Trends in
Cognitive Sciences

CellPress

Review

Attention in the wild: balancing flexibility and stability

Maëlle Lerebourg [1,*], Surya Gayet [2], Floris P. de Lange [1], and Marius V. Peelen [1,*]

To prioritize the visual processing of task-relevant objects in our surroundings, we rely on an attentional template—an internal representation of object features that guides attention toward potential targets. Decades of research have characterized attentional templates for simple targets in artificial arrays. How could templates function in real-world search, where target appearance is variable and objects are embedded in complex, dynamic scenes? We consider two possibilities: (i) flexible templates that are adapted to changing scene contexts and (ii) stable ('one-size-fits-all') templates that generalize across contexts. We review recent behavioral and neuroimaging evidence for both possibilities and discuss how optimal search depends on balancing the relative costs and benefits of template adaptation, enabling efficient attention 'in the wild'.

Highlights

To guide attention toward potential target objects, searchers rely on an attentional template, an internal representation of task-relevant features. However, prioritizing specific visual features is challenging in natural scenes, where the target's appearance is highly variable.

Real-world attentional templates can be highly flexible, optimized to fit the current scene based on contextual expectations.

Alternatively, searchers can also form context-invariant, stable templates. These provide less accurate attentional guidance but reflect an efficient one-size-fits-all approach for detecting familiar targets across contexts and occurrences.

Whether attentional templates are flexibly updated depends on the reliability of contextual expectations, the relative cost of potential misguidance, and the cost of generating flexible templates.

Attentional templates for naturalistic search

Is it safe to cross the street? Where are the apples in the supermarket? To answer such questions, we perform countless visual searches throughout the day, often within highly complex and cluttered environments, and typically with remarkable efficiency and ease. Successful search (and goal-directed behavior more generally) relies on attentional mechanisms that enable us to selectively process behaviorally relevant objects while discarding currently irrelevant ones. To do so effectively, a key idea in current theories of selective attention and visual search is that searchers can proactively prepare their visual system to facilitate sensory processing of upcoming targets. This idea is captured by the theoretical concept of an **attentional template** (see Glossary)—an internal (memory) representation of the to-be-attended object that guides attention and biases subsequent visual processing in favor of target-like input (Box 1).

By asking participants to perform simplified versions of visual search tasks in the lab, we have gained a good understanding of the consequences and neural basis of such template-based prioritization. This research has shown that when specific target features are known (e.g., by cueing the target's color or shape), attention is drawn or guided toward matching objects, restricting further processing to target-like objects and thereby reducing interference by nonmatching **distractors** [12,13]. On the neural level, preparing to search for a target object has been shown to involve selective pre-activation of target-selective neural populations in visual cortex [4,14], followed by increased sensory gain for feature-matching objects in the search scene [15]. After potential target candidates have been located in this way, they will be sequentially attended and processed further for their eventual identification. Both **attentional guidance** and object identification typically rely on at least partially overlapping features. However, attention can only be guided by a limited number of distinct objects/features at once, while we are able to memorize and identify many more separate targets (e.g., not only the apple we are currently searching for but multiple items on our shopping list). Recent theories therefore distinguish between the attentional template guiding attention and a more precise, high-capacity target representation used for target identification (the target template), based on

[1]Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands
[2]Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, The Netherlands

*Correspondence:
maelle.lerebourg@donders.ru.nl
(M. Lerebourg) and
marius.peelen@donders.ru.nl
(M.V. Peelen).

Box 1. Attentional templates, working memory, and imagery

An attentional template is often regarded (and described) as a mental image—a visual representation of the target object held in mind. This description links attentional templates to both visual working memory (VWM) and mental imagery, thus raising the question of how the three are related or different.

Similar to attentional templates, VWM can bias attention and eye movements toward visual input matching the content of VWM [1,2] and prioritize its conscious access [3,4]. In addition, both attention and working memory may involve top-down activation of object representations in visual cortex [5,6] (Box 3). Importantly, however, not every working memory representation acts as an attentional template and biases sensory processing, for instance, depending on whether the memory representation is currently in a prioritized state, and where and how working memory content is encoded in the brain [7]. Most likely, working memory representations and attentional templates rely on overlapping storage mechanisms but differ in terms of stored content, depending on what aspect(s) of a stimulus are memorized and for which behavioral goal.

While some search (or working memory) tasks may allow us to maintain an almost 'imagery-like' representation of an object in mind, introspectively, at least most of us do not seem to experience vivid mental imagery of the target object when searching a scene. Indeed, people without vivid imagery (aphantasics) often perform comparably to controls in template-based visual search tasks [8,9]. And although others have reported modest differences in VWM or search performance between observers that vary in imagery vividness [10,11], aphantasic individuals exhibit normal functioning in daily behavior. Thus, the effectiveness of an attentional template does not seem to strongly rely on the creation of a vivid mental image. Moreover, during naturalistic search, templates can incorporate properties of the target object that cannot be directly visualized because they are location-unspecific, invariant to viewpoint or size, or abstracted away from any single visual feature.

In conclusion, while both imagery and attentional templates depend on VWM to be actively maintained, the representational content of these internal representations is optimized for the task at hand (to imagine, to memorize, or to search) and may further vary between individuals [12]. For attentional templates, this optimization takes the form of biasing the visual processing hierarchy toward target-like input at the expense of distractor input, by selectively increasing the excitability of neural populations tuned to diagnostic target properties.

detailed long-term memory for the target [17,18]. In this review, we focus specifically on the content of the template guiding attention.

While it is relatively well understood how we can prioritize the processing of singular low-level features cued by the experimenter, much less is still known about how attention is quickly guided to objects in rich and dynamically changing real-world environments. Here, specific target features are often unknown, ever-changing, and strongly context-dependent. Even when searching for a simple object such as an apple, the retinal input evoked by the very same fruit—and many of its diagnostic visual features—will change dramatically when it is placed on a kitchen counter or hanging from a tree, depending on (dynamic) changes in viewpoint, lighting conditions, distance from the observer, surrounding (potentially occluding) objects, etc. (Figure 1). Despite this apparent challenge, most real-world searches are remarkably fast and efficient [19,20]. What does the attentional template represent when the specific features best characterizing the target object are largely unknown? Is the concept of an attentional template even still useful?

Here, we review behavioral and neural evidence showing that the content of attentional templates can be highly flexible—that is, adapted to the current scene or task context—capitalizing on predictable environmental structure to prioritize different features in different contexts. Then, we review seemingly conflicting evidence showing that attentional templates for naturalistic search are often stable, capitalizing on high-level object representations that are either invariant to contexts or that already incorporate context during visual processing. Finally, we bring these research lines together and discuss the balance between flexible and stable templates, arguing that stable templates, incorporating long-term statistics of visual environments, often provide useful defaults for naturalistic visual search.

Glossary

Attentional guidance: the process of directing spatial attention (and possibly eye movements) toward the location of relevant targets in the visual scene based on our top-down goals.

Attentional template: a working memory representation reflecting features of a to-be-attended object that is activated when preparing to attend/search for a given object. To function as attentional template, this representation should guide attention toward objects in the visual field matching template features, at the expense of other nonmatching objects. Also referred to as guiding template.

Distractors: objects in a scene or search display that are not current targets.

Selection history: strong attentional biases in favor of recent search targets and away from recent distractors, largely independent of the observer's search goals.

	Flexible template	Stable template
Context-specificity	Context-specific: adapted to the current location and scene	Context-invariant: not adapted to the current location and scene
Feature-weighting	Feature weights (e.g., color) are variable across contexts	Feature weights (e.g., color) are fixed across contexts
Veridicality	Possibly distorted to reflect diagnostic features	Coarse match to the target
Benefits	Early prioritization of target-like input	Generalizes across locations and contexts
Costs	Can be slow and effortful to set up (detailed context processing, long-term memory retrieval, and template switch costs)	Less efficient guidance, nontargets may be processed up to higher representational level

Trends in Cognitive Sciences

Figure 1. Characteristics of stable and flexible templates. Schematic depiction of flexible and stable attentional templates and some of their characteristics.

Adapting the template to context: evidence for flexible templates

Despite their apparent complexity, real-world environments are highly structured and predictable, and these regularities can be exploited for efficient visual search [21–23]. One way in which contextual expectations facilitate search is by restricting where and when attentional templates should be deployed, as reviewed elsewhere [24–26]. Here, we focus on how the feature-based content of the template may be influenced by contextual expectations. For example, when searching for an apple on the kitchen counter, we may not know its precise shade of red but can easily anticipate its approximate size, shape, and other objects surrounding it (e.g., a fruit basket and oranges, not fire trucks). A flexible template would incorporate these expectations to optimize the match between the template and visual input evoked by the target in the current context (Figure 2), resulting in fast and efficient prioritization. In addition, contextual expectations can change how different target features are weighted in the template to reflect those that are the most diagnostic in the current context (e.g., relying more on the apple's color than its shape when it is placed next to oranges).

A rapidly growing number of studies have provided behavioral, neuroimaging, and eye-tracking evidence that attentional templates can be flexibly optimized for the current search.

Flexible templates adapt to target and distractor expectations

Important evidence for flexible templates comes from studies showing that searchers can rapidly learn and exploit newly introduced predictable structure in the environment, adapting their attentional templates to the distribution of low-level features belonging to targets and distractors. Such expectations are sometimes established by explicit trial-wise context cues (e.g., that the target on this trial will have a unique shape or color) but can also be quickly formed based on repeated searches [27,28]. These expectations shape how attention is

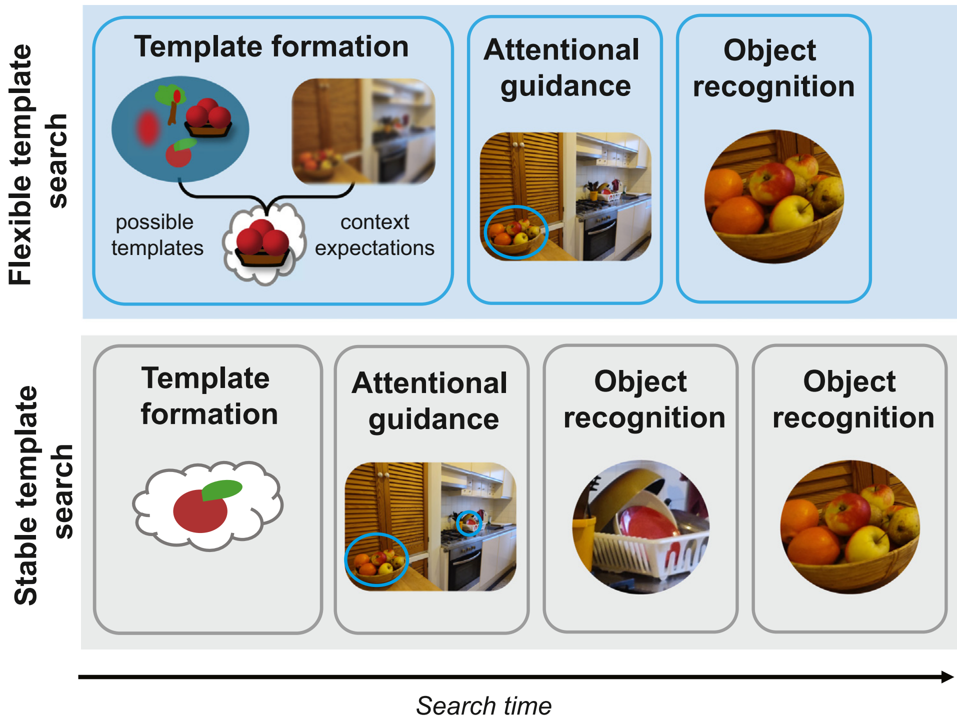

Figure 2. Timecourse of search with a flexible or stable template. If the attentional template is flexible (upper row), searchers may use different templates, varying in precision and including different features, for different contexts. For a given search, the specific scene determines which features are included in the template. For instance, when searching for an apple in the kitchen, the template may include the fruit basket in which the apples are likely placed, and the retinal size of the apples given the expected viewing distance. Once this template is formed, attention and eye movements are efficiently guided toward matching visual input in the scene, here resulting in a single target candidate (circled in blue), which then proceeds to an object recognition stage. A stable template (lower row), by contrast, is not adapted to the specific context, reflecting only more generic apple features, for example, that they are red and round. This leads to less precise attentional guidance, thus prioritizing more target candidates. These can include distractors sharing similar features, such as the red, round plate in the dish rack, which may need to be sequentially recognized and rejected before the target is eventually found. While stable templates may be set up more quickly, this frequent misguidance may increase overall search time.

distributed across possible target features. For instance, when asked to search for a shape-defined target, which across trials is more likely to be blue than red, searchers will update their template accordingly, resulting in faster and more efficient attentional guidance for shapes in the expected color [29–33]. While much of this research has focused on simple, low-dimensional stimuli (e.g., colored, oriented lines), recent evidence suggests that attentional templates for highly familiar, naturalistic object categories (e.g., cars, animals, and tools) can be similarly updated to include, for example, the color or orientation of recently seen exemplars [34–36].

For efficient search, flexible templates should not only reflect the most likely target features but rather the most diagnostic ones—features that set the target apart from expected distractors, as in the example of searching for apples among similarly shaped and sized oranges above. Importantly, while expectations about target features increase the template's precision and veridicality, distractor expectations may distort this representation, potentially making the template less similar to the target [17,37–39]. One way of increasing the distinction between the template and distractors is by adapting the attentional weights of different target features or entire feature dimensions (e.g., color, irrespective of specific feature values such as red), thereby prioritizing features that are unique to the target and deprioritizing those that are not [37,40–44]. In some especially difficult searches, the most diagnostic features may not be veridical target features at all [17]. For instance, when searching for the ripest apple in a supermarket display full of apples of similar shape and size, reducing template-distractor overlap is best achieved by enhancing sensory gain away from distractors (e.g., by searching for a darker shade of red than any of the apples, including the target) [45]. Such a template-shift away from distractors has been found across a broad range of targets, defined by orientation [46,47], color [48–50], dynamic properties such as motion direction or speed [51], and even emotional facial expression [52].

In most real-world situations, distractors are not fully predictable in terms of their low-level features. Search performance for naturalistic objects, however, still improves when targets are consistently paired with distractors varying in low-level features but drawn from a unique category (e.g., searching for mammals among reptiles) [53–55]. This suggests that attentional templates for real-world objects can also be adapted to distractor context, likely reflecting a change in template precision and reliance on different mid- or high-level features depending on expected distractors.

Overall, these studies highlight our ability to quickly extract consistent target and distractor statistics to support efficient search. Contextual expectations may refine attentional templates for simple, low-dimensional stimuli but also for multifeature, familiar objects. It should be noted, however, that optimizing attentional templates in this way may only be adaptive—or even possible—when targets or target–distractor relationships remain consistent across immediately repeating searches, a rare case in daily-life search, where we instead typically search for various targets in changing environments.

Flexible templates adapt to scene context

Moving beyond artificially arranged, isolated search arrays, real-world scenes provide a wealth of contextual information that can facilitate search and object recognition more broadly [56–58]. Naturalistic scenes change the type of contextual regularities that can be exploited and the timeframe over which these regularities are learned, ranging from a single experimental session to potentially a lifetime. Many studies have shown that naturalistic scenes provide powerful spatial guidance, directing attention toward likely target locations (e.g., the sky when searching for a bird). This spatial scene guidance is distinct from the guidance provided by the attentional template [23–25]. Interestingly, recent evidence suggests that scene-based expectations can also directly interact with and flexibly update attentional templates. Specifically, recent studies have shown multiple ways in which scene context (provided by scene previews not yet including the target [59–62]) can be used to optimize attentional templates for the current situational demands.

First, natural scenes can support the structuring and retrieval of context-specific memories for real-world regularities that are highly consistent within contexts but vary across contexts (e.g., the color of taxis in different countries). Searchers can, for instance, learn associations between features and specific object categories [35] or rewards [63] in such a hierarchically structured manner, biasing attention toward different features based on the associated scene preview.

Second, the structure of real-world scenes dictates which objects are typically found within a given scene and how they relate to one another [21,23]. For example, when searching for small and less salient targets, attention may first need to be directed to larger objects that are associated with the target (e.g., the kitchen counter on which apples are likely to be found) [61,64–71], suggesting that these objects may be reflected in the template even though they are not targets themselves. Recent studies support this idea, showing that neural representations of these guiding objects are activated already prior to search, in addition to—or instead of—specific target features [61,64], driving the first fixation in the scene. Interestingly, this was also the case when object associations were scene-specific, suggesting that preparatory search biases can be highly flexible [61].

Finally, and perhaps most strikingly, naturalistic environments directly constrain target appearance. Accordingly, when looking for something to sit on in the kitchen, attention may be guided toward a simple wooden chair but not an armchair [62]. Importantly, even for the very same object, the retinal image it projects changes predictably depending on its position in the scene, for example, being larger when the object is nearby compared with far away, and similarly changing with viewpoint or lighting. Breaking such canonical scene–object relationships impairs detection [72], potentially leading participants to even miss giant targets when their size is inconsistent with their surroundings (e.g., a toothbrush as large as the sink [73]). To account for changes in retinal size, attentional templates may be flexibly rescaled depending on viewing distance. Support for this idea comes from studies showing that preparatory target representations in visual cortex reflect the object's expected retinal size at the current location [59] and that attention is guided specifically toward size-matching objects [60]. Finally, when observers (or objects) move, flexible templates would have to be updated dynamically. A recent study provided evidence that, when viewing a moving scene, neural representations of occluded objects within object-selective cortex were updated and rotated congruently with their surrounding context [74]. While this particular study did not require participants to search, the findings suggest that attentional templates could be similarly adapted to account for dynamic changes in viewpoint or distance.

Altogether, the attention system can leverage our extensive familiarity with the structure of real-world scenes to not only flexibly determine where we search [23–25] but also what we search for.

When one size fits all: guidance by stable templates

While naturalistic environments provide a wealth of regularities that allow searchers to anticipate target-diagnostic features, many aspects of target appearance remain uncertain and only partially predictable from the current context. Rather than dynamically optimizing and adapting individual features and attentional weights to ever-changing contexts, searchers may alternatively rely on context-invariant templates, prioritizing those features that remain stable across contexts (e.g., an apple is typically red and round, irrespective of its retinal size, specific hue, and the context-dependent diagnosticity of these features). Compared with highly context-specific flexible templates, stable templates thus reflect a 'one-size-fits-all' approach (Figure 1), optimized for generalizability across contexts and target occurrences. The efficacy of such templates does not depend on the availability and precision of context expectations, and they guide attention to the target even when specific features remain unknown, are highly variable, or unexpected in the current context.

Stable templates emphasize context-invariant features

To provide accurate, context-independent approximations of target appearance, stable templates should preferentially reflect those features that remain diagnostic across many different occurrences. In line with this, studies have shown that searchers are not only sensitive to specific (or average) target features but also to their variability [27,75]. Importantly, more consistent (i.e., less variable) features are preferentially reflected in the template compared with less consistent ones [75–77]. Furthermore, while low target variability results in highly specific templates, searchers adopt a broader, more general template when target features are more variable and uncertain [36,78].

In the real world, we often search for visually heterogeneous objects or object categories for which the specific features within a scene are highly variable and cannot all be accurately predicted from context (e.g., searching for any car, where color, shape, or distance are unknown in advance). In these cases, searchers rely on relatively broad templates reflecting category-diagnostic mid- or high-level visual features [76,77,79–81]. Such high-level templates are derived from cortical object representations that are tuned to objects across a wide variety of exemplars and contexts, particularly for familiar object categories [22,82–84]. These object representations are shaped by their common contexts and can, for instance, reflect specifically those features that distinguish a given category from its most common distractors [54,85]. In addition, they are influenced by long-term object associations and canonical object–context relationships [86,87]. Despite not being dynamically adapted to the current context, stable templates for real-world objects are therefore still shaped by long-term statistics of naturalistic environments and incorporate features that are diagnostic in most (i.e., default) contexts. While a stable template would, for instance, not rely on context-specific object associations, it could still guide attention toward objects that frequently co-occur with the target across contexts [65,67,70,71,88,89].

Attentional guidance based on these high-level templates will generally be less precise compared with when context-specific low-level features of the target are known (Figure 2). Importantly, however, stable high-level templates are well suited to account for changes across exemplars, size, or viewpoint, or to deal with partial occlusion. For instance, they have been shown to include a range of diagnostic object parts [79], allowing for attentional selection even if parts of the object are occluded or not visible from the current perspective. Similarly, they may, at least to some degree, be size- and viewpoint-invariant, such that attention is still guided to shape-matching objects of unexpected size or orientation [60,79,90–92 but see 73]. Neuroimaging evidence further suggests that searching for a given category in visually heterogeneous scenes activates neural populations tuned to objects of the target category across a range of exemplars, poses, and sizes [80,84,93,94] but see [59]. Stable templates also apply across a large search space within a scene, guiding attention to highly familiar categories (e.g., bodies and faces) even when they appear at unexpected locations [83,84,90,94].

Stable, context-invariant templates therefore support efficient detection across a wide range of visually distinct exemplars, scenes, and viewpoints, even in the absence of specific contextual expectations for example, in very briefly flashed scenes [17,80,92,93]. This strongly contributes to the efficiency of naturalistic visual search, without requiring the template to be flexibly updated based on the specific context.

Stable templates can act on context-modulated object representations

While we have focused on how attentional templates incorporate contextual regularities, scene context facilitates and modulates visual processing also for objects that are not current search targets. For example, context creates expectations about the identity of objects at specific positions in a scene (e.g., a distant blob on the road is likely a car). These expectations can be used to inform the attentional template, as reviewed earlier, but also directly modulate visual processing during perceptual encoding [56–58,95], both for attended and unattended objects [96,97]. This raises the possibility that stable attentional templates operate on object representations after they are modulated by context, thereby reducing the need to incorporate contextual expectations proactively and flexibly in the attentional template.

This idea is supported by a recent study [98], in which two identically sized objects were placed at near and far locations in a scene (with the near object perceived as larger due to size constancy [99]). The near object only captured attention when the template consisted of a large object, and the distant object only captured attention when the template was a small object, indicating that distance inferred from scene context modulated these object representations before attentional selection. These results highlight a visual search mechanism that relies on a stable template but is nonetheless context-sensitive.

Balancing flexibility and stability

Flexible and stable templates offer varying costs and benefits and place distinct demands on the attention system. How are flexibility and stability balanced for effective attentional guidance?

When searching for familiar objects in real-world scenes, stable templates often provide useful defaults, especially when targets are characterized by context-invariant diagnostic features and when contextual priors are either weak or already incorporated during sensory encoding. While growing evidence suggests that templates can be further refined based on contextual expectations, there is clear evidence that searchers often do not do so, even when given the opportunity [42,46,47,54,69], suggesting that flexibility may come at a cost.

Indeed, adapting the template to frequently varying contexts can be effortful and computationally costly (Box 2). Flexibility involves additional control processes, deeper processing of the context [108], and the retrieval of context-associations from memory [109]. In addition, setting up a precise, detailed template may take longer than a coarser one [110] and the act of changing

Box 2. Automatic and controlled updating of the template

Attentional templates have been shown to flexibly adapt to current context. Is this an automatic process, occurring with little control and cognitive effort, or does it instead reflect a more controlled and effortful process?

To establish contextual expectations, many studies reviewed here asked participants to repeatedly search for the same or similar targets within relatively unchanging contexts over a large number of trials. Within such blocked contexts, target and distractor statistics can often be quickly learned and adapted to, also without explicit instructions to do so [28,40,42,100]. These adaptations may, however, be partially explained by **selection history** or priming [101,102]—relatively strong attentional biases toward recent target features or feature dimensions, and away from previous distractors. These biases can be highly adaptive in repeated searches [28,103] but arise in the absence of strategic control or even awareness. They also often quickly fade away after a change in context [34,42,53] (although sometimes persisting over a long time [30]). Importantly, however, the degree to which regularities in repeated search are exploited still depends on their usefulness for search and participants' strategy [37]. In contrast to repeated search in unchanging contexts, adapting the template on a moment-to-moment basis based on explicit trial-by-trial context cues in the lab requires stronger top-down control [104–106].

Flexibly adapting to hierarchically structured and rule-based regularities during real-world search can sometimes be a similarly voluntary and thus effortful process; for example, when your friend reminds you that taxis in London are black and not yellow as in Germany, when stepping out of the airport. With sufficient experience, we may, however, begin to automatically anticipate an either black or yellow cab, based on rich and multimodal context cues offered by naturalistic contexts (e.g., the language of the announcements or the brands of shops we encounter in the airport). Adapting the template to highly familiar scene–object relationships may then become relatively automatic and could even be taken into account when they are not useful for the current task [62,107].

For many contextual regularities, the degree to which we automatically and flexibly adapt attentional templates to them, that is, whether this occurs also when not strictly necessary for the task or when resources are low, has not yet been systematically investigated (see Outstanding questions). Generally, however, this will depend on the specific type, consistency, and complexity of the regularity involved, as well as our experience with it.

the template itself may lead to additional switch costs [111,112]. Even when aware of relevant contextual regularities, searchers may therefore not adapt their template based on trial-by-trial context cues [42,46,47,54]. Also outside of search tasks, attentional control settings are often not flexibly updated on the fly but rather based on continuous task experience [113], suggesting that the moment-to-moment flexibility of attentional templates is limited.

Further limiting the flexible updating of templates is the difficulty of learning context-specific regularities in the first place. While consistent feature and co-occurrence statistics (e.g., fire trucks are red or pens are found on office desks) are learned quickly, regularities that vary across time or contexts (e.g., the varying taxi color across countries) are slower to learn and retrieve [114,115]. While such context-specific memories can be used to guide attention [35,42,63], and their formation and retrieval may be facilitated when rich, naturalistic environments serve as context cues, this is likely still slower and more difficult compared with highly stable statistics.

Thus, while attentional templates can be relatively easily (potentially automatically) adapted to regularities that are highly consistent in the short or long term, many other regularities are only relied upon when the incentives are high enough or when it is important to find the target quickly (e.g., looking for our lost child in a busy park). This view is supported by studies showing that attentional templates incorporate expectations about changing distractor features or object associations specifically in especially difficult searches, such as when distractor interference is high [46,47] or when other (context-independent) features cannot be relied upon [37]. In such cases, searchers spend additional time and effort refining their template, allowing for more efficient guidance and reduced interference by distractors.

More generally, the degree to which attentional templates are flexibly adapted in a controlled (i.e., effortful) manner will be determined by the cost of setting up a context-specific template relative to the cost and likelihood of potential misguidance (e.g., fixating on a nontarget object). Considering the relatively low cost of making eye movements (but see [116]) and the strong cues already provided by spatial scene layout, attentional templates typically only need to be 'good enough' [17]. We propose that the range of regularities searchers spontaneously adapt their templates to in most daily life searches is likely limited to overlearned regularities (e.g., the relationship between size and distance) and regularities that are highly consistent in the short term (e.g., across repeated searches).

Generalizing to a dynamically changing world

Though studies use increasingly more naturalistic stimulus material, almost all of the current evidence for the flexibility or stability of attentional templates is based on simplified lab tasks, which still differ from truly naturalistic searches in several ways. Even when specific targets and distractors vary on a trial-by-trial basis, typical lab-based search tasks still involve many repetitions of a limited number of stimuli and often highly consistent contextual regularities, favoring the reliance on a context-dependent template. By contrast, real-world search is often one-shot; that is, we only search for an object within a specific context once before moving on to search for another object. Consequently, searchers rarely benefit from expectations about targets, target–distractor relations, or feedback from immediately preceding searches to iteratively fine-tune their template. Flexible real-world templates would thus necessarily need to rely more on either long-term experience or analysis of the immediate scene context instead.

In addition, lab experiments typically encourage participants to put maximum effort into the task for fast and accurate search, or even require the use of flexible attentional templates to find the target, due to meticulously chosen foils or extremely rapid presentation times. Real-world search typically does not operate under the same constraints, and both effort and reliance on memory are often minimized within these more naturalistic settings [103,117,118]. And while individual glances (fixations) during real-world viewing may be comparable in duration to these rapid stimulus presentations, the visual input from an individual fixation period is often more predictable than the content of the search display in a randomized experimental design.

Another important difference between lab-based search and real-world search is that searchers in the real world move and interact with the objects around them. This allows them, for instance, to move their heads or bodies to reduce occlusion and view the object from a preferred viewpoint [119], thereby reducing the need for a flexible template.

Furthermore, in naturalistic settings, objects are not only attended to be identified but are also selected as part of specific actions. Accordingly, attentional templates may emphasize action-relevant features of the target [120]. For example, while driving, an orientation-specific template is useful for spotting pedestrians facing toward the street, whereas orientation is not relevant when searching for a friend [121]. When attending to objects becomes an intrinsic part of behavior (e.g., when driving, shopping, or cooking), feature, spatial, and temporal expectations may also work in concert to efficiently prioritize relevant objects in sequence.

In sum, while attentional templates remain a useful mechanism to select objects in real-world environments, a range of additional factors shapes the balance between flexibility and stability within fully naturalistic and dynamic search. This leaves open important avenues for future research to understand the boundary conditions of this trade-off and the content and neural correlates of naturalistic templates (Box 3). Understanding the efficient prioritization of objects in dynamic, real-world scenes will require new experimental paradigms that simulate the available contextual cues, the time and resource constraints of naturalistic search, and the selection of objects as part of naturalistic action sequences.

Box 3. Combining flexible and stable templates in the brain

One way in which attentional templates could be implemented in the brain is through pre-activation of object representations in visual cortex [5,15,80]. Such a baseline shift in neuronal populations coding for target features has been proposed to constitute a top-down attention signal that biases competition between multiple stimuli in favor of the target [16]. Preparatory activity is often measured before stimulus onset but has also been observed while searchers concurrently view scenes [59,61], continuously biasing the processing of visual input, thereby making it a viable mechanism for real-world search. Interestingly, objects are represented in a highly distributed manner across the cortex, with these representations varying in both precision and task-dependency [122]. For instance, when viewing or remembering an object, regions in the early visual cortex will typically encode an object's low-level features relatively independently of current task demands, while object representations in the parietal and frontal cortex are much more task-dependent [123–125]. Similarly, even within object-selective brain regions such as the lateral occipital complex (LOC), subregions differ substantially in their degree of invariance (e.g., with posterior-dorsal LOC responses being mostly size-specific, and anterior-ventral LOC responses being mostly size-invariant) [126]. Not all of these representations may be equally able or useful to affect early attentional guidance and sensory gain as attentional templates. For instance, in order to guide spatial attention, object feature representations may need to be topographically organized in the visual cortex [127]. Importantly, however, this distributed organization may allow for the coexistence of both context-specific and context-invariant template features in the brain [128].

Behavioral studies investigating visual search for objects in scenes provide some evidence that templates can consist of multiple feature representations. For example, when looking for a specific object at an unknown location (or at an entire scene at once), attentional templates were shown to be highly shape-specific but location invariant [90], akin to classic feature-based attention [127]. By contrast, when looking for a person or object at a specific location, this creates strong priors about their expected (object-specific) shape as well as their (distance-dependent) size [60], while other properties (such as the direction they are facing) remain uncertain. Such situations may result in a size-dependent, but rotationally invariant shape template [79,91]. Therefore, attentional templates may, in some cases, simultaneously combine specificity (for known features) with invariance (for unknown features), bridging multiple levels of representation to fully capitalize on available contextual information.

Concluding remarks

(How) are short- and long-term statistics of the environment incorporated into the attentional template for efficient search in the real world? The lab-based studies reviewed here provide convincing evidence that top-down attentional templates 'can' be flexibly updated based on short- and long-term statistics of the environment. This flexible account of attentional selection aligns well with the general view that prediction is ubiquitous in perception [129]. However, lab-based experiments may have overestimated the moment-to-moment flexibility of templates and the need for highly flexible templates in real-world search. For many real-world searches, relying on stable templates that incorporate long-term priors and act on context-modulated object representations may already provide an efficient selection mechanism. At the same time, the many regularities characterizing real-world environments offer ample opportunities to adapt the template, and searchers likely use these opportunities partly automatically when they are highly consistent and overlearned, and, when needed, also in a more controlled manner (Boxes 2 and 3). The degree to which templates are flexibly adjusted to the current context depends on the associated computational costs relative to the costs and likelihood of potential misguidance. Moving beyond the use of naturalistic scenes, future studies should examine attentional templates in truly naturalistic search conditions, where searchers can freely move and where searching for objects is an intrinsic part of behavior (e.g., preparing a sandwich or riding a bike) [118] (see Outstanding questions).

Acknowledgments

M.V.P. and F.P.d.L. received support from the Dutch Research Council (NWO-VICI: 231-057 and NWO-VICI: 231-043, respectively).

Declaration of interests

The authors declare no competing interests.

Outstanding questions

What are the costs associated with generating and adjusting (flexible) attentional templates, and how can these costs be measured (e.g., using pupillometry as an index of cognitive effort)?

Do we rely more on stable (and less on flexible) templates when cognitive resources are exhausted, such as when concurrently performing another challenging task?

How does the context-dependency of target representations evolve across the entire search process, from search preparation to object identification?

How stable (or flexible) are target templates that are used for identification?

During real-world search, humans may adapt their attentional template to the current scene context, but they can also move freely, thereby predictably altering their viewpoint on the scene. How does this impact the use and updating of attentional templates?

How do attentional templates and their neural correlates differ for searches that are externally cued (as is the case in most lab-based experiments) and searches that are internally cued (e.g., as part of naturalistic action sequences such as making coffee)?

Can flexible and stable templates operate in parallel, and do they bias visual processing at different timepoints? For instance, does a viewpoint-specific flexible template affect sensory processing earlier than an invariant stable template?

Can attentional templates shift between stable and flexible versions during the search process, for example, when an initially stable template turns out to be too generic, guiding attention toward too many distractor objects?

How does the trade-off between stability and flexibility change across the lifespan, among individuals, and across different environments?

References

1. Downing, P.E. (2000) Interactions between visual working memory and selective attention. *Psychol. Sci.* 11, 1182–1191
2. van der Stigchel, S. and Hollingworth, A. (2018) Visuospatial working memory as a fundamental component of the eye movement system. *Curr. Dir. Psychol. Sci.* 27, 136–143
3. Gayet, S. *et al.* (2013) Information matching the content of visual working memory is prioritized for conscious access. *Psychol. Sci.* 24, 2472–2480
4. van Moorselaar, D. *et al.* (2018) Competitive interactions in visual working memory drive access to awareness. *Cortex* 102, 6–13
5. Battistoni, E. *et al.* (2017) Preparatory attention in visual cortex. *Ann. N. Y. Acad. Sci.* 1396, 92–107
6. Stokes, M.G. (2011) Top-down visual activity underlying VSTM and preparatory attention. *Neuropsychologia* 49, 1425–1427
7. Olivers, C.N.L. *et al.* (2011) Different states in visual working memory: when it guides attention and when it does not. *Trends Cogn. Sci.* 15, 327–334
8. Cabbai, G. *et al.* (2023) Mental imagery and visual attentional templates: a dissociation. *Cortex* 169, 259–278
9. Keogh, R. and Pearson, J. (2017) The perceptual and phenomenal capacity of mental imagery. *Cognition* 162, 124–132
10. Monzel, M. *et al.* (2021) Imagine, and you will find – lack of attentional guidance through visual imagery in aphantasics. *Atten. Percept. Psychophysiol.* 83, 2486–2497
11. Milton, F. *et al.* (2021) Behavioral and neural signatures of visual imagery vividness extremes: aphantasia versus hyperphantasia. *Cereb. Cortex Commun.* 2, tgab035
12. Reeder, R.R. (2017) Individual differences shape the content of visual representations. *Vis. Res.* 141, 266–281
13. Eimer, M. (2014) The neural basis of attentional control in visual search. *Trends Cogn. Sci.* 18, 526–535
14. Wolfe, J.M. (2020) Visual search: how do we find what we are looking for? *Annu. Rev. Vis. Sci.* 6, 539–562
15. Chelazzi, L. *et al.* (1993) A neural basis for visual search in inferior temporal cortex. *Nature* 363, 345–347
16. Desimone, R. (1995) Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222
17. Yu, X. *et al.* (2023) Good-enough attentional guidance. *Trends Cogn. Sci.* 27, 391–403
18. Wolfe, J.M. (2021) Guided Search 6.0: an updated model of visual search. *Psychon. Bull. Rev.* 28, 1060–1092
19. Kirchner, H. and Thorpe, S.J. (2006) Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. *Vis. Res.* 46, 1762–1776
20. Wolfe, J.M. *et al.* (2011) Visual search for arbitrary objects in real scenes. *Atten. Percept. Psychophysiol.* 73, 1650–1671
21. Kaiser, D. *et al.* (2019) Object vision in a structured world. *Trends Cogn. Sci.* 23, 672–685
22. Peelen, M.V. and Kastner, S. (2014) Attention in the real world: toward understanding its neural basis. *Trends Cogn. Sci.* 18, 242–250
23. Võ, M.L.H. *et al.* (2019) Reading scenes: how scene grammar guides attention and aids perception in real-world environments. *Curr. Opin. Psychol.* 29, 205–210
24. Wolfe, J.M. *et al.* (2011) Visual search in scenes involves selective and nonselective pathways. *Trends Cogn. Sci.* 15, 77–84
25. Castelhano, M.S. and Krzyś, K. (2020) Rethinking space: a review of perception, attention, and memory in scene processing. *Annu. Rev. Vis. Sci.* 9, 1–24
26. Nobre, A.C. and van Ede, F. (2018) Anticipated moments: temporal structure in attention. *Nat. Rev. Neurosci.* 19, 34–48
27. Chetverikov, A. *et al.* (2019) Feature distribution learning (FDL): a new method for studying visual ensembles perception with priming of attention shifts. In *Spatial Learning and Attention Guidance* (Vol. 151) (Pollmann, S., ed.), pp. 37–57, Springer US

28. Kristjánsson, Á. (2023) Priming of probabilistic attentional templates. *Psychon. Bull. Rev.* 30, 22–39
29. Ortego, K. *et al.* (2025) Early cortical sensitivity and speeded target selection underlie incidentally learned prioritization of visual features. *J. Neurosci.* 45, e0607252025
30. Kruijne, W. and Meeter, M. (2016) Long-term priming of visual search prevails against the passage of time and counteracting instructions. *J. Exp. Psychol. Learn. Mem. Cogn.* 42, 1293–1303
31. Sha, L.Z. *et al.* (2017) Short-term and long-term attentional biases to frequently encountered target features. *Atten. Percept. Psychophysiol.* 79, 1311–1322
32. Grubert, A. and Eimer, M. (2023) Do we prepare for what we predict? How target expectations affect preparatory attentional templates and target selection in visual search. *J. Cogn. Neurosci.* 35, 1919–1935
33. Anderson, B.A. (2026) Statistical learning and the efficiency of visual search. *Atten. Percept. Psychophysiol.* 88, 47
34. 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
35. 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. Psychophysiol.* 84, 1304–1316
36. Addleman, D.A. *et al.* (2024) Attention to object categories: selection history determines the breadth of attentional tuning during real-world object search. *J. Exp. Psychol. Gen.* 153, 1568–1581
37. Yu, X. *et al.* (2024) Task-adaptive changes to the target template in response to distractor context: separability versus similarity. *J. Exp. Psychol. Gen.* 153, 564–572
38. Geng, J.J. *et al.* (2019) Distractor ignoring: strategies, learning, and passive filtering. *Curr. Dir. Psychol. Sci.* 28, 600–606
39. van Moorselaar, D. and Slagter, H.A. (2020) Inhibition in selective attention. *Ann. N. Y. Acad. Sci.* 1464, 204–221
40. Lee, J. and Geng, J.J. (2020) Flexible weighting of target features based on distractor context. *Atten. Percept. Psychophysiol.* 82, 739–751
41. Boettcher, S.E.P. *et al.* (2020) Functional biases in attentional templates from associative memory. *J. Vis.* 20, 1–10
42. Lerebourg, M. *et al.* (2023) Expected distractor context biases the attentional template for target shapes. *J. Exp. Psychol. Hum. Percept. Perform.* 49, 1236–1255
43. Alexander, R.G. *et al.* (2019) Specifying the precision of guiding features for visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 45, 1248–1264
44. Xu, Z.J. *et al.* (2024) Top-down instructions influence the attentional weight on color and shape dimensions during bidimensional search. *Sci. Rep.* 14, 31376
45. Navalpakkam, V. and Itti, L. (2007) Search goal tunes visual features optimally. *Neuron* 53, 605–617
46. Scolari, M. and Serences, J.T. (2009) Adaptive allocation of attentional gain. *J. Neurosci.* 29, 11933–11942
47. Scolari, M. *et al.* (2012) Optimal deployment of attentional gain during fine discriminations. *J. Neurosci.* 32, 7723–7733
48. Becker, S.I. *et al.* (2023) Tuning to non-veridical features in attention and perceptual decision-making: an EEG study. *Neuropsychologia* 188, 108634
49. Kerzel, D. and Huynh Cong, S. (2024) Trial history contributes to the optimal tuning of attention. *J. Exp. Psychol. Hum. Percept. Perform.* 50, 1000–1009
50. Becker, S.I. *et al.* (2025) Visual search is relational without prior context learning. *Cognition* 260, 106132
51. Boettcher, S.E.P. and Nobre, A.C. (2025) Going through the motions: Biasing of dynamic attentional templates. *Journal of Experimental Psychology: General* 154, 111–127
52. Won, B.Y. *et al.* (2020) Flexible target templates improve visual search accuracy for faces depicting emotion. *Atten. Percept. Psychophysiol.* 82, 2909–2923
53. Lau, J.S.H. *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. Psychophysiol.* 83, 1435–1454
54. Bravo, M.J. and Farid, H. (2016) Observers change their target template based on expected context. *Atten. Percept. Psychophysiol.* 78, 829–837
55. Bravo, M.J. and Farid, H. (2012) Task demands determine the specificity of the search template. *Atten. Percept. Psychophysiol.* 74, 124–131
56. Peelen, M.V. *et al.* (2023) Predictive processing of scenes and objects. *Nat. Rev. Psychol.* 3, 13–26
57. Bar, M. (2004) Visual objects in context. *Nat. Rev. Neurosci.* 5, 617–629
58. Oliva, A. and Torralba, A. (2007) The role of context in object recognition. *Trends Cogn. Sci.* 11, 520–527
59. Gayet, S. and Peelen, M.V. (2022) Preparatory attention incorporates contextual expectations. *Curr. Biol.* 32, 687–692.e6
60. Gayet, S. *et al.* (2024) Searching near and far: the attentional template incorporates viewing distance. *J. Exp. Psychol. Hum. Percept. Perform.* 50, 216–231
61. Lerebourg, M. *et al.* (2024) Attentional guidance through object associations in visual cortex. *Sci. Adv.* 10, eado6226
62. 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
63. Anderson, B.A. (2015) Value-driven attentional priority is context specific. *Psychon. Bull. Rev.* 22, 750–756
64. Zhou, Z. and Geng, J. (2025) Preparatory attentional templates in prefrontal and sensory cortex encode target-associated information. *Elife* 14, RP104041
65. 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, 1–13
66. 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
67. 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, 1–16
68. Koehler, K. and Eckstein, M.P. (2017) Beyond scene gist: objects guide search more than scene background. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 1177–1193
69. Zhou, Z. and Geng, J.J. (2024) Learned associations serve as target proxies during difficult but not easy visual search. *Cognition* 242, 105648
70. Souza-Wiggins, M. and Geng, J.J. (2025) Anchor objects guide spatial attention during visual search. *Atten. Percept. Psychophysiol.* 88, 19
71. Yeh, L.-C. *et al.* (2025) Contextual associations impact visual search across multiple processing stages. *bioRxiv* https://doi.org/10.64898/2025.12.12.693961
72. Biederman, I. *et al.* (1982) Scene perception: detecting and judging objects undergoing relational violations. *Cognit. Psychol.* 14, 143–177
73. Eckstein, M.P. *et al.* (2017) Humans, but not deep neural networks, often miss giant targets in scenes. *Curr. Biol.* 27, 2827–2832.e3
74. Aldegheri, G. *et al.* (2026) Dynamic context–based updating of object representations in the visual cortex. *Sci. Adv.* 12, eadw6726
75. Witkowski, P.P. and Geng, J.J. (2019) Learned feature variance is encoded in the target template and drives visual search. *Vis. Cogn.* 27, 487–501
76. 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. Psychophysiol.* 79, 1578–1592
77. Yu, C.-P. *et al.* (2016) Searching for category-consistent features: a computational approach to understanding visual category representation. *Psychol. Sci.* 27, 870–884
78. Ajith, S. *et al.* (2026) Visual search is constrained by the variability of object-category templates. *bioRxiv* https://doi.org/10.64898/2026.02.02.702780
79. Reeder, R.R. and Peelen, M.V. (2013) The contents of the search template for category-level search in natural scenes. *J. Vis.* 13, 1–13

80. Peelen, M.V. and Kastner, S. (2011) A neural basis for real-world visual search in human occipitotemporal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 108, 12125–12130
81. Yang, H. and Zelinsky, G.J. (2009) Visual search is guided to categorically-defined targets. *Vis. Res.* 49, 2095–2103
82. Brady, T.F. (2019) Scaling up visual attention and visual working memory to the real world. In *Psychology of Learning and Motivation* (Vol. 70) (Federmeier, K.D. *et al.*, eds), pp. 29–69, Academic Press
83. Störmer, V.S. *et al.* (2019) Tuning attention to object categories: spatially global effects of attention to faces in visual processing. *J. Cogn. Neurosci.* 31, 937–947
84. Thorat, S. and Peelen, M.V. (2022) Body shape as a visual feature: evidence from spatially-global attentional modulation in human visual cortex. *Neuroimage* 255, 119207
85. Sigala, N. and Logothetis, N.K. (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415, 318–320
86. Kaiser, D. and Peelen, M.V. (2018) Transformation from independent to integrative coding of multi-object arrangements in human visual cortex. *Neuroimage* 169, 334–341
87. Welbourne, L.E. *et al.* (2021) The transverse occipital sulcus and intraparietal sulcus show neural selectivity to object-scene size relationships. *Commun. Biol.* 4, 768
88. Moores, E. *et al.* (2003) Associative knowledge controls deployment of visual selective attention. *Nat. Neurosci.* 6, 182–189
89. Bahle, B. *et al.* (2025) Combined conceptual and perceptual control of visual attention in search for real-world objects. *Atten. Percept. Psychophys.* 88, 59
90. Reeder, R.R. *et al.* (2015) Involuntary attentional capture by task-irrelevant objects that match the search template for category detection in natural scenes. *Atten. Percept. Psychophysiol.* 77, 1070–1080
91. Bravo, M.J. (2009) The specificity of the search template. *J. Vis.* 9, 1–9
92. Zhang, M. *et al.* (2018) Finding any Waldo with zero-shot invariant and efficient visual search. *Nat. Commun.* 9, 3730
93. Soon, C.S. *et al.* (2013) Preparatory patterns of neural activity predict visual category search speed. *Neuroimage* 66, 215–222
94. Peelen, M.V. *et al.* (2009) Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* 7251, 94–97
95. Peelen, M.V. (2025) The neural basis of visual search in scene context. *Curr. Dir. Psychol. Sci.* 34, 114–121
96. Leticevscaia, O. *et al.* (2024) Scene context and attention independently facilitate MEG decoding of object category. *Vis. Res.* 224, 108484
97. Munneke, J. *et al.* (2013) The influence of scene context on object recognition is independent of attentional focus. *Front. Psychol.* 4, 552
98. Gayet, S. and Peelen, M.V. (2019) Scenes modulate object processing before interacting with memory templates. *Psychol. Sci.* 30, 1497–1509
99. Sperandio, I. *et al.* (2012) Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. *Nat. Neurosci.* 15, 540–542
100. Hout, M.C. and Goldinger, S.D. (2010) Learning in repeated visual search. *Atten. Percept. Psychophysiol.* 72, 1267–1282
101. Awh, E. *et al.* (2012) Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends Cogn. Sci.* 16, 437–443
102. Maljkovic, V. and Nakayama, K. (1994) Priming of pop-out: 1. Role of features. *Mem. Cognit.* 22, 657–672
103. Kristjánsson, Á. and Draschkow, D. (2021) Keeping it real: looking beyond capacity limits in visual cognition. *Atten. Percept. Psychophysiol.* 83, 1375–1390
104. Schneider, W. and Shiffrin, R.M. (1977) Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol. Rev.* 84, 1–66
105. Goldstein, R.R. and Beck, M.R. (2018) Visual search with varying versus consistent attentional templates: effects on target template establishment, comparison, and guidance. *J. Exp. Psychol. Hum. Percept. Perform.* 44, 1086–1102
106. Theeuwes, J. (2018) Visual selection: usually fast and automatic; seldom slow and volitional. *J. Cogn.* 1, 1–15
107. Aldegheri, G. *et al.* (2023) Scene context automatically drives predictions of object transformations. *Cognition* 238, 105521
108. Hansen, H.A. *et al.* (2019) Taking stock: the role of environmental appraisal in the strategic use of attentional control. *Atten. Percept. Psychophysiol.* 81, 2673–2684
109. Le-Hoa Võ, M. and Wolfe, J.M. (2015) The role of memory for visual search in scenes. *Ann. N. Y. Acad. Sci.* 1339, 72–81
110. Meyyappan, S. *et al.* (2025) Hierarchical organization of human visual feature attention control. *J. Neurosci.* 45, e2073242025
111. Monsell, S. (2003) Task switching. *Trends Cogn. Sci.* 7, 134–140
112. Grubert, A. (2024) Target switch costs in visual search arise during the preparatory activation of target templates. *Psychophysiology* 61, e14658
113. Braem, S. *et al.* (2024) One cannot simply 'be flexible': regulating control parameters requires learning. *Curr. Opin. Behav. Sci.* 55, 101347
114. Collins, A.G.E. (2017) The cost of structure learning. *J. Cogn. Neurosci.* 29, 1646–1655
115. Frost, R. *et al.* (2025) Statistical learning subserves a higher purpose: novelty detection in an information foraging system. *Psychol. Rev.* 133, 237–252
116. Schütz, A.C. and Stewart, E.E.M. (2025) A review of the costs of eye movements. *Nat. Rev. Psychol.* 4, 625–638
117. Kumle, L. *et al.* (2025) Sensorimnemonic decisions: choosing memories versus sensory information. *Trends Cogn. Sci.* 29, 311–313
118. Ballard, D.H. *et al.* (1995) Memory representations in natural tasks. *J. Cogn. Neurosci.* 7, 66–80
119. Wu, T.C. and Tsotsos, J.K. (2025) Real-world visual search goes beyond eye movements: active searchers select 3D scene viewpoints too. *PLoS One* 20, e0319719
120. Trentin, C. *et al.* (2023) Visual working memory representations bias attention more when they are the target of an action plan. *Cognition* 230, 105274
121. Britt, N. *et al.* (2025) Context-dependent modulation of spatial attention: prioritizing behaviourally relevant stimuli. *Cogn. Res. Princ. Implic.* 10, 4
122. Kourtzi, Z. and Connor, C.E. (2011) Neural representations for object perception: structure, category, and adaptive coding. *Annu. Rev. Neurosci.* 34, 45–67
123. Xu, Y. (2018) A tale of two visual systems: invariant and adaptive visual information representations in the primate brain. *Annu. Rev. Vis. Sci.* 4, 311–336
124. Bracci, S. *et al.* (2017) Task context overrules object- and category-related representational content in the human parietal cortex. *Cereb. Cortex* 27, 310–321
125. Christophel, T.B. *et al.* (2017) The distributed nature of working memory. *Trends Cogn. Sci.* 21, 111–124
126. Eger, E. *et al.* (2008) Graded size sensitivity of object-exemplar-evoked activity patterns within human LOC subregions. *J. Neurophysiol.* 100, 2038–2047
127. Maunsell, J.H.R. and Treue, S. (2006) Feature-based attention in visual cortex. *Trends Neurosci.* 29, 317–322
128. Schulz, M.-C. *et al.* (2024) Parallel gain modulation mechanisms set the resolution of color selectivity in human visual cortex. *Sci. Adv.* 10, eadm7385
129. Lange, F.P.D. *et al.* (2018) How do expectations shape perception? *Trends Cognit. Sci.* 22, 764–779