Habitual versus goal-driven attention

Yuhong V. Jiang

Department of Psychology
University of Minnesota

For the special issue on the “Unconscious Guidance of Attention” (CORTEX-D-17-00091R2)

Abstract
Recent research has expanded the list of factors that control spatial attention. Beside current goals and perceptual salience, statistical learning, reward, motivation and emotion also affect attention. But do these various factors influence spatial attention in the same manner, as suggested by the integrated framework of attention, or do they target different aspects of spatial attention? Here I present evidence that the control of attention may be implemented in two ways. Whereas current goals typically modulate where in space attention is prioritized, search habits affect how one moves attention in space. Using the location probability learning paradigm, I show that a search habit forms when people frequently find a visual search target in one region of space. Attentional cuing by probability learning differs from that by current goals. Probability cuing is implicit and persists long after the probability cue is no longer valid. Whereas explicit goal-driven attention codes space in an environment-centered reference frame, probability cuing is viewer-centered and is insensitive to secondary working memory load and aging. I propose a multi-level framework that separates the source of attentional control from its implementation. Similar to the integrated framework, the multi-level framework considers current goals, perceptual salience, and selection history as major sources of attentional control. However, these factors are implemented in two ways, controlling where spatial attention is allocated and how one shifts attention in space.

Keywords: spatial attention; implicit learning; search habit; spatial reference frame;

1. Introduction
The visual world is complex; not all input can be perceived or acted upon at once. Spatial attention allows one to prioritize locations that are most relevant or significant. Extensive research has examined factors that drive spatial attention, leading to the theoretical development of an attentional “priority map” (Bisley & Goldberg, 2010; Fecteau & Munoz, 2006; Gottlieb & Balan, 2010). Early studies focused on current goals and perceptual salience as factors that influence attentional priority (Egeth & Yantis, 1997). More recent studies have considered selection history, including inter-trial priming, reward learning, statistical learning, and motivational and emotional factors, as a third source of spatial attention (Chelazzi, Perlato, Santandrea, & Della Libera, 2013; Kristjánsson & Campana, 2010; Turk-Browne, 2012; Vuilleumier, 2015). This research cumulates into an integrated framework of spatial attention, in which current goals, perceptual salience, and selection history jointly determine the priority weights one assigns to locations (Awh, Belopolsky, & Theeuwes, 2012). But how are these various sources implemented to control spatial attention? Here I present evidence that spatial attention may be implemented in two ways – by affecting where to attend and how to move attention in space. Whereas the integrated framework primarily explicates the various sources of attentional control, I argue that these sources may be implemented in two ways. Thus, sources of attention are one level of attentional control; implementation of these sources is another (the multi-level framework).

The multi-level framework stems from the observation that spatial attention serves both perception and action. At the perceptual level, spatial attention allows one to filter out irrelevant information from perceptual processing. For example, Neurophysiological studies show that spatial
selection biases neural activity in visual areas of the brain, such as V4, Inferior Temporal Lobe, and the Middle Temporal cortex (Desimone & Duncan, 1995). These findings are extended to early visual areas (such as V1) in human neuroimaging studies (for reviews, see Nobre & Kastner, 2014). Spatial attention also affects action selection such as hand or eye movements. Regions that control covert allocation of attention, such as the Lateral Intraparietal Sulcus and Frontal Eye Fields, are also involved in oculomotor control (Corbetta et al., 1998). The extensive overlap between overt and covert allocation of spatial attention, and the functional consideration of attention, supports the idea that attention is not just for perception, but also for action (Allport, 1989). In fact, the premotor theory of attention proposes that the perceptual and motor functions of attention are integrated (Rizzolatti, Riggio, Dascola, & Umiltá, 1987). For example, it is difficult for one to covertly attend to a location that cannot also be selected by an eye movement (Craighero & Rizzolatti, 2005).

Given that spatial attention serves both perception and action, it is important to consider how space is coded for these functions. Research on the duplex system of vision suggests that spatial information, such as the shape of an object, can be coded by both the perceptual system and the visuomotor system (Goodale & Westwood, 2004). Similarly, spatial locations may be represented in two ways: a map-like code where each location is assigned a Cartesian coordinate, and an action-based code in which locations are specified as vectors of movements (Moore & Fallah, 2001). It follows that the various sources of spatial attention may be implemented in two ways: they may affect where attention is allocated and how one moves attention.

In this review, I provide evidence supporting the idea that spatial attention may be implemented in multiple ways. I propose that frequently attending to a region of space leads to the formation of a search habit – the how of spatial attention. Habit is a prevailing concept in skill learning, executive control, and addiction (Graybiel, 2008; Wood & Rünger, 2016). I suggest that it is also a major way through which the control of spatial attention is implemented. To this end, I will first present the experimental paradigm that produces habitual spatial attention – location probability cuing – followed by an overview of the key features of probability cuing that distinguish it from goal-driven attention. A third section will discuss the nature of attentional allocation in a related but distinct implicit learning paradigm – contextual cuing. Finally, I will develop the multi-level framework of attention and consider how various sources of attention may be implemented.

2. Location probability learning
2.1. Historical overview

As early as the 1970’s, research showed that locations frequently containing a visual search target are prioritized. Shaw and Shaw (1977) asked participants to identify a target letter that may occur in 8 different locations. Some locations were more likely to contain the target than others. Performance was better in the high-probability locations than the low-probability locations. Miller (1988) extended this research to show that the high-probability locations could be both an absolute screen location and a relative location within an array (see also Hoffmann & Kunde, 1999). Miller’s latter finding anticipated subsequent work on contextual cueing, in which participants learn the target locations in search arrays they encountered before (Chun & Jiang, 1998). Interest in location probability learning was further stimulated in the 2000’s, when Geng and Behrmann (2002) extended the finding to patients with hemifield neglect. Despite a deficit in attending to the left visual field, these patients acquired probability learning when the visual search target more often appeared in the left than the right visual field. Learning reduced, but did not eliminate, left hemifield neglect. Geng and Behrmann (2002) termed this type of attention “probability cuing” and suggested that it could be distinguished from other forms of attention. Location probability learning has since been extended to the learning of distractor locations. Goschy, Bakos, Müller, and Zehetleitner (2014) show that locations frequently containing a salient distractor are more easily ignored (see also Leber, Gwinn, Hong, & O’Toole, 2016 for implicit learning of distractor locations). However, as discussed below, the implication of location probability learning with respect to attentional control was obscured by
several concerns. These concerns are only recently addressed, linking this simple but powerful finding to a fundamental way in which the control of spatial attention is implemented.

2.2 Role of explicit knowledge

The first concern is that the location probability effect is the same as attention driven by current goals. If people know which locations are likely to contain the target, they can deliberately prioritize those locations, rendering the effect indistinguishable from endogenous cuing in Posner’s classic work (Posner, 1980). However, when queried after the experiment, few healthy control participants in Geng and Behrmann (2002) reported that they noticed the target’s spatial distribution. Formal tests of explicit recognition yielded similar results. Here I review one representative study to illustrate location probability learning and its implicit nature.

Jiang, Swallow, Rosenbaum, & Herzig (2013) asked participants to search for a target letter (“T”) among distractor letters (“L”). Participants reported the orientation (or color) of the target upon detection. There was only one target on the display during each trial. Unbeknownst to participants, across multiple trials, the target was more often placed in some locations than in others. To reduce the likelihood that participants could explicitly remember where the target-rich locations were, the study used a large number of potential locations – 100 in total – and associated high target probability with an entire region rather than a few locations (see also Druker & Anderson, 2010). The target appeared in one visual quadrant on 50% of the trials, and in any of the other three quadrants on 16.7% of the trials. The high-probability quadrant was counterbalanced across participants. Figure 1 illustrates the location probability manipulation and the visual search results. Participants became significantly faster finding the target in the high-probability quadrant.

![Figure 1. Location probability learning. Left: The training-testing two-phase location probability design. In the training phase the target more often appears in one visual quadrant (the “rich”, high-probability quadrant). In the testing phase, the target’s location is random. Right: Reaction time (RT) becomes faster when the target appears in the high-probability quadrant. This effect persists during the testing phase. Each phase has about 500 trials in the experiment depicted here (adapted from Jiang, Swallow, Rosenbaum, & Herzig, 2013).](image)

To test whether the location probability effect resulted from deliberate goals, at the end of the experiment participants were asked whether they thought the target’s location was random. Similar to Geng and Behrmann (2002) and Druker and Anderson (2010)’s findings, few participants spontaneously realized that the target was biased toward some locations. Participants were then asked to identify the high probability quadrant. The number of participants correctly identifying the high probability quadrant was not greater than chance (Jiang, Swallow, & Rosenbaum, 2013). Furthermore, participants who correctly identified the high probability quadrant did not show greater location probability learning than those who failed the recognition test (Jiang, Koutstaal, & Twedell, 2016). Thus, location probability learning results largely from implicit learning rather than explicit goals.

2.3 Statistical learning or short-term repetition priming?

A second concern about the location probability effect relates to its underlying mechanism. Even if not goal-driven, this effect may not be products of learning the target’s location probability over a long time. It is well known that when a target repeats its location across consecutive trials, people find the target faster due to inter-trial repetition priming (Maljkovic & Nakayama, 1996). If a location more often contains the target, location repetition also tends to happen more often in that
location than elsewhere. On this account, short-term inter-trial priming, rather than learning of the target’s location probability over a sustained period of time, underlies the location probability effect.

Evidence supporting the inter-trial priming account initially came from Walthew and Gilchrist (2006). They presented items in a ring of 8 locations and asked participants to direct the first saccadic eye movement to the target’s location. The target appeared more often in one visual field than the other. Walthew and Gilchrist constrained the trial sequence such that the target could not repeat its exact location on consecutive trials. Under this condition, participants were no more successful landing their first saccade on targets in the high-probability visual field than the other visual field. Similar results were found in a study in which participants identified a target presented among three distractors. The target appeared in one of the four locations 40% of the time and in each of the other locations 20% of the time. Probability learning was unreliable when the target’s location was constrained such that it could not repeat on consecutive trials (Kabata & Matsumoto, 2012).

However, constraining the target’s location did not eliminate the location probability effect in other studies. For example, in a failure to replicate Walthew and Gilchrist, Jones and Kaschak (2002) found significant probability learning even when the target’s location never repeated on consecutive trials. Similarly, Goschy et al. (2014) found that probability learning of a salient distractor’s location was robust even when its location could not repeat on consecutive trials. In considering these inconsistencies, Druker and Anderson (2010) noted that the methodology used in these studies may have introduced new statistical learning that masked probability cuing. Specifically, restricting the target’s location repetition introduced non-random statistics that may bias attention away from where the target was found before. Methodologies that do not depend on restricting location repetition are needed to resolve these inconsistencies.

Recently, we developed a new approach to isolate long-term probability learning from short-term inter-trial priming (Jiang, Swallow, Rosenbaum, et al., 2013). This paradigm relies on the logic that long-term learning is likely to persist even after the target’s location has become random. When the target location is random, the odds of a target repeating its location are the same in all regions. If the location probability effect reflects inter-trial priming, an effect that lasts no longer than 5-8 trials (Maljkovic & Nakayama, 1996), then the reaction time advantage in the high-probability quadrant should disappear shortly after the target’s location becomes random. As shown in Figure 1, this is not the case. The location probability effect is surprisingly persistent. Several hundred trials of random target locations fail to extinguish the RT advantage in the previously high probability quadrant. A later study estimated the size of location repetition priming by contrasting location-switch and location-repeat trials. This study showed that inter-trial repetition priming significantly underestimated the RT advantage in the high-probability quadrant (Jiang, Sha, & Remington, 2015). Thus, the location probability effect includes not just short-term inter-trial priming but also long-term statistical learning.

2.4 Attentional guidance or response selection?

A third concern about the location probability effect is whether or not it constitutes a form of attention. It is possible that search time is the same regardless of where the target is. The RT advantage may arise because people are more confident when the target is found in the high-probability quadrant. Their faster RT results from late facilitation of response decision rather than early facilitation of visual search. This concern was raised in the context of another experimental paradigm – contextual cuing (Kunar, Flusberg, Horowitz, & Wolfe, 2007). Does location probability affect how we attend to visual space, or does learning influence processes much later, perhaps after the target is found?

Two methods can be used to distinguish an early attention effect from a late response decision effect. The first is to vary the number of distractors and examine the efficiency of search. In difficult search tasks like the T-among-L task, adding distractors on the display slows people down. The slope of the regression line relating RT to the number of items corresponds to the speed of shifting attention from one item to another. The intercept of the line corresponds to a constant such as response
decision and motor execution (Wolfe, 1998). If location probability affects late response decision, then it should reduce the intercept, producing a constant reduction in RT across all set sizes. Empirical data, however, do not support the response decision account. Search slope becomes shallower when the target appears in the high probability quadrant, suggesting that learning facilitates the speed of attentional allocation rather than response decisions (Jiang, Swallow, & Rosenbaum, 2013).

A second method to dissociate early attentional processes from late response decision processes is by examining eye movements. The eyes usually go to where the current focus of attention is. If location probability learning affects the allocation of attention, then the first eye movement, initiated within 200ms of trial onset, should be more often directed towards the high probability quadrant than expected by chance. This is indeed the case (Jiang, Won, & Swallow, 2014; Jones & Kaschak, 2012). These data show that an attentional bias toward the high probability locations emerges very early in the search process.

Does location probability effect occur when people are not allowed to move their eyes? Several findings show that it does. In some experiments participants were required to keep their eyes fixated at the center of the screen. An eye tracker verified that they did not move their eyes in visual search (Jiang et al., 2014; Geng & Behrmann, 2005). In other experiments, the display was presented briefly, leaving no time for participants to move their eyes (Jiang & Swallow, 2013b, 2013a). In these cases, participants did not make more frequent eye movements to where the high probability locations were. Nonetheless, they demonstrated faster RT and higher accuracy in finding the target there. Not only is frequently moving one’s eyes to a location unnecessary, but also is it insufficient, to induce probability cuing. In one study, a central arrow endogenously cued participants to attend to one region of space (Jiang, Swallow, & Rosenbaum, 2013). The arrow’s direction changed from trial to trial and whichever quadrant it pointed to had a 50% probability of containing the target. In these experiments, the arrow was frequently directed towards one specific quadrant. Even though participants frequently moved their eyes toward that quadrant, they failed to acquire location probability learning. Thus, the use of an arrow in a goal-directed manner interfered with the development of a search habit. This finding is reminiscent of other habitual behaviors - instructing someone to carry out those behaviors hinders habit formation (Wood & Rünger, 2016).

Together, these studies provide strong evidence that location probability learning is a form of attention not under deliberate control. Learning affects both covert attentional shift and overt eye movements. Following Geng and Behrmann (2002), we will refer to location probability learning as “probability cuing.”

3. Habitual versus goal-driven attention

In the integrated framework, all sources of attention, including selection history and current goal, channel into the attentional priority map to influence spatial selection (Awh et al., 2012). This framework implies that despite differences in how they originate, probability cuing and current goals should affect spatial attention in similar ways. Here I present evidence that challenges this assumption. These findings suggest, instead, that probability cuing and goal-driven attention are implemented differently in the control of spatial attention.

3.1. Persistence/flexibility

One hallmark of goal-driven attention is its flexibility. Where one attends depends on what is important at the moment, and this changes as current goals change. Is probability cuing similarly flexible? That is, can people rapidly readjust where to attend when the visual statistics about the target’s location have changed? Section 2 already hints at the answer: probability cuing is persistent and does not rapidly readjust.

In one experiment, participants performed the training phase on Day 1, where they searched for a target that frequently appeared in one visual quadrant. One week later, they returned for the testing phase, during which the target’s location was completely random. The previous attentional bias toward the high-probability quadrant was no longer valid. Nonetheless, participants continued
to exhibit a search advantage in the (previously) high probability quadrant (Jiang, Swallow, Rosenbaum, et al., 2013). Another experiment used a reversal training procedure to probe the extinction of a previously learned attentional bias. In the training phase participants acquired an attentional bias toward one quadrant. In the testing phase the high-probability quadrant changed to a new region. Over several hundred trials of testing, participants acquired a new attentional bias toward the currently high-probability quadrant. The previously acquired bias toward the other quadrant slowly weakened, though it remained significant at the end of the testing phase (Jiang, Swallow, & Rosenbaum, 2013).

Probability cuing not only persists over time but also across tasks. Contextual changes fail to extinguish the learned attentional bias (Jiang, Swallow, Won, Cistera, & Rosenbaum, 2015). These experiments use the “training-testing” two-phase design, where the training phase induces an attentional bias toward a high-probability quadrant, and the testing phase probes the persistence of this bias when the target’s location is random. Probability cuing persists when different stimuli are used in the two phases. For example, one phase involves finding a T among Ls and the other phase involves finding a 2 among 5’s. Or one phase uses black items and the other phase uses white items. Furthermore, even when task difficulty varies across the two phases, probability cuing transfers between an easy task (such as finding a T among dissimilar Ls) and a difficult task (such as finding a T among similar Ls).

The persistence of probability cuing suggests that it is habit-like. Like habits, the transfer of probability cuing is limited to similar tasks. Whether it is searching for a 2 among 5’s or a T among Ls, the task requires participants to attend to each item serially and evaluate whether it is a target. Probability cuing acquired in a T-among-L search task does not transfer to tasks that do not involve serial search. For example, one study examined cross-task transfer between the T-among-L search task and a treasure-hunt task (Jiang, Swallow, et al., 2015). In the treasure-hunt task, participants view several Ls and click on one of them to discover hidden treasure. The treasure hunt task requires no search – just high-level decision about which item to click on. Participants learn to favor locations that are more likely to contain the hidden treasure. However, the location preference acquired from the treasure hunt task does not transfer to visual search. Similarly, probability cuing acquired from visual search does not transfer to treasure hunt. These findings show that probability cuing is moderately task-specific.

### 3.2. Spatial reference frame

A second difference between probability cuing and explicit, goal-driven attention is the reference frame in which attended locations are coded. When we attend to a region of space, is the attended location coded relative to ourselves as in our upper left visual field, or is it coded relative to the environment, as in a particular region of space independent of our viewpoint? Because most forms of attention do not linger long enough to permit a complete change in viewpoint, previous research has addressed a more limited version of this question, namely, whether spatial attention is retinotopic or spatiotopic (Cavanagh, Hunt, Afraz, & Rolfs, 2010; Wurtz, 2008). For example, participants may be asked to fixate at the center of the screen while a spatial cue directs their attention to the left side of the screen. Participants then make a saccadic eye movement to another location. With this setup, investigators can test whether effects of attentional cuing linger in the same screen location (spatiotopic) or move with the eyes (retinotopic). However, “spatiotopic” in this design is not the same as environment-centered coding. Because only the eye fixation changes, a spatiotopic effect can be centered on the environment or on parts of the viewer that remain stable (e.g., head, body). Because it survives several hundred trials of extinction testing, probability cuing provides a unique opportunity to dissociate environment-centered coding of spatial attention from various viewer-centered frames of reference.
An initial study on this question introduced a viewpoint change between the training and testing phases (Jiang & Swallow, 2013b). This study presented the search array on a computer monitor laid flat on a desk (Figure 2). Participants sat at one side of the desk and identified the target’s color in the T-among-L search task. Unbeknownst to them, the T was most often placed in one quadrant of the monitor. Following the training phase, participants moved their seating position to an adjacent side of the desk, rotating their viewpoint by 90°. They then performed the same search task from the new perspective, except that this time the target’s location was entirely random. Based on studies described in Section 3.1, we know that the previously learned attentional bias should persist in the testing phase. But to where would attention be directed? If the high-probability locations are coded in an environment-centered reference frame, then people should continue to respond faster when the target appears in that part of the screen. But if the high-probability locations are coded relative to the participants, such as in their lower right, then after a change in viewpoint, the attentional bias should be redirected to a new part of the monitor corresponding to the same visual field as before. Finally, the attended locations may be coded relative to both the environment and the viewer, in which case people should be faster finding the target in either the viewer-centered or environment-centered high-probability quadrant.

Results show that participants are significantly faster when the target appears in the viewer-centered high-probability quadrant relative to the low-probability quadrants (Figure 2). In addition, when the target appears in the environment-centered high-probability quadrant, there is no advantage in search RT. This finding suggests that the attended locations are coded only relative to the viewer and not to the environment.

The viewer-centered coding of probability cuing is observed in many situations. For example, even with the addition of a salient landmark, such as a red border on one edge of the monitor, probability cuing is viewer-centered rather than environment-centered. In addition, when the T-among-L search task is performed against a constant natural scene, participants still do not code the high-probability quadrant relative to the scene. These data also hold when the search display is presented so briefly that people have no time to move their eyes (Jiang & Swallow, 2013b). Finally, instructing participants to attend to the environment-centered quadrant after reseating does not eliminate the automatic bias toward the viewer-centered quadrant (Jiang, Swallow, & Sun, 2014).

We found three scenarios under which attended locations are coded, in part, in an environment-centered reference frame. First, when participants acquire explicit knowledge about where the target is most often placed, they tend to favor the environment-centered region after reseating. For example, a subset of the college students tested by Jiang et al. (2016) can correctly identify the high-probability quadrant. The other participants have no explicit awareness. In the testing phase, the unaware participants show only a viewer-centered attentional bias. However, the
aware participants show faster search RT in both the viewer-centered and the environment-centered high-probability quadrants. This finding suggests that goal-driven attention can be flexibly directed toward a stable region of space after a viewpoint change.

A second scenario in which probability cuing is partly environment-centered is when viewpoint changes are introduced by tilting the participants' body, such as when the person goes from sitting upright to lying down. In one study, we asked participants to perform the T-among-L search task while lying on a slant, tilting their body and head 45° away from vertical (e.g., clockwise) as they viewed items on an upright monitor. Following location probability training, we tested the persistence of their learned attentional bias by flipping the slant. Now participants tilted their body and head in the opposite direction (e.g., counter-clockwise), 45° away from vertical. We found that probability cuing persisted in both the same screen locations as the previously high-probability region, and the same visual field as before. Similar results were found when participants sat upright, but tilted their head 45° against a headrest during training, and in the opposite direction 45° during testing (Jiang & Swallow, 2013a). What accounts for the presence of an environment-centered component with a body or head tilt, but not in studies in which people move their seating position around a desk? A plausible account is that in the body-tilt setup, the gravitational axis provides a strong and stable environmental cue. The gravitational axis remains stable regardless of how the participants' body or head is tilted. Evidence that the gravitational axis is used can be found in object recognition. For example, the outline shape of the continent of Africa remains recognizable when one tilts his or her head, but is difficult to recognize when the shape itself is rotated, disrupting its alignment with the gravitational axis (Rock, 1997). Similar results are found when testing hemifield neglect patients. Locations neglected by these patients fall both on the left of the patient's body and on the left of the gravitational axis (Farah, Brunn, Wong, Wallace, & Carpenter, 1990). These findings suggest that although probability cuing is strongly viewer-centered, it has some tolerance to changes in body or head tilt.

A third condition in which environment-centered coding is observed is when people conduct visual search in a large outdoor space. On one hand, probability cuing is partially viewer-centered in the outdoor task. Consistently referencing the high-probability region relative to the participant can induce learning, even though that region is random in the external environment (Jiang, Won, Swallow, & Mussack, 2014). On the other hand, unlike in computerized tasks, changes in viewpoint from trial-to-trial in the outdoor task does not completely disrupt location probability learning. Two reasons may contribute to an environment-centered coding in outdoor search: spatial scale (Jiang & Won, 2015) and awareness. Spatial scale is known to influence tasks such as reorientation and navigation (Wolbers & Wiener, 2014). Outdoor space also affords richer environmental cues that can facilitate environment-centered coding. Alternatively, the difference between computerized and outdoor search tasks may arise from different degrees of explicit awareness. Nearly all participants tested in large-scale tasks can identify the high-probability region (Jiang, Won, Swallow, et al., 2014; Smith, Hood, & Gilchrist, 2010). As noted earlier, with explicit awareness participants can mobilize goal-driven attention to flexibly attend to the external environment even as their viewpoint changes. Whether spatial scale has an effect over and above increased awareness remains to be tested.

Thus, whereas explicit, goal-driven attention codes space in a flexible manner, including the use of an environment-centered reference frame, implicit probability cuing is largely viewer-centered.

3.3. Constraints on probability cuing

Coding attended locations relative to the viewer constrains what people can possibly learn. In one study, we asked participants to perform visual search on a monitor laid on a stand. Participants can stand at various locations around the monitor, allowing them to change viewpoint from trial to trial (Jiang, Swallow, & Capistrano, 2013). Figure 3 illustrates three trials viewed from above. The green footprint indicates where the participants stand on that trial, and the 50% region indicates where the target is most likely presented. In all trials the target is most often placed in a fixed part of the monitor, so there is consistent information about where the target is likely to be in the external
environment. But because participants stand at random locations around the monitor, the high-probability locations are random relative to their viewpoint.

Under this condition, and in the absence of explicit instructions about where the target may be, participants show no learning. They are not faster at finding the target in the high-probability locations than in the other locations. In contrast, when the design is modified by rendering the high-probability quadrant consistent relative to the participants, participants acquire probability cuing (Jiang & Swallow, 2014). Learning occurs even though the high-probability quadrant is random in the external environment. As long as it can be consistently referenced to the viewer, a strong search habit emerges.

Under which conditions can participants acquire probability cuing toward an environmentally stable, but viewer-independent, high-probability quadrant? Section 3.2 discussed two conditions: when participants have explicit awareness about where the target is most likely presented, and when search is conducted in a large outdoor space. When these conditions are not met, probability cuing invariably fails. Failure is seen even when participants change their viewpoint in steps of just 30º from one trial to the next, in a predictable direction (Jiang & Swallow, 2014). Such a small incremental change should have facilitated spatial updating, but it is not incorporated into probability learning. Failure also occurs when the search display itself contains rich visual cues, as when people search for a traffic icon on a realistic Google Map. Finally, compelling people to move during search (e.g., by blocking some of the search items from view with a tall occluder), a condition that mimics what they do in the outdoor task, does not lead to environment-centered probability learning (Jiang & Won, 2015). These findings highlight a critical feature of probability cuing – its function is not to extract environmental statistics. Statistical regularities are only learned when they are represented in a format decodable by the learning system. In location probability learning, this system codes information in a viewer-centered reference frame. In the absence of explicit awareness, only viewer-centered statistical regularities support learning.

3.4. Reliance on working memory resources

The next major difference between probability cuing and goal-driven attention is their reliance on working memory resources. Previous research on attention and working memory considers these two constructs as closely related. Holding a location in spatial working memory biases spatial attention to that location (Awh & Jonides, 2001) and holding a color or shape in working memory influences subsequent search for other colors or shapes (Olivers, Peters, Houtkamp, & Roelfsema, 2011). In fact, working memory is sometimes considered as attention directed toward an internal representation (Chun, 2011).
Consistent with this idea, imposing a spatial working memory load impairs goal-driven spatial attention (Won & Jiang, 2015). In one study, participants perform the T-among-L search task. Before the onset of the search display, a central arrow is presented as an endogenous cue. Participants are informed that the target is more likely located in the quadrant cued by the arrow. In the absence of a secondary working memory load, participants are effective in using the central arrow to guide attention. RT is faster and the slope of the RT-set size function is shallower when the target appears in the cued quadrant than in the other quadrants. Goal-driven, endogenous cuing becomes less effective when participants conduct visual search under a secondary working memory load. In this condition, participants encode an array of dot locations in working memory. During the retention interval, they perform the arrow-cued T-among-L search task. The cue-validity effect in RT and search slope decline in the secondary load condition. Thus, goal-driven attention depends on working memory resources.

Different results are observed when spatial attention is cued through location probability learning. If, instead of instructing participants to attend to a specific quadrant first, we simply let participants acquire location probability learning by placing the target more often in one quadrant. RT is faster and the slope of the RT-set size function is shallower in the high-probability relative to the low-probability quadrants. When the task is conducted under a working memory load, such that participants perform the search task during the dot-array retention interval, the RT advantage in the high-probability quadrant is maintained. Even though probability cuing introduces a spatial attentional bias, it does not depend on the availability of spatial working memory resources (Won & Jiang, 2015).

### 3.5. Effects of aging

Aging is accompanied by significant changes in brain and cognitive functions. A decline in the frontoparietal network is known to impair cognitive control, in tasks that involve executive function (e.g., switching, shifting, or manipulating information in working memory; Braver & Barch, 2002; Salthouse, 2012). Spatial attention is also affected – visual clutter is more detrimental to older adults’ search behavior than to young adults’ (Grahame, Laberge, & Scialfa, 2004). Other studies show that aging affects the basal ganglia (Howard & Howard, 2013), which is important for acquiring and maintaining habitual behavior. In fact, older adults are impaired when performing some probabilistic learning tasks, such as learning a sequence of locations and responses (Howard Jr., Howard, Dennis, Yankovich, & Vaidya, 2004). However, most studies show that aging has a more deleterious effect on explicit learning than implicit learning.

Consistent with this idea, in a recent study we showed that aging affects goal-driven attention to a greater extent than probability cuing (Twedell et al., 2016). This study also probes the spatial reference frame older adults use to code frequently attended locations. Using the “training-testing” two-phase procedure, we first train older (60-80 years) and young adults (18-30 years) to perform the T-among-L search task on a monitor laid flat on the desk. Similar to young adults, older adults demonstrate successful probability cuing. Their RT is faster, and the slope of the RT-set size function is shallower, when the target appears in the high-probability, rather than the low-probability, quadrants. Following training, participants are informed of where the high-probability quadrant was on the screen. They then move their sitting position to an adjacent side of the desk, changing their viewpoint by 90°. The screen location where the target is most often found previously is marked with a blue outline square. Participants are asked to prioritize that quadrant for the remainder of the experiment. During this phase the target’s location is random. Despite this instruction, both young and older adults exhibit persisting attentional bias toward the same visual field as where the previously high-probability quadrant was (viewer-centered). The magnitude of this bias is comparable between the two age groups, suggesting that habitual attention is relatively intact in older adults. At the same time, both groups are faster at finding the target in the instructed, environment-centered high-probability quadrant than in the low-probability quadrants. However, the magnitude of this effect differed: the RT advantage is about twice as large in young adults as in older adults. Thus, although older adults are able to rely on instructions to prioritize a region of space, their ability
to do so is impaired relative to the young adults. This suggests that as we age, the balance between
goal-driven attention and habitual attention is tipped toward the latter.

Table 1 summarizes differences between probability cuing and goal-driven attention. These
differences suggest that the control of spatial attention may be implemented in different ways. Before
turning to a theoretical account, we will consider mechanisms underlying a related statistical learning
paradigm – contextual cuing.

Table 1. Characteristics of goal-driven attention and probability cuing

<table>
<thead>
<tr>
<th></th>
<th>Goal-driven attention</th>
<th>Probability cuing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awareness</td>
<td>Explicit</td>
<td>Implicit</td>
</tr>
<tr>
<td>Flexibility</td>
<td>Can rapidly readjust</td>
<td>Highly persistent</td>
</tr>
<tr>
<td>Spatial reference frame</td>
<td>Flexible, can be environment-centered</td>
<td>Largely viewer-centered</td>
</tr>
<tr>
<td>Working memory demand</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Aging</td>
<td>Impaired in older adults</td>
<td>Intact in older adults</td>
</tr>
</tbody>
</table>

4. Contextual cuing

A widely used paradigm on implicit learning and attention is contextual cuing (Chun & Jiang, 1998). This paradigm differs from probability cuing in several ways. In standard contextual cuing
tasks, participants perform a T-among-L search task across 20-30 experimental blocks. Unbeknownst to participants, some of the search displays are repeatedly shown, typically once per block. These repeated, “old” displays are randomly intermixed with “new” displays. The new displays are not entirely new – the location of the T target among these displays also repeats across blocks. Thus, repetition of the target’s location is controlled between the old and new displays. In addition, the old and new displays use various target locations; when all trials are considered, no region of space is more likely to contain the target. Because the target’s location is evenly distributed across all quadrants, and because the specific target locations repeat equally often for old and new displays, contextual cuing does not support location probability learning. Instead, participants have to rely on specific spatial layouts to cue the target’s location. Many studies using this paradigm show that participants find the target among old displays more quickly than among new displays (for a review, see Goujon, Didierjean, & Thorpe, 2015). Though occasional slippage of awareness occurs in contextual cuing (Vadillo, Konstantinidis, & Shanks, 2016), learning is largely implicit. Instructing participants about the display repetition does not increase the RT advantage. If anything, conscious attempts to memorize repeated displays interfere with contextual cuing (Chun & Jiang, 2003).

Even though participants learn the target-context association rather than a general spatial bias in contextual cuing, their learning shows similar properties to probability cuing. In addition to being implicit, contextual cuing is also persistent. The search advantage for old displays over new displays is detectable one-week after training (Chun & Jiang, 2003). Contextual cuing shows a limited degree of transfer across stimuli and tasks. Changes in surface properties, such as from a T-among-L task to a 2-among-5 task, do not disrupt contextual cuing (Chun & Jiang, 1998). However, contextual cuing can become specific to item color or task difficulty if different spatial contexts are associated with different colors or task difficulty during training (Jiang & Song, 2005a). In addition, contextual cuing shows limited transfer between visual search and change detection (Jiang & Song, 2005b). Other properties characteristic of probability cuing are sometimes, though not always, found in contextual cuing. Viewpoint-specificity of the spatial context is demonstrated in some studies (Chua & Chun, 2003), though small changes in viewpoint are tolerated (Tsuchiai, Matsumiya, Kuriki, & Shioiri, 2012). Imposing a working memory load leaves contextual cuing intact in some studies (Vickery, Sussman, & Jiang, 2010), but impaired in others (Travis, Mattingley, & Dux, 2013). Contextual cuing appears to depend on attending to the repeated context for its expression (Jiang & Chun, 2001), but not for its acquisition (Jiang & Leung, 2005). Healthy older adults show intact contextual cuing in some studies.
(Howard Jr. et al., 2004; Lyon, Scialfa, Cordazzo, & Bubric, 2014), but a reduced effect in others (Smyth & Shanks, 2011).

Not all inconsistencies across studies on contextual cuing have been resolved. This state of affair may originate from the multifaceted nature of contextual cuing. An RT advantage depends on two mechanisms: learning the repeated context, and allocating attention to the associated target location. The first, context learning, likely depends on the hippocampus and other structures of the medial temporal lobe. The latter, attentional cuing, may depend on the basal ganglia and other habit-related brain regions. Consistent with this possibility, research on patients with brain damage shows that both amnesic patients with medial-temporal damage (Chun & Phelps, 1999; Manns & Squire, 2001) and Parkinson’s patients with basal ganglia damage (van Asselen et al., 2009) are impaired in contextual cuing. When older adults are found to show reduced contextual cuing, the deficit appears to be related to a reduction in hippocampal function rather than to aging itself (Negash et al., 2015). Other mechanisms may also contribute to contextual cuing, notably facilitation of later processes, such as response selection and decision (Kunar et al., 2007).

The complex nature makes it tenuous to relate contextual cuing to a search habit. Nonetheless, some findings suggest that learned search habits may be a component of contextual cuing. First, though cuing is contextual, distractors adjacent to the target are more important than distractors far away from the target (Olson & Chun, 2002). In fact, a full contextual cuing effect is found when distant items are placed in random locations, keeping only the locations of items in the target’s quadrant consistent (Brady & Chun, 2007). The near items are likely the ones people visited shortly before encountering the target; any kind of attentional shift among these items is most likely reinforced when people successfully find the target. In this regard, contextual cuing can be related to probability cuing. Participants do not just learn an abstract map of where the target is in relation to the context; they learn how to shift attention among a local context to find the target.

Consistent with the search habit idea, one study shows that contextual cuing is observed only when people actively search a display. In this study, each search display is preceded by place-holder items in locations that will be occupied by the search elements. The place-holders provide a preview of the repeated spatial context. If people are able to use this context to cue the target’s location, they should benefit from the preview of the place-holders. The longer the place-holders are previewed, the more time people have to shift attention to the associated target’s location. This should have led to a greater contextual cuing effect. Results reveal no preview benefit (Jiang, Sigstad, & Swallow, 2013). This finding suggests that attentional cuing starts only when active visual search is underway, a finding that differs from explicit scene-target learning (Brockmole & Henderson, 2006). Previewing the scene before the presentation of the target enhances RT in scene-based contextual cuing (Summerfield, Lepsien, Gitelman, Mesulam, & Nobre, 2006). In this case, people acquire explicit association between the scene and the target location. Previewing the scene allows them to shift attention to the target’s location in anticipation of its presentation. This is not the case with implicit, array-based contextual cuing. Cuing in the implicit task is not anticipatory and does not initiate until search is underway.

In sum, although what is learned in contextual cuing differs from location probability learning, the two paradigms may both induce search habits. In contextual cuing, the habit is specific to individual spatial contexts. In location probability learning, the habit applies to all spatial contexts. Contextual cuing, however, includes other processes, including context learning and enhanced response decision.

5. The multi-level framework of attention

5.1. Sources and implementation of attention

The empirical data reviewed so far suggest that the control of spatial attention may be conceptualized at two levels. One level specifies sources of attentional control, the other specifies how these sources are implemented.
Sources of attentional control include current goals, perceptual salience, emotion and motivation, as well as many factors related to previous experience, such as inter-trial priming, reward, and learned semantic association. Important distinctions can be made among the various sources. For example, goal-driven attention and salience-driven attention exhibit different time courses both in behavioral measures and in neurophysiological recordings (Buschman & Miller, 2007; Jonides, 1980). Emotional, motivational, and reward-based attention rely on additional (and often unique) parts of the brain, such as the amygdala and the striatum (Anderson et al., 2016; Vuilleumier, 2015). At the same time, the various sources of attention are interactive. For example, top-down goals modulate the potency of attentional capture by a perceptually salient stimulus (Folk, Remington, & Johnston, 1992; Ipata, Gee, Gottlieb, Bisley, & Goldberg, 2006), and explicit, goal-driven attention may overshadow implicit contextual or probability learning (Jiang, Swallow, & Rosenbaum, 2013; Rosenbaum & Jiang, 2013). Given the complexity, it is not surprising that most research on spatial attention is directed at understanding the various sources of attentional control.

The research reviewed here suggests, however, that not all sources affect spatial attention in the same manner. At the implementation level, attentional control may be considered a dual-system, modulating where priority should be given and how one shifts spatial attention. Figure 4 illustrates the two ways in which the control of spatial attention can be implemented (Jiang, Swallow, & Capistrano, 2013). Where people attend may be sets up as a priority map, illustrated in the figure as the heat map. The hot spots are the locations assigned with greater priority, according to current and long-term goals (including emotional and motivational incentives). Its setup may occur prior to and during task performance, and may be modeled as a baseline shift in neurons coding different locations. This map affects subsequent perceptual selection and action planning. In addition, attention also has a procedural component – the how component of spatial attention. Visual search, for example, involves frequent shifts of spatial attention from one location to another. Every time a target is successfully detected, the most recent shifts of attention that lead to target detection are reinforced, increasing the likelihood that such vectors of attentional shift would occur again.

The dual-system view of how spatial attention is implemented fits with the division of vision into perception and visuomotor functions (Goodale & Westwood, 2004). Whereas the perceptual system is typically object-centered, the visuomotor system is egocentric (Goodale & Haffenden, 1998). This property is consistent with what we found on spatial attention. The dual-system view echoes Allport’s (1989) proposal that attention is for action as well as for perception. It is consistent with the
idea that information processing is not abstracted from the viewer, but is sometimes embodied and constrained by the viewer’s sensorimotor processing (Wilson, 2002). By acknowledging both the where and how components of spatial attention, the dual-system view is also broadly consistent with neurophysiological findings on spatial attention. Although spatiotopic coding is possible (Melcher & Morrone, 2003), many neurons in the parietal cortex code space in an eye-centered, head-centered, or body-centered manner (Cohen & Andersen, 2002). One function of such egocentric codes may be to implement attentional selection in action space.

5.2 Potential neural substrate

Geng and Behrmann (2002)’s study on neglect patients reviewed earlier suggests that probability cuing relies on distinct brain mechanisms than goal-driven attention. In two subsequent studies, Shaqiri and Anderson (2012, 2013) also found evidence of probability cuing in patients with right-hemisphere damage, including those with neglect symptoms. In another study, Lucas et al. (2013) asked neglect patients to search for multiple targets among distractors and touch the one under which a high reward is hidden. By placing the high-reward target more often in the left visual field, Lucas et al. were able to induce a left-ward spatial bias in neglect patients and healthy controls.

The brain damage in neglect patients tested in the studies reviewed above is not restricted to a single brain site. Most patients have damage in the right parietal cortex, frontal cortex, and the basal ganglia. Their preserved ability to acquire a learned attentional bias suggests that these brain regions are not essential for probability cuing. Which brain regions are important remain to be tested. On theoretical grounds, if the learned attentional bias reflects a change in how one shifts spatial attention, then the relevant brain regions are likely those involved in both covert and overt orienting of attention, such as the frontal eye fields and the superior colliculus. In fact, an early study recorded neuronal activity in superior colliculus under two conditions: when the target’s location changes frequently or when it is fixed. Neurons in the superior colliculus show greater activity in the fixed condition, even before an eye movement is made (Basso & Wurtz, 1998).

5.3 Sources vs. implementation of attentional control

How do the various sources of spatial attention relate to the two ways in which control is implemented? Although current goals map primarily onto where and implicit location probability learning primarily onto how, exceptions likely exist. In our view, all known sources of attention are likely to modulate both the where and how of spatial attention.

Consider current goals. When instructed to attend to one region of space, people are likely to set up higher priority weights for that region; this can be done offline, before the actual movement of attention. In this sense, current goals affect where in space is prioritized. However, research on habit learning suggests that what initially starts out as goal-oriented behavior often becomes habitual with extensive practice (Wood & Rünger, 2016). If this is the case, goal-driven attention may become habit-like over time. In addition, once the spatial priority is set, the priority map can also influence how people move attention. Thus, current goals primarily affect where one prioritizes processing, but can also modulate how attention is allocated.

Conversely, though implicit location probability learning primarily affects how people shift spatial attention, the emergence of explicit awareness may additionally influence the where component. In fact, studies reviewed above suggest that once people become aware of the target’s location probability, they use this information to set up higher priority weights in the high-probability locations. This is characteristic of controlling the where component of attention.

Other sources of spatial attention, such as inter-trial repetition priming and monetary reward, may also affect both the where and the how components of spatial attention. For example, the tendency to orient toward the same location as the preceding target’s location may reflect both a bias toward where the previous target was, and reinforcement learning of how to orient in space. To identify the connection between a specific source of attention and how its control is implemented,
future studies may examine characteristics that distinguish the *where* and *how* components, including flexibility, spatial reference frame, and reliance on working memory resources.

6. **Conclusion and Future Directions**

   The above review suggests that the control of spatial attention can be understood at two levels: the source of attention and its implementation. As research on the various sources of attention deepens, it is also important to consider whether each source affects attentional priority given to a spatial location (*where* to attend) or the movement of attention (*how* to find targets). After all, the ultimate goal of spatial selection is to act on the selected information. Including both the perceptual and the action components of attention is necessary for future models of attention.

   This review provides some initial evidence linking goal-driven and habitual attention to the dual-system view of attentional implementation. Going forward, it will be important to identify how other sources of attention are implemented, such as reward-based or emotional/motivational influences of attention. Such research may help resolve discrepancies in previous findings on how reward, particularly implicitly learned reward, affects spatial attention (Chelazzi et al., 2014; Hickey, Chelazzi, & Theeuwes, 2014; Jiang, Sha, et al., 2015; Won & Leber, 2016). Future studies should also investigate the underlying neural mechanism for habitual attention. Research on individual differences, healthy and pathological aging, and brain-damaged patients may shed light on how these different systems are mapped onto the brain.

**Acknowledgments**

   Research reviewed in this article was funded in part by NIH R03 MH102586 and an Engdahl family research fund. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health. I thank Leo Chelazzi, Roger Remington, Sha Li, Doug Addleman, Abby Barthel, Emily Twedell, and Nikita Salovich for suggestions and discussions. Correspondence should be directed to Yuhong Jiang, 75 East River Road, S504 Elliott Hall, Minneapolis, MN 55455. Email: jiang166@umn.edu.

**References**


