Wolfe, J.M. *et al.* (2011) Visual search in scenes involves selective and nonselective pathways. *Trends Cogn. Sci.* 15, 77–84
105. Boettcher, S.E.P. *et al.* (2018) Anchoring visual search in scenes: assessing the role of anchor objects on eye movements during visual search. *J. Vis.* 18, 11

106. Helbing, J. *et al.* (2022) Auxiliary scene-context information provided by anchor objects guides attention and locomotion in natural search behavior. *Psychol. Sci.* 33, 1463–1476
107. Castelhano, M.S. and Krzys, K. (2020) Rethinking apace: a review of perception, attention, and memory in scene processing. *Annu. Rev. Vis. Sci.* 6, 563–586
108 Z. Zhang and N.B. Carlisle, Explicit attentional goals unlock implicit spatial statistical learning, *J. Exp. Psychol. Gen.*, In press.
109. Beitner, J. *et al.* (2021) Get your guidance going: investigating the activation of spatial priors for efficient search in virtual reality. *Brain Sci.* 11, 44
110. Mack, S.C. and Eckstein, M.P. (2011) Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. *J. Vis.* 11, 9
111. Võ, M.L. *et al.* (2019) Reading scenes: how scene grammar guides attention and aids perception in real-world environments. *Curr. Opin. Psychol.* 29, 205–210
112. Wolfe, J.M. *et al.* (2011) Visual search for arbitrary objects in real scenes. *Atten. Percept. Psychophys.* 73, 1650–1671
113. Parr, T. and Friston, K.J. (2018) Attention or salience? *Curr. Opin. Psychol.* 29, 1–5
114. Kershner, A.M. and Hollingworth, A. (2022) Real-world object categories and scene contexts conjointly structure statistical learning for the guidance of visual search. *Atten. Percept. Psychophys.* 84, 1304–1316
115. Gayet, S. and Peelen, M.V. (2019) Scenes modulate object processing before interacting with memory templates. *Psychol. Sci.* 30, 1497–1509
116. Collegio, A.J. *et al.* (2019) Attention scales according to inferred real-world object size. *Nat. Hum. Behav.* 3, 40–47
117. Bahle, B. *et al.* (2021) Categorical cuing: object categories structure the acquisition of statistical regularities to guide visual search. *J. Exp. Psychol. Gen.* 150, 2552–2566
118. Zhou, Z. and Geng, J.J. (2023) Learned associations serve as target proxies during difficult but not easy during visual search. *PsyArXiv* Published online January 24, 2023. http://dx.doi.org/10.31234/osf.io/y26qx
119. Eickhoff, S.B. *et al.* (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335
120. Nobre, A.C. and Stokes, M.G. (2019) Premembering experience: a hierarchy of time-scales for proactive attention. *Neuron* 104, 132–146
121. Lleras, A. *et al.* (2020) A target contrast signal theory of parallel processing in goal-directed search. *Atten. Percept. Psychophys.* 82, 394–425
122. Krajbich, I. (2019) Accounting for attention in sequential sampling models of decision making. *Curr. Opin. Psychol.* 29, 6–11
123. Zivony, A. and Eimer, M. (2022) Categorization templates modulate selective attention. *J. Exp. Psychol. Hum. Percept. Perform.* 48, 1294–1312
124. Ma, D.S. *et al.* (2015) The Chicago face database: a free stimulus set of faces and norming data. *Behav. Res. Methods* 47, 1122–1135
125. Carlisle, N.B. *et al.* (2011) Attentional templates in visual working memory. *J. Neurosci.* 31, 9315–9322
126. Drew, T. *et al.* (2017) One visual search, many memory searches: an eye-tracking investigation of hybrid search. *J. Vis.* 17, 5
127. Moore, T. *et al.* (2003) Visuomotor origins of covert spatial attention. *Neuron* 40, 671–683
128. Ester, E.F. *et al.* (2016) Feature-selective attentional modulations in human frontoparietal cortex. *J. Neurosci.* 36, 8188–8199
129. Long, N.M. and Kuhl, B.A. (2018) Bottom-up and top-down factors differentially influence stimulus representations across large-scale attentional networks. *J. Neurosci.* 38, 2495–2504
130. Lorenc, E.S. and Sreenivasan, K.K. (2021) Reframing the debate: the distributed systems view of working memory. *Vis. Cogn.* 29, 416–424
131. Stokes, M.G. *et al.* (2020) Theoretical distinction between functional states in working memory and their corresponding neural states. *Vis. Cogn.* 28, 420–432
132. Teng, C. and Postle, B.R. (2021) Understanding occipital and parietal contributions to visual working memory: commentary on Xu (2020). *Vis. Cogn.* 29, 401–408
133. Xu, Y. (2020) Revisit once more the sensory storage account of visual working memory. *Vis. Cogn.* 28, 433–446
134. Schafer, R.J. and Moore, T. (2011) Selective attention from voluntary control of neurons in prefrontal cortex. *Science* 332, 1568–1571
135. Ruff, C.C. *et al.* (2006) Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. *Curr. Biol.* 16, 1479–1488
136. Aly, M. and Turk-Browne, N.B. (2017) How hippocampal memory shapes, and is shaped by, attention. In *The Hippocampus from Cells to Systems* (Hannula, D.E. and Duff, M.C., eds), pp. 369–403, Springer International
137. Olmos-Solis, K. *et al.* (2021) Content or status: frontal and posterior cortical representations of object category and upcoming task goals in working memory. *Cortex* 135, 61–77
138. Nee, D.E. (2021) Integrative frontal-parietal dynamics supporting cognitive control. *Elife* 10, e57244
139. Feredoes, E. *et al.* (2011) Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. *Proc. Natl. Acad. Sci. U. S. A.* 108, 17510–17515
140. Wutz, A. *et al.* (2018) Different levels of category abstraction by different dynamics in different prefrontal areas. *Neuron* 97, 716–726
141. Shah, S. *et al.* (2022) Attention differentially modulates multiunit activity in the lateral geniculate nucleus and V1 of macaque monkeys. *J. Comp. Neurol.* 530, 1064–1080
142. Luck, S.J. *et al.* (1997) Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J. Neurophysiol.* 77, 24–42
143. Keller, A.S. *et al.* (2022) Attention enhances category representations across the brain with strengthened residual correlations to ventral temporal cortex. *Neuroimage* 249, 118900
144. Panichello, M.F. and Buschman, T.J. (2021) Shared mechanisms underlie the control of working memory and attention. *Nature* 592, 601–605
145. Boshra, R. and Kastner, S. (2022) Attention control in the primate brain. *Curr. Opin. Neurobiol.* 76, 102605
146. Bocincova, A. *et al.* (2020) A common neural network architecture for visual search and working memory. *Vis. Cogn.* 28, 356–371
147. Reynolds, J.H. and Heeger, D.J. (2009) The normalization model of attention. *Neuron* 61, 168–185
148. Yu, Q. *et al.* (2020) Different states of priority recruit different neural representations in visual working memory. *PLoS Biol.* 18, e3000769
149. van Loon, A.M. *et al.* (2018) Current and future goals are represented in opposite patterns in object-selective cortex. *Elife* 7, e38677
150. Jigo, M. *et al.* (2018) Neural determinants of task performance during feature-based attention in human cortex. *eNeuro* 5 ENEURO.0375-17.2018
151. Saenz, M. *et al.* (2002) Global effects of feature-based attention in human visual cortex. *Nat. Neurosci.* 5, 631–632
152. Brouwer, G.J. and Heeger, D.J. (2013) Categorical clustering of the neural representation of color. *J. Neurosci.* 33, 15454–15465
153. Jehee, J.F. *et al.* (2011) Attention improves encoding of task-relevant features in the human visual cortex. *J. Neurosci.* 31, 8210–8219
154. O'Craven, K.M. *et al.* (1999) fMRI evidence for objects as the units of attentional selection. *Nature* 40, 584–587
155. Võ, M.L. *et al.* (2019) Reading scenes: how scene grammar guides attention and aids perception in real-world environments. *Curr. Opin. Psychol.* 29, 205–210