Changing viewer perspectives reveals constraints to implicit visual statistical learning

Yuhong V. Jiang

Department of Psychology, University of Minnesota, Minneapolis, MN, USA

Khena M. Swallow

Department of Psychology, Cornell University, Ithaca, NY, USA

Statistical learning—learning environmental regularities to guide behavior—likely plays an important role in natural human behavior. One potential use is in search for valuable items. Because visual statistical learning can be acquired quickly and without intention or awareness, it could optimize search and thereby conserve energy. For this to be true, however, visual statistical learning needs to be viewpoint invariant, facilitating search even when people walk around. To test whether implicit visual statistical learning of spatial information is viewpoint independent, we asked participants to perform a visual search task from variable locations around a monitor placed flat on a stand. Unbeknownst to participants, the target was more often in some locations than others. In contrast to previous research on stationary observers, visual statistical learning failed to produce a search advantage for targets in high-probable regions that were stable within the environment but variable relative to the viewer. This failure was observed even when conditions for spatial updating were optimized. However, learning was successful when the rich locations were referenced relative to the viewer. We conclude that changing viewer perspective disrupts implicit learning of the target's location probability. This form of learning shows limited integration with spatial updating or spatiotopic representations.

Introduction

The human mind's ability to extract and use regularities in complex environments, often in the absence of awareness, is stunningly powerful. From language acquisition to visual perception (Fiser & Aslin, 2001; Reber, 1993; Saffran, Aslin, & Newport, 1996), statistical learning has been characterized as ubiquitous (Turk-Browne, 2012), powerful (Reber, 1993; Stadler & Frensch, 1998), and useful for perceptual and attentive processing (Brady & Chun, 2007; Chun & Jiang, 1998; Goujon, Brockmole, & Ehinger, 2012; Kunar, Flusberg, Horowitz, & Wolfe, 2007; Zhao, Al-Aidroos, & Turk-Browne, 2013). These findings suggest that statistical learning is a critical factor in perceiving and adapting to environmental regularities. Yet only a few studies have directly tested the idea that implicit visual statistical learning (VSL) allows mobile observers to extract visual statistics that are environmentally stable. This study examines the roles of explicit awareness, spatial updating, and reference frames in the ability of mobile observers to learn statistical regularities in the environment.

Perhaps because it can be acquired implicitly and is observed in a variety of domains, statistical learning is often considered an evolutionarily old capability (Reber, 1993). As such, the suggestion that it facilitates basic survival behavior (such as visual search) is appealing. The ability to rapidly learn and use knowledge of where valuable resources are located should be important for survival, as it could optimize effort allocation (Chukoskie, Snider, Mozer, Krauzlis, & Sejnowski, 2013; Smith, Hood, & Gilchrist, 2010). Consistent with this proposal, when a visual search target is more often found in some screen locations than others, people prioritize the "rich" locations even though they are unable to explicitly report where those locations are (Geng & Behrmann, 2005; Jiang, Swallow, Rosenbaum, & Herzig, 2013; Umemoto, Scolari, Vogel, & Awh, 2010).

However, many search tasks involve viewer movements, changing the locations of the rich regions relative to the observer. For VSL to facilitate search in moving observers, it must overcome a difficult computational challenge: representing important locations in a manner that allows them to consistently influence behavior from multiple perspectives. One solution to this problem could be the formation of a map that
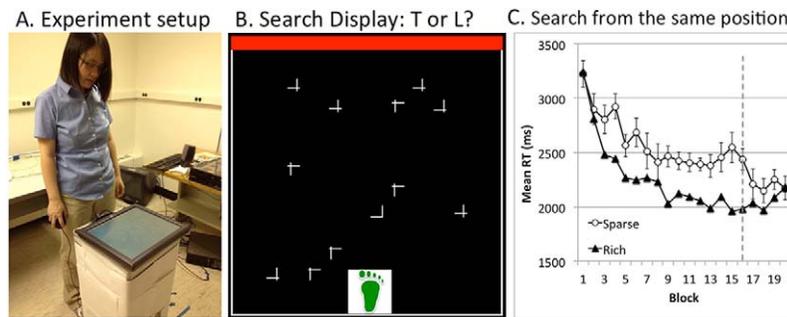
Figure 1. (A) Experimental setup. Participants search for a target on a monitor that was laid flat on a stand. (B) A sample search display from Experiment 1. The green footprint indicates where participants should stand for that trial (in the actual experiment, the footprint preceded, rather than overlaid, the search display). (C) Results from a previous study on observers who always searched from the same position. Adapted from Jiang, Swallow, and Capistrano (2013), Copyright the Association for Research in Vision and Ophthalmology (ARVO©).

codes the rich locations relative to landmarks in the external environment (a "spatiotopic map" or other environment-centered coding; Burr & Morrone, 2012). Alternatively, locations may be coded relative to the viewer but updated or remapped following viewer locomotion or eye movements (Cavanagh, Hunt, Afraz, & Rolfs, 2010; Colby & Goldberg, 1999; Wang & Spelke, 2000; Wurtz, 2008). Both solutions are evidently used to code space in navigation and localization tasks (Burr & Morrone, 2012; Wang & Spelke, 2000; Wurtz, 2008). However, for VSL to facilitate search in moving observers, implicit learning must be integrated with spatiotopic representations or with spatial updating mechanisms. Several recent studies have examined how VSL is used when people move through space. These studies have produced inconsistent findings regarding whether implicit VSL can facilitate search when people search from variable perspectives.

Participants in one recent study searched for a target character among distractor characters on a monitor that was laid flat on a stand (Jiang, Swallow, & Capistrano, 2013). Unbeknownst to the participants, the target was more often in one visual quadrant (50%) than in any one of the other three quadrants (17%). Despite being unable to report which quadrant was more likely to contain the target, participants who always searched from the same position found the target faster when it was in the high-probability, rich quadrant rather than a low-probability, sparse quadrant (Figure 1). These data changed when participants searched from random locations around the monitor on each trial. Under these conditions, participants failed to prioritize the target-rich quadrant. They also lacked explicit awareness about where the target-rich quadrant was. Movement itself did not disrupt learning; when participants moved between trials but always returned to the same starting position: They had no difficulty prioritizing the target-rich quadrant. These

findings showed that changing the viewer's perspective is detrimental to the acquisition of environmentally stable visual statistics.

Two other studies provide suggestive evidence of environment-centered VSL, however. In one study, participants searched for a hidden target light embedded in the floor (Smith et al., 2010). When participants reached a light, they switched it on to discover if it was the target (defined as a particular color). The hidden target was more likely to be in one side of the room (80%) than the other (20%). When their starting position was fixed, participants were able to use this fact to more quickly find a target in the rich side of the room. They showed no learning, however, when their starting position was random and all lights had the same color, suggesting that changes in viewpoint interfered with learning. This was not always the case, however. If different colored lights marked the two sides of the room, participants were able to prioritize the rich side. In this situation, many participants spontaneously reported noticing that the target was unevenly distributed. Another study asked participants to find a coin in a large (64 m²) outdoor environment (Jiang, Won, Swallow, and Mussack, in press). The coin was more often placed in one region of the environment than in other regions. Participants were able to prioritize the target-rich region. Much like the study by Smith et al. (2010) study, participants in the outdoor task were highly accurate in identifying the rich region. These data showed that mobile participants were sometimes capable of acquiring environmental regularities. However, such learning was accompanied by explicit awareness of what was learned.

The contradiction between the three studies reviewed above raises questions about when and why implicit VSL is insensitive to environmentally stable visual statistics. To address this question, the current study systematically examines the roles of explicit awareness, spatial updating, and reference frames in one form of

VSL: location probability learning. We selected location probability learning as our testing paradigm because it is considered an important mechanism for foraging and visual search (Chukoskie et al., 2013; Smith et al., 2010). In addition, spatial location is the key element in location probability learning, making this paradigm ideal for examining the spatial reference frame of implicitly learned visual statistics. Finally, the large effect size of location probability learning under fixed-viewing conditions facilitates the interpretation of potential null results. At the end of the article, we will discuss the generalizability of our findings to other forms of VSL.

First, we tested the hypothesis that the implicit VSL of the target's likely location in a visual environment is more successful if conditions for spatial updating are optimized. In the study by Jiang, Swallow, and Capistrano (2013), participants moved to a random side of the search space on each trial. These unpredictable changes in viewer perspective could have increased the difficulty of spatial updating over the course of the task. Moreover, the visual environment was relatively sparse, reducing the likelihood that landmarks could be used to form an environment-centered reference frame. One goal of the current study is to optimize conditions for environment-centered VSL. To this end, participants made small, predictable perspective changes from one trial to the next (Experiment 1). In addition, we compared performance in a visually rich map search task (Experiment 2) with that in a visually sparse letter search task.

Second, we tested the role of explicit awareness in learning environmental regularities. Although two studies have shown some evidence for environment-centered learning (Jiang et al., in press; Smith et al., 2010), neither examined the relationship between explicit awareness and environment-centered learning. In Experiment 3, we therefore used a task in which participants would be likely to acquire various levels of explicit awareness of the visual statistics. If explicit awareness is important for acquiring an environment-centered learning, those participants who showed greater awareness should also evidence more learning.

Finally, previous studies had little to say about why implicit location probability learning was insensitive to environment-centered visual statistics. One possibility is that this type of learning is intrinsically viewer centered. If this is the case, then visual statistics that are referenced relative to the viewer should be readily acquired. Alternatively, if other factors (such as disorientation) interfered with learning, then moving observers should be unable to acquire any type of visual regularities, including viewer-centered visual statistics. A third goal of this study is to test the source of failure for implicit learning (Experiment 4).

Experiment 1

To examine whether implicit VSL can result in environment-centered learning in mobile observers, participants in Experiment 1 changed their standing position on a trial-by-trial basis in a visual search task. They searched for a rotated letter target (T or L) among symbol distractors (distorted +) and reported whether it was a T or an L. There was one target on each trial. The items were presented on a monitor that was laid flat on a stand (Figure 1A). At the beginning of each trial, participants were cued to stand at one position around the stand (Figure 1B). Experiment 1A served as a replication of the study by Jiang, Swallow, and Capistrano (2013) and also enabled cross-experiment comparisons. In this experiment, the standing position was chosen randomly from the four sides of the monitor. Consequently, the participants' viewpoint could change 0°, 90° clockwise or counter-clockwise, or 180° from one trial to the next. The goal of Experiment 1B was to optimize conditions for spatial updating. In this experiment, the standing position changed in 30° increments along a single direction around the monitor (e.g., clockwise). Perspective change was therefore small and predictable. The two versions of the experiment were otherwise identical.

We divided the experiment into 20 blocks of trials. In the first 16 blocks (mobile phase), across multiple trials, the target was more often located in one target-rich quadrant than in any one of the target-sparse quadrants. The high-probable, target-rich locations were stable within the environment. However, because the standing position changed from one trial to the next, the location of the rich quadrant relative to the participant was variable. Note that changes in perspective occurred between trials rather than during a trial. In the last four blocks (stationary phase), the target was equally likely to appear in any quadrant. In addition, participants always searched from the same position to minimize interference from movements. If an attentional bias toward the target-rich quadrant had developed in the mobile phase, then it should manifest as a persisting preference for the (previously) target-rich quadrant in the stationary phase.

Participants had full access to stable environmental landmarks in the testing room. Room furniture, an experimenter, and a lamp at one corner of the room were constantly in view. One side of the computer monitor was colored red to provide an additional landmark that was readily visible during search. These cues could be used to code the target-rich locations in an environment-centered representation or to update viewer-centered representations following movement. If implicit VSL extracts regularities in an environment-centered fashion, then it should result in faster search response time (RT) when the target appears in the rich

quadrant rather than the sparse quadrants. In contrast, if implicit VSL is viewpoint specific, then it may fail to produce a search advantage in the rich quadrant. Finally, it is possible that a search advantage in the rich quadrant would be found when the perspective change is small and predictable (Experiment 1B) but not when it is large and unpredictable (Experiment 1A). Such findings would suggest that implicit VSL can support environment-centered learning if conditions for spatial updating are optimized.

Method

Participants

Participants in all experiments reported in this study were college students between the age of 18 and 35 years. They were naïve to the purpose of the study, had normal vision, and received $10/hour or extra course credit for their time. No participants performed more than one experiment.

There were 16 participants in each of Experiments 1A and 1B. The estimated statistical power was greater than 0.99 in detecting an effect size as large as those found in previous studies on stationary observers (Cohen's $d = 2.93$; Jiang, Swallow, & Capistrano, 2013, experiment 3).

Equipment

A 17-in. touchscreen monitor (1200 × 860 pixels resolution) was placed flat on a 35-in.-tall stand (Figure 1A). Tape on the floor marked four sides of the stand. Participants responded with a wireless mouse. A 25-watt lamp illuminated the room. Viewing distance varied according to the participant's height and was approximately 55 to 90 cm.

Stimuli

Each display contained 12 items (one target and 11 distractors) placed in randomly selected locations in an invisible 10 × 10 matrix (19 × 19 cm). The items were white presented against a black background. Three items were placed in each quadrant. The target was either a T or an L, and the distractors were distorted plus symbols (+; all items 1.3 × 1.3 cm). All items had a random orientation of 0°, 90°, 180°, or 270°. A red bar (1.3 × 19 cm) on one side of the display provided a consistent landmark (Figure 1B).

Design and procedure

After 10 trials of practice involving random standing positions and random target locations, participants completed 20 blocks of experimental trials, with 24 trials per block. Before each trial, a cue (a green footprint icon in Experiment 1A, or an arrow in Experiment 1B) on the monitor indicated the position that participants should move to (Figure 1B). This position changed from trial to trial in the first 16 blocks but remained the same in the last four blocks. In Experiment 1A, the standing position cue changed randomly to one of four equidistant locations around the monitor. Therefore, from one trial to the next, the participants' viewpoint changed 0°, 90° clockwise or counter-clockwise, or 180°. In Experiment 1B, the standing position cue changed in 30° increments along a consistent direction, clockwise for half of the participants and counter-clockwise for the other half. An experimenter stayed in the room to monitor compliance.

Once in position, participants touched a square in the middle of the monitor to initiate a trial. The touch response required eye-hand coordination and ensured that the eye position returned to the center of the display. The search display appeared 200 ms later and remained in view until participants clicked the mouse to indicate which target (T or L) was there. Trials that lasted more than 10 s were considered outliers, which happened less than 1% of the time in all experiments but Experiment 3. The display was erased after the response. A tone provided feedback about response accuracy.

In the first 16 blocks, the target appeared in one target-rich quadrant on 50% of the trials and in any of the other three target-sparse quadrants on 16.7% of the trials. Which quadrant was rich was counterbalanced across participants but remained the same for a given participant. The target-rich quadrant was environmentally stable and did not change when participants moved to different standing positions. In the last four blocks, the target appeared in each quadrant on 25% of the trials. Participants were not informed about where the target was likely to be.

Recognition

After the search task, participants were asked to select the quadrant where the target was most often found.

Results and discussion

Accuracy

One participant in Experiment 1B had low accuracy (less than 80%); this person's data were excluded from the analysis. For the other participants, search performance was highly accurate (greater than 97%) and was unaffected by the target's quadrant condition ($p > 0.10$ in both Experiments 1A and 1B). This was also the case in all subsequent experiments. Therefore, the analysis
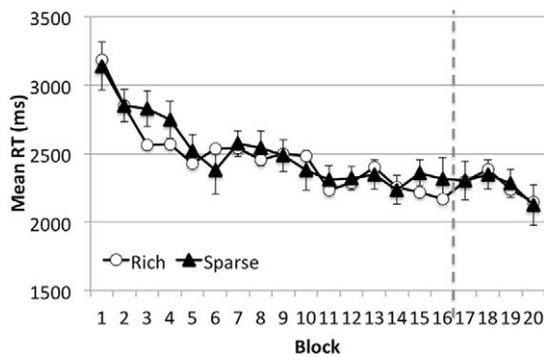
Figure 2. Results from Experiment 1A. Error bars show ±1 SEM of the difference between the target-rich and target-sparse conditions.

in all experiments used the mean RT from correct trials as the dependent measure.

Experiment 1A

In Experiment 1A, participants moved randomly to any side of the monitor before each visual search trial. The target-rich quadrant (in the first 16 blocks) was environmentally stable. Despite the presence of multiple stable environmental landmarks, participants failed to prioritize the search in the rich quadrant (Figure 2). Analysis of variance (ANOVA) on the target's location (rich or sparse quadrant) and block (1–16) showed that RT improved as the experiment progressed, $F(15, 225) = 12.61$, $p < 0.001$, $\eta_p^2 = 0.46$, but it did not differ between target-rich and target-sparse quadrants, $F(1, 15) = 1.25$, $p > 0.25$, and neither did the target quadrant interact with block, $F < 1$. Furthermore, it is unlikely that participants learned where the target was likely to be but were unable to use that knowledge when moving. For the last four blocks, participants stood at a single position, but the target was equally likely to appear in any quadrant. Under these

conditions, VSL should manifest as a persistent attentional bias toward the previously rich quadrant (Jiang, Swallow, Rosenbaum, et al., 2013; Umemoto et al., 2010). This, however, did not occur. The RT in the stationary phase was unaffected by whether the target was in the formerly rich or sparse quadrant, $F < 1$.

Experiment 1B

In Experiment 1A, participants moved a relatively large distance to an unpredictable location on each trial. Both the large change in the perspective and its unpredictability could disrupt spatial updating (Tsuchiai, Matsumiya, Kuriki, & Shioiri, 2012). In Experiment 1B, changes to the participant's search position were both predictable (e.g., they moved in a single direction around the display) and small (30° change from one trial to the next; Figure 3A). However, in the first 16 blocks, RT was similar when the target appeared in the rich quadrant and the sparse quadrants, $F < 1$, and this effect did not interact with block, $F < 1$. The main effect of block was significant, $F(15, 210) = 7.01$, $p < 0.001$, $\eta_p^2 = 0.33$. The stationary phase also did not reveal an attentional bias toward the target-rich quadrant, $F(1, 14) = 1.59$, $p > 0.20$ (Figure 3B).

A direct comparison between Experiments 1A and 1B revealed no main effect of target quadrant, $F < 1$, and no interaction between experiment and target quadrant, $F(1, 29) = 1.16$, $p > 0.25$. Thus, even though participants made small incremental changes in viewpoint, were exposed to many viewpoints around the display, and could predict where they were going next, they were unable to prioritize the target-rich locations of the environment.

Recognition

The percentage of participants who correctly identified the rich quadrant was 12.5% in Experiment 1A
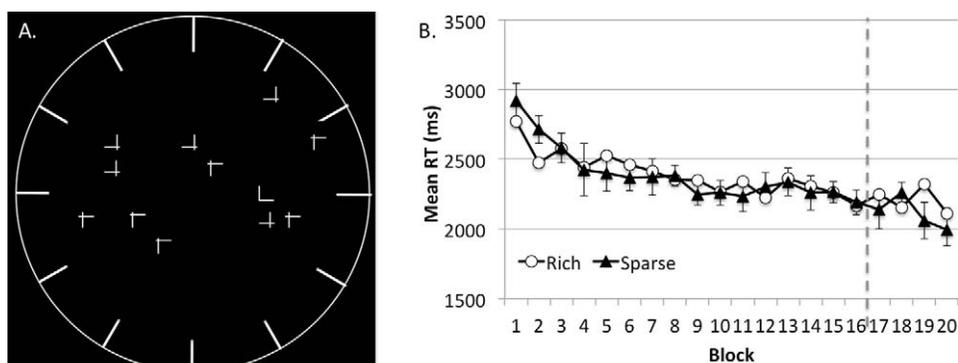

Figure 3. (A) A bird's eye view of the search display and possible standing positions around the monitor (in the actual experiment, the standing position cue preceded, rather than overlaid, the search display). (B) Results from Experiment 1B. Error bars show ±1 SEM of the difference between the target-rich and target-sparse conditions.

and 20% in Experiment 1B. These values did not differ from chance (chance: 25%), ps > 0.15 on a binomial test. Thus, participants did not acquire explicit awareness about where the target was most often found. Recognition performance also did not interact with probability learning (p > 0.10 for data combined between Experiments 1A and 1B).

Comparison with previous findings

Data from Experiment 1 should be contrasted with that of a previous study in which participants searched similar displays from the same standing position (Jiang, Swallow, & Capistrano, 2013, experiment 3). In that study, participants walked halfway toward another side of the monitor and back between trials. Despite this movement, participants found the target faster when it was in the rich quadrant rather than in the sparse quadrants (Figure 1C). The effect was large, Cohen's d = 2.93, and provides a baseline against which null results can be interpreted. We therefore compared data from Experiment 1A with those from our previous study (Jiang, Swallow, & Capistrano, 2013, experiment 3) in an ANOVA using experiment as the between-subject factor and target quadrant (rich or sparse) and block (1–16) as within-subject factors. This test revealed a significant interaction between experiment and target quadrant, $F(1, 30) = 21.37$, $p < 0.001$, $\eta_p^2 = 0.42$.

Discussion

Changing the viewer's perspective from trial to trial disrupts learning, both when those changes are random and large and when they are predictable and small. These findings significantly expanded conditions under which implicit VSL was disrupted. They show that even when conditions for spatial updating are optimal, participants are unable to acquire environmental regularities that cannot be consistently represented relative to the viewer.

Experiment 2

People typically search for useful items in visually rich environments. Although many potential landmarks were available in Experiment 1 (e.g., room furniture), the search display itself was impoverished. Other than a red bar on one side of the monitor, participants did not have other task-relevant cues to reference the target's location. Enriching the search display may be a critical factor for producing environment-centered VSL. In fact, making two sides of a room perceptually distinct produced environment-centered VSL in a previous study (Smith et al., 2010). Participants in Experiment 2 were therefore asked to search a visually rich display (a Google map) for a traffic icon. The target—an icon of a car or a gas station—was most often found in one quadrant of the map, allowing the participants to reliably reference the target's location to the map itself. Similar to Experiment 1A, participants stood at a random location before each trial, so the target-rich locations were variable relative to their perspective. If a visually rich environment is sufficient to produce environment-centered VSL, then participants in Experiment 2 should find the icon faster when it appears in the quadrant that is most likely to contain it.

Methods

Participants

Sixteen new participants (18–35 years old) completed Experiment 2.

Stimuli

Four satellite images (45° aerial view; 1 in. on the map = 20 m in real space) of an unfamiliar university campus (19 × 19 cm) were acquired through Google Maps. Street labels were removed. The maps were visually rich and not symmetrical along any axis. One map was randomly assigned to each participant. This map was displayed for the entire experiment in the same orientation relative to the monitor. The search target was an icon of a car or a gas station (0.6 × 0.6 cm; in one of four orientations, 0°, 90°, 180°, or 270°, randomly selected for each trial). The icon was placed in a randomly selected location of the map. There were 100 possible locations from a 10 × 10 invisible grid that subtended 19 × 19 cm.

Design and procedure

This experiment used the same design as Experiment 1A, except for the stimuli and task. Participants completed 20 blocks of testing (24 trials per block). Just like Experiment 1A, for the first 16 blocks, the target icon was more often placed in one quadrant of the display (50%) than in any one of the other quadrants (16.7%). The rich quadrant was stable relative to the larger environment (e.g., the room) as well as the map itself. Participants changed their standing position from trial to trial randomly around the four sides of the monitor. For the last four blocks, participants stood at the same position to perform the search. In addition, the target icon could appear in any quadrant with equal probability (25%).
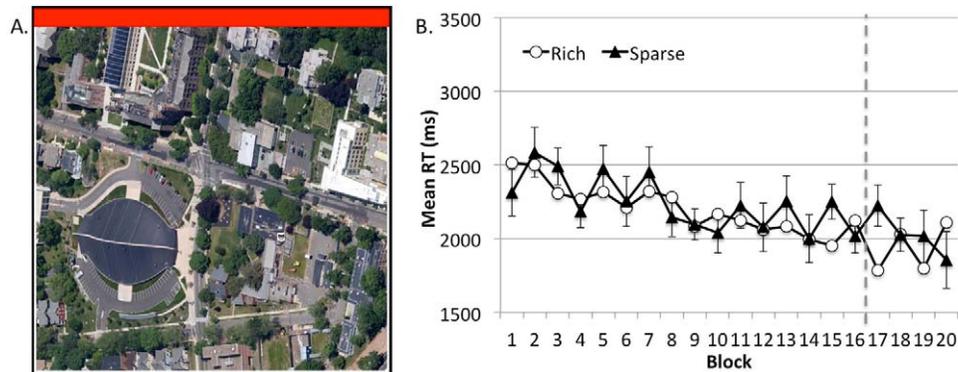
Figure 4. A sample Google Map display (A) and results (B) from Experiment 2. Error bars show ±1 SEM of the difference between the target-rich and target-sparse conditions.

Recognition

At the completion of the visual search task, participants were shown four maps simultaneously (one in each quadrant). They were asked to choose the map that they saw in the experiment. Following this response, the correct map was displayed in the same way as during the search task. Participants were asked to touch the quadrant of the map where the target icon was most often found.

Results

Search RT

The use of a rich search display failed to yield VSL in Experiment 2 (Figure 4). In the first 16 blocks, RT was comparable when the target appeared in a rich quadrant or sparse quadrant, $F < 1$. This similarity in search times did not interact with experimental block, $F(15, 225) = 1.01$, $p > 0.40$, and no effect emerged in the stationary phase either, $F(1, 15) = 1.78$, $p > 0.20$.

Recognition

All participants were able to identify the map that they saw from a set of four maps. However, only 31.3% of the participants were able to identify the target-rich quadrant on the map, which did not differ significantly from chance (25%; $p > 0.15$ on a binomial test). Quadrant recognition accuracy did not interact with location probability learning, $p > 0.10$.

Discussion

The map search task of Experiment 2 provided rich environmental cues that could facilitate landmark-based (including spatiotopic) coding of the environment. Participants actively explored the Google map to spot a traffic icon—a car or a gas station. Over several

hundred trials of visual search, participants had acquired high familiarity with the map. However, they failed to prioritize search in one quadrant of the map that frequently contained the search target. The failure of VSL was also accompanied by a lack of explicit awareness about where the search target was most often found. These data showed that the richness of visual cues, by itself, is unable to produce environment-centered VSL.

Experiment 3

The failure of implicit VSL in Experiments 1 and 2 suggests that this form of learning is not environment centered. The observation that changing viewer perspective disrupts implicit VSL poses significant constraints on how much it could influence search in everyday contexts (e.g., when searching through a produce section for a preferred kind of apple). As we will discuss later, these data have important implications for understanding the function of implicit VSL. However, they also run counter to the intuition that people should be able to prioritize important locations during search even when their standing positions change. In fact, two previous studies that tested participants in a large environment have evidenced some degree of environment-centered VSL (Jiang et al., in press; Smith et al., 2010). In both of those studies, however, participants were highly accurate in explicitly recognizing the target-rich region.

The goal of Experiment 3 is to examine whether participants can spontaneously acquire environment-centered VSL and, if so, whether explicit awareness correlates with the degree of learning. Experiment 3 used a statistical learning paradigm that was known to produce explicit knowledge (Brockmole, Castelhano, & Henderson, 2006; Ehinger & Brockmole, 2008). In this experiment, participants searched for a small green

letter overlaid on a natural scene. Multiple scenes and multiple target locations were used, each appearing 20 times over the course of the experiment. For half of the scenes (the *old* condition), the target's location was consistent within the scene across all repetitions. For the other half of the scenes (the *shuffled* condition), the mapping between the scene and the target's location was variable over time. By associating a single target location with a unique scene in the *old* condition, this design increased the likelihood that participants would become aware of where the target was in the scene (Brockmole et al., 2006). Similar to Experiments 1 and 2, participants' standing position changed randomly from trial to trial. Thus, the scene and the target's location were variable relative to the participants. If explicit awareness facilitates environment-centered VSL, then as long as participants became aware of the scene-target association, they should acquire learning in Experiment 3. Furthermore, individuals who showed greater awareness of the scene-target association should demonstrate greater VSL. In contrast, if participants are completely unable to acquire environment-centered learning in our setup (perhaps as a result of disorientation), then VSL should fail even in people who became aware of the scene-target association.

Method

Participants

Sixteen new participants (18–25 years old) completed Experiment 3.

Stimuli

Sixteen scenes were selected randomly for each participant from a set of 48 indoor and outdoor scenes. In addition, 16 target locations were randomly chosen from 100 possible locations (10 × 10 invisible grid subtending 19 × 19 cm), with the constraint that there were an equal number of possible target locations in each visual quadrant. The target was a small green letter (0.4 × 0.4 cm), either T or L, presented in one of four possible orientations (0°, 90°, 180°, or 270°). Because the participant's standing position could be at any side of the monitor, the scenes had various orientations such that four scenes appeared upright to participants when standing at any side. The orientation of the scene was environmentally stable and did not change when participants moved.

Design and procedure

The experiment was divided into 20 blocks of trials with 16 trials per block. Each trial started with the standing position cue that signaled the participants to move to one side of the monitor. The standing position was randomly chosen on each trial. Once in position, participants initiated the trial by touching the central fixation point. Two hundred milliseconds later, a scene was presented along with a green letter overlaying on it. Participants reported whether the green letter was a T or an L by pressing the left or right mouse button. The display was erased upon the mouse click response or after 10 s, whichever occurred earlier. The 10-s time-out cutoff was chosen to keep the experiment from becoming excessively long. An illustration of the procedure and stimuli can be found at http://jianglab. psych.umn.edu/ViewpointSpecificity.mov. It may be necessary to pause the video to find the target.

Each block of trials involved 16 unique scenes and 16 unique target locations. The same scenes and target locations were used in all blocks. Half of the scenes and target locations were assigned to the *old* condition. The other half were assigned to the *shuffled* condition. In the *old* condition, a different target location was assigned to each of the eight scenes; this mapping was held constant across the 20 blocks of trials. In the *shuffled* condition, the mapping between the remaining eight scenes and the target's locations was variable. Participants received no information about the experimental manipulation and were asked to find the letter as quickly and as accurately as possible.

There were four practice trials before the experiment to familiarize participants with the search task. Scenes used in the practice trials differed from those used in the main experiment.

Recognition test

At the completion of the visual search task, participants were presented with 24 scenes one at a time: eight new scenes, eight scenes from the *old* condition, and eight scenes from the *shuffled* condition, in a random order. Participants were asked to first report whether they had seen the scene in the experiment. Following that response, they were asked to touch the location of the target letter on that scene. Participants were told to make a guess about the target's location if they were not sure or if they thought they had not seen the scene before. To estimate localization accuracy, we calculated the Euclidian distance between the participant's response and the actual target's location. The correct target location for *old* scenes was the location where the target was placed on that scene during the experiment. The correct target location for *shuffled* and *new* scenes was a randomly selected location (from the set of eight possible target locations used for the *shuffled* scenes or from the set of 16 possible target locations for the *new* scenes).
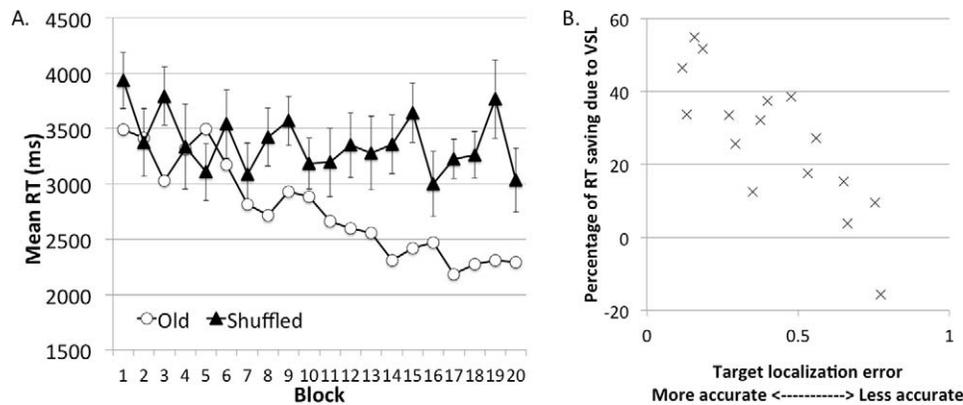
Figure 5. Results of Experiment 3. (A) RT during the visual search task. Error bars show ±1 SEM of the difference between the old and shuffled conditions. (B) The size of VSL during the visual search (shuffled RT minus old RT) correlated with explicit awareness (the localization error on new scenes minus the localization error on old scenes). Each "x" represents data from one participant.

Results and discussion

Search accuracy

Participants made few incorrect responses: 2.5% in the *old* condition and 2.4% in the *shuffled* condition. This difference was not significant, $p > 0.50$. However, participants were timed-out (failure to respond within 10 s) on 6.1% of the trials in the *old* condition and 15.9% of the trials in the *shuffled* condition, a difference that was significant, $p < 0.001$. In the following RT analysis, we report data from trials receiving a correct response. The pattern of results was the same if timed-out trials were included and were assigned the maximum time (10 s).

Search RT

Across all participants (Figure 5A), RT was significantly faster in the *old* condition than in the *shuffled* condition, $F(1, 15) = 18.96$, $p < 0.001$, $\eta_p^2 = 0.56$. This effect increased as the experiment progressed, yielding a significant interaction between condition (old or shuffled) and block, $F(19, 285) = 3.11$, $p < 0.001$, $\eta_p^2 = 0.17$. The RT also improved in later blocks compared with earlier blocks, $F(19, 285) = 4.90$, $p < 0.001$, $\eta_p^2 = 0.25$. Thus, unlike Experiments 1 and 2, VSL facilitated visual search in Experiment 3 even though participants changed their perspective from trial to trial.

Scene recognition and scene-target localization error

Participants were highly accurate in the scene recognition task. They correctly recognized 88.3% of the *old* scenes and 94.5% of the *shuffled* scenes, with a 3.3% false alarm rate for *new* scenes. The correct recognition rate was comparable between the *old* and *shuffled* scenes ($p > 0.08$) and was substantially higher than the false alarm rate ($p < 0.001$).

To examine whether people became explicitly aware of the consistent scene-target association in the *old* condition, we measured the target localization error. In units of pixels, the mean localization error was 141 pixels on *old* scenes, which was significantly less than 347 pixels on *new* scenes ($p < 0.001$) or 325 pixels on *shuffled* scenes ($p < 0.001$). Thus, averaged across all participants, there was evidence that people had acquired explicit knowledge of the target's location on old scenes.

Explicit awareness and VSL

To examine whether explicit awareness about the scene-target association had supported the learning shown in the search task, we computed the correlation between the size of VSL and the degree of explicit awareness. VSL was indexed by the difference in search RT between the *old* and *shuffled* conditions. This was calculated for blocks 11 to 20, during which learning had stabilized. Explicit awareness was indexed by target localization error (in units of pixels) in *new* scenes minus localization error in *old* scenes. As seen in Figure 5B, individuals who were better able to explicitly report where the target was in the old scene also showed a larger RT benefit in the *old* condition, Pearson's $r = 0.79$, $p < 0.001$.

Discussion

Experiment 3 demonstrates that changes in viewpoint do not always interfere with environment-centered VSL. These results were found even when the search task was conducted on a computer monitor, rather than in a large space. Results from Experiment 3 can be contrasted with those of Experiments 1 and 2. We believe that the main difference lies in explicit
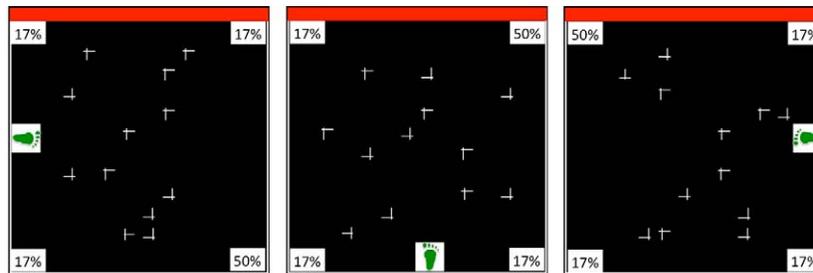
Figure 6. An illustration of the location probability manipulation used in Experiment 4. The displays show three different trials. As indicated by the footprint location, participants stand at different positions around the flat monitor. The target-rich quadrant was referenced relative to the participants' standing position (in this example, always in their upper right).

awareness. Whereas participants in Experiments 1 and 2 were unaware of where the target was most likely located, participants in Experiment 3 acquired explicit awareness. In fact, those who had greater awareness (as indexed by more accurate localization of the target on a scene) also showed greater VSL.

The approached used in Experiment 3 was intrinsically correlational: participants who showed greater VSL also had greater awareness of where the target was in the scene. The experiment did not address whether explicit knowledge causes environment-centered learning (or vice versa). In addition, it did not rule out the possibility that learning the one-to-one mapping between a scene and a target location (Experiment 3) was easier than learning the rich quadrant across all trials (Experiments 1 and 2). However, it seems unlikely that visual statistics used in Experiment 3 were intrinsically easier to learn. The ease of computing summary statistics might lead to the opposite prediction (e.g., Alvarez & Oliva, 2008). In fact, when explicit learning is required, participants did not rapidly learn to associate a specific location with a specific spatial configuration (Chun & Jiang, 2003) or a specific scene (Brockmole & Henderson, 2006). In addition, an earlier study provided a stronger test of whether explicit knowledge enables environment-centered learning in the same task as that of Experiments 1 and 2 (Jiang, Swallow, & Capistrano, 2013, experiment 4). This experiment was identical to Experiment 1, except that participants were explicitly told where the target was likely to appear before they began the experiment. If explicit knowledge of the target's likely location promotes environment-centered learning, then these participants should show an advantage for the target-rich quadrant even though they moved to a random search position for each trial. This was the case.

In contrast to this earlier study, Experiment 3 is the first to show that participants are able to spontaneously acquire environment-centered learning even in the absence of explicit instructions. These data show that incidental learning of environment-centered visual statistics is possible. However, the strong correlation between the size of learning and the participants' level

of awareness points to the explicit nature of the learning. Together with previous findings on explicit learning in VSL (Jiang, Swallow & Capistrano, 2013; Jiang et al., in press), Experiment 3 provides strong evidence that (a) environment-centered VSL is possible but (b) such learning likely depends on explicit awareness of the underlying statistics. As we will show next, these results contrast with viewer-centered learning, which does not depend on explicit awareness.

Experiment 4

The first three experiments suggested that in mobile observers, environment-centered learning depends on explicit awareness of where target-rich locations are. Yet these experiments say little about the spatial representation that implicit VSL uses. One possibility is that unlike explicit learning, implicit VSL codes attended locations in a viewer-centered coordinate system that is not updated when participants move to a new location (Jiang & Swallow, 2013b; Jiang, Swallow, & Sun, 2014). If this is true, then moving observers should be able to learn statistical regularities that are referenced relative to their perspective. Therefore, in Experiment 4, we modified Experiments 1A (letter search) and Experiment 2 (map search) by referencing the target-rich quadrant relative to the participant. The display and the task were identical to those of Experiments 1 and 2. What "moved" was the target's location probability relative to the world. Figure 6 illustrates the experimental design. For example, for a quarter of the participants, the target-rich quadrant would always be in their upper right visual field regardless of where they stood. Because the searcher's viewpoint was random, the target-rich quadrant's location was random relative to the screen and environment.

If implicit VSL is viewer centered, then a search advantage for the target-rich, viewer-centered quadrant should develop during the training session (i.e., RTs should improve more when the target appears in the
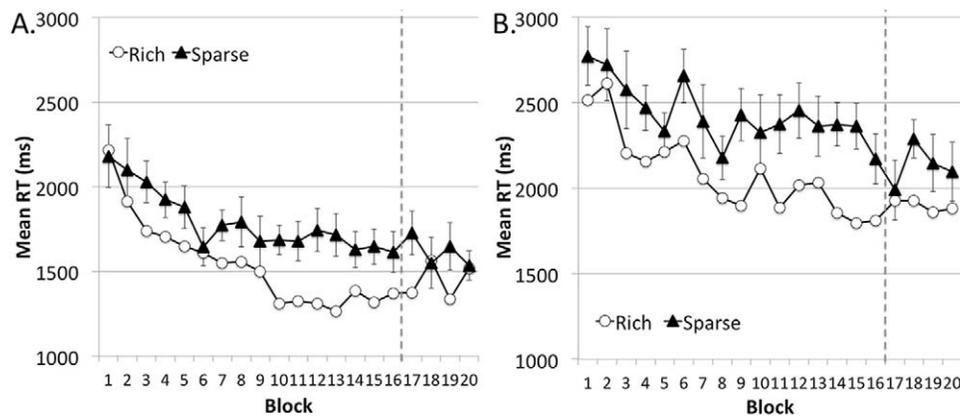
Figure 7. Results from the letter search task of Experiments 4A (A) and the map search task of Experiment 4B (B). The target-rich region was variable in the external environment but consistent relative to the viewer's standing position. Error bars show ±1 SEM of the difference between the target-rich and target-sparse conditions.

target-rich quadrant than in the target-sparse quadrant). In contrast, if the visual statistics must be stable in the environment for learning to occur, or if other factors (such as disorientation) interfere with performance, then no learning should be evident.

Method

Participants

Thirty-two new participants were tested, 16 in Experiment 4A and 16 in Experiment 4B.

Stimuli

Experiment 4A used the same letter-among-symbol stimuli as in Experiment 1. Experiment 4B used the same icon-on-a-map stimuli as in Experiment 2.

Design and procedure

Experiment 4A was the same as Experiment 1A, except that the target-rich quadrant was defined relative to the viewer (e.g., it was always to the participant's upper left, regardless of his or her standing position). In Blocks 1 to 16, the target appeared in the viewer-centered rich quadrant 50% of the time (the other quadrants were equally likely to contain a target, 16.7%). In addition, participants moved to a new, randomly determined standing position before each trial. In Blocks 17 to 20, the target was equally likely to appear in any quadrant (25%), and participants stood in one position. Like Experiment 1A, the first 16 blocks tested whether participants could acquire implicit VSL when moving, and the last four blocks tested whether the learned attentional bias persisted in a stationary phase.

Experiment 4B was the same as Experiment 2, except that the target-rich quadrant was consistently referenced relative to the participant and therefore was random relative to the map. Similar to Experiment 2, the Google map itself did not change orientation. However, where the target was most often found depended on the participant's standing position. To make this concrete, imagine the screen as a clock with its face pointed to the ceiling. If the target-rich quadrant was to the participant's upper left, then it would be the area between the 9 and the 12 when he or she stood at the 6-o'clock position. If the participant stood at the 9-o'clock position the target-rich quadrant would be the area between the 12 and the 3, and so on.

Recognition tests

Recognition tests were the same as those of Experiments 1A and Experiment 2. However, the target-rich quadrant was defined relative to the participant rather than the external environment. For example, if the target-rich quadrant was in the upper left visual field relative to the participant, then the correct recognition response would be to touch the part of the screen corresponding to the upper left visual field.

Results

Experiment 4A search RT

The data provided clear evidence that VSL developed rapidly when statistical regularities were stable relative to the viewer (Figure 7A). In Experiment 4A's letter search task, participants were significantly faster when the target appeared in the viewer-centered, target-rich quadrant rather than the target-sparse quadrants in the mobile phase, $F(1, 15) = 17.12$, $p < 0.001$, $\eta_p^2 =$

0.53. This advantage was significant as early as Block 3 ($p < 0.03$) and did not significantly interact with block, $F(15, 225) = 1.21$, $p > 0.25$. The attentional bias toward the target-rich quadrant persisted even after the target became evenly distributed in the stationary phase, demonstrating long-term persistence of viewer-centered VSL, $F(1, 15) = 7.90$, $p < 0.02$, $\eta_p^2 = 0.35$.

Experiment 4B search RT

The findings in the letter search task were replicated in the map search task of Experiment 4B (Figure 7B). Even though a visually rich environment (the map) might be expected to interfere with viewer-centered learning, participants were significantly faster finding the car/gas icon when it appeared in the viewer-centered rich quadrant in the first 16 blocks, $F(1, 15) = 32.72$, $p < 0.001$, $\eta_p^2 = 0.69$. This effect stabilized early and did not interact with block, $F < 1$. It also persisted in the stationary phase, $F(1, 15) = 15.44$, $p < 0.001$, $\eta_p^2 = 0.51$.

Recognition test

The percentage of participants who correctly identified the rich quadrant was 18.7% in Experiment 4A and 18.7% in Experiment 4B, which did not differ significantly from chance, $p > 0.10$. Recognition performance did not interact with the size of VSL, $p > 0.10$ in both experiments.

Thus, the presence of VSL in Experiment 4 could not be attributed to increased explicit awareness of the experimental manipulation. Instead, these data suggest that implicit VSL is egocentric. Information that can be consistently referenced relative to the observer accumulates over time to facilitate visual search, even though this information is random in an environment-centered reference frame.

General discussion

In most studies, statistical learning occurs without an intention to learn or an awareness of what was learned. In fact, implicit learning is a powerful mechanism for extracting statistical regularities, even in young infants, the elderly, and brain-damaged patients (Perruchet & Pacton, 2006; Reber, 1993; Saffran et al., 1996). Indeed, the statistics used in Experiments 1 and 2 are easily acquired when participants are stationary (Geng & Behrmann, 2005; Jiang, Swallow, Rosenbaum, et al., 2013) and when they move halfway to one position and back between each trial (Jiang, Swallow, & Capistrano, 2013). Effect sizes were large in those studies. Importantly, VSL is functionally beneficial in these

tasks. Because it has such a large effect on search times (it sped up RT by ~20% in Experiments 3 and 4), VSL could optimize effort allocation.

This study demonstrates, however, that under some circumstances, implicit VSL fails to extract environmental regularities. Environmentally stable regularities are learned and used to facilitate visual search when the viewer is stationary (Geng & Behrmann, 2005; Jiang, Swallow, Rosenbaum, et al., 2013; Umemoto et al., 2010). Once the viewer assumes variable perspectives, location probability learning occurs only for visual statistics that maintain viewpoint consistency, or when the viewer becomes explicitly aware of the environmental regularities. The lack of environment-centered implicit VSL for environmentally stable statistics is striking. It also provides significant constraints on what VSL might reflect and what it might be used for.

The failure of implicit VSL to acquire environmental regularities in Experiments 1 and 2 indicates that it is not well integrated with spatiotopic coding or spatial updating mechanisms. Two conditions are considered important for establishing spatiotopic representations (Burr & Morrone, 2012). First, attention should be available to form the spatiotopic map. Second, there should be sufficient time for establishing a viewer-invariant representation. Experiments 1 and 2 met these conditions (participants attended to the display, they had time to update their representations after moving to the new position, and environmental landmarks were present). However, learning was disrupted. This is not to say that spatiotopic representations could not form in some tasks and conditions. They just do not appear to support implicit VSL in conditions tested in our study.

Even without a spatiotopic map of the environment, however, it is possible that implicit VSL could prioritize a region of space relative to the external environment through spatial updating. Spatial updating often occurs in anticipation of, or following, an eye movement or perspective change (Cavanagh et al., 2010; Colby & Goldberg, 1999; Wurtz, 2008). Experiments 1 and 2 were designed to support spatial updating: viewpoint changes were introduced by viewer locomotion rather than display rotation, plenty of environmental cues were present, and in one experiment, small, predictable changes in perspective were used. Yet participants failed to learn where the target-rich region was. These data indicate that implicit VSL of the type investigated here is poorly integrated with spatial updating.

Because the viewpoint manipulation is not commonly used in visual search experiments, it is important to consider whether it might have led participants to expect that we were looking for viewpoint dependence (or, alternatively, viewpoint independence). This seems unlikely for several reasons. First, any expectations participants develop as a result of trying to guess the

purpose of the experiment should result in explicit attentional biases. Yet there was no evidence that learning depended on explicit knowledge of the target's location probability. Other data further emphasize the distinction between explicit expectations and probability cuing: Viewer-centered probability cuing persists even after participants are explicitly told to expect the target in an environment-rich quadrant (Jiang et al., 2014). Just as importantly, viewpoint manipulations do not exclusively lead to viewer-centered learning. In previous studies, tilting one's body and head through the vertical plane induced some environment-centered learning (Jiang & Swallow, 2013a), as did performing the task in an outdoor environment (Jiang et al., in press). In both of these studies, environment-centered learning was associated with above-chance recognition rates. These data suggest that if participants had set up expectations about the experimenter's intent, such expectations would more likely produce environment-centered rather than viewer-center learning.

Our data suggest that the functional significance of implicit VSL may be very different than previously supposed. Implicit statistical learning is often considered an important mechanism for extracting statistical regularities that are stable in the environment, such as frequently co-occurring sounds (Saffran et al., 1996) or consistently paired objects (Fiser & Aslin, 2001; Turk-Browne, 2012). As a result, unsupervised VSL may be useful for representing the hierarchical structure of features, objects, and their relationship in the external world (Fiser & Aslin, 2005; Orbán, Fiser, Aslin, & Lengyel, 2008). Although the view of VSL as a mechanism for "extracting what is out there" is satisfactory in paradigms that involve perception only, it is more difficult to resolve with tasks that involve covert or overt action. Visual search, for example, is more than just perceiving what is in the visual world. It also involves actively shifting covert and overt attention between items until the target object is found. Unlike visual perception (which may be centered on objects or environmental locations), visual action is inherently egocentric (Goodale & Haffenden, 1998). Consistent with this claim, spatial attention is referenced primarily egocentrically (Golomb, Chun, & Mazer, 2008; Jiang & Swallow, 2013b; Mathôt & Theeuwes, 2010). Therefore, the function of VSL in active tasks is unlikely to be simply about perception. Rather, VSL may serve to increase the likelihood that successful actions will repeat in the future. On this interpretation, VSL did not accumulate in Experiments 1 and 2 because the target-rich region was randomly located in a viewer-centered (possibly even action-centered) reference frame. In contrast, VSL was present in Experiment 4 because the statistical regularities were stable relative to the viewer. As a result, successful

movements of attention through space were relatively consistent across trials.

Locations that are rich in food and other resources should be independent of the viewer's perspective. It is therefore puzzling that a statistical learning system would remain egocentric when evolutionary pressures should encourage the formation of environment-centered representations. Three explanations may jointly explain this puzzle. First, the external pressure for extracting truly viewer-independent representations may not be as great as imagined. Both natural and manmade environments are highly constrained and limit the number of possible viewpoints one may have. A viewer-centered representation may suffice in these situations. Second, viewer-centered representations can be advantageous. They are easy to compute because neurons in occipital and parietal cortices are predominantly retinotopic (Saygin & Sereno, 2008). In addition, a major purpose of vision is or guiding visuomotor action. Because actions carried out by one's eyes, hands, and body are predominantly referenced relative to the viewer (Goodale & Haffenden, 1998), an egocentric visual learning system can more easily and rapidly interface with motor systems. Finally, the egocentric system is complemented by additional mechanisms that can be referenced to the external environment. Specifically, awareness of environmental regularities allows a person to prioritize locations that are likely to contain a target, even as their viewpoint changes.

The current study is an initial step toward understanding when and why implicit VSL fails. Two previous studies have shown some (limited) success for environment-centered VSL (Jiang et al., in press; Smith et al., 2010), whereas a third study has demonstrated a complete failure (Jiang, Swallow, & Capistrano, 2013). Our study showed that the failure of environment-centered learning is unlikely attributable to the lack of environmental cues (Experiment 2), to abrupt changes in viewpoint (Experiment 1B), or to general disruptions such as spatial disorientation (Experiment 4). Rather, explicit awareness about the underlying statistical regularity may be key to overcoming viewpoint dependency.

This study leaves open the possibility that large-scale environments, such as the ones used by Jiang et al. (in press) and Smith et al. (2010), may facilitate environment-centered representations. In a large environment, most of the search space is beyond the arm's reach; it is in the person's "perception space" rather than "action space." In addition, search often entails active viewer movement within the search space. Under such conditions, visual search may rely more on external environmental cues and may be less constrained by the viewer's perspective than what was found here. To firmly test the role of implicit VSL in natural behaviors

such as foraging, it is important to extend the current findings from laboratory testing to the real world while minimizing explicit awareness of the underlying statistics.

Furthermore, our study has not exhaustively tested conditions that promote environment-centered VSL. Successful learning in Experiment 3 suggests that people can acquire environment-centered VSL. A potential explanation for this learning is that participants may have viewed the background images as task relevant, and this could have promoted environment-centered learning. However, it is unclear why participants in Experiment 2 would not also use the background scene, as it was also predictive of the target's likely locations. Further research is needed to directly examine the effects of task relevance on environment-centered learning. This research will also need to examine whether the effects of task relevance are mediated by explicit awareness.

VSL involves diverse paradigms and mechanisms. In some paradigms, spatial locations are entirely irrelevant, such as learning the association between two novel shapes (Fiser & Aslin, 2001; Turk-Browne, 2012). In other paradigms, the locations of targets are defined relative to the spatial context of the locations of other items (e.g., contextual cueing; Chun & Jiang, 1998). Compared with location probability learning, other paradigms have received less systematic investigation regarding the viewpoint specificity of learning. Several findings have suggested that viewpoint specificity may generalize to other paradigms of implicit VSL. For example, in two studies, contextual cueing was disrupted when participants' viewpoint changed (Chua & Chun, 2003; Tsuchiai et al., 2012). In addition, associative learning of a pair of visual objects was eliminated when the pair was turned upside down (Vickery & Jiang, 2009). Additional research is needed to test whether all forms of implicit spatial learning are viewer centered.

Conclusion

This study shows that changes in an observer's perspective disrupt the ability to implicitly learn visual statistical regularities that are stable in the environment. When testing is done on a computer, VSL of spatial locations is egocentric unless explicit awareness of the regularities is acquired. Thus, the conditions for successful VSL may be more limited than what is suggested by contemporary research. The role of VSL may not be to extract statistical regularities that are stable in the environment but to increase the likelihood that successful behavior will be repeated. Future

research should examine whether implicit VSL remains egocentric in large-scale, real-world search tasks.

Keywords: statistical learning, visual attention, spatial updating, viewpoint specificity

Corresponding author: Yuhong V. Jiang.
Email: Jiang166@umn.edu.
Address: Department of Psychology, University of Minnesota, Minneapolis, MN, USA.

References

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science, 19*, 392–398. doi:10.1111/j.1467-9280.2008.02098.x.

Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: Modeling contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 798–815. doi:10.1037/0096-1523.33.4.798.

Brockmole, J. R., Castelhano, M. S., & Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(4), 699–706. doi:10.1037/0278-7393.32.4.699.

Brockmole, J. R., & Henderson, J. M. (2006). Recognition and attention guidance during contextual cueing in real-world scenes: evidence from eye movements. *Quarterly Journal of Experimental Psychology (2006), 59*, 1177–1187. doi:10.1080/17470210600665996.

Burr, D. C., & Morrone, M. C. (2012). Constructing stable spatial maps of the world. *Perception, 41*, 1355–1372.

Cavanagh, P., Hunt, A. R., Afraz, A., & Rolfs, M. (2010). Visual stability based on remapping of

attention pointers. *Trends in Cognitive Sciences, 14*, 147–153. doi:10.1016/j.tics.2010.01.007.

Chua, K.-P., & Chun, M. M. (2003). Implicit scene learning is viewpoint dependent. *Perception & Psychophysics, 65*, 72–80.

Chukoskie, L., Snider, J., Mozer, M. C., Krauzlis, R. J., & Sejnowski, T. J. (2013). Learning where to look for a hidden target. *Proceedings of the National Academy of Sciences, USA, 110*(Suppl. 2), 10438–10445. doi:10.1073/pnas.1301216110.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*, 28–71. doi:10.1006/cogp.1998.0681.

Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 224–234.

Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience, 22*, 319–349. doi:10.1146/annurev.neuro.22.1.319.

Ehinger, K. A., & Brockmole, J. R. (2008). The role of color in visual search in real-world scenes: Evidence from contextual cuing. *Perception & Psychophysics, 70*, 1366–1378. doi:10.3758/PP.70.7.1366.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*, 499–504.

Fiser, J., & Aslin, R. N. (2005). Encoding multielement scenes: statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General, 134*, 521–537. doi:10.1037/0096-3445.134.4.521.

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics, 67*, 1252–1268.

Golomb, J. D., Chun, M. M., & Mazer, J. A. (2008). The native coordinate system of spatial attention is retinotopic. *Journal of Neuroscience, 28*, 10654–10662. doi:10.1523/JNEUROSCI.2525-08.2008.

Goodale, M. A., & Haffenden, A. (1998). Frames of reference for perception and action in the human visual system. *Neuroscience and Biobehavioral Reviews, 22*, 161–172.

Goujon, A., Brockmole, J. R., & Ehinger, K. A. (2012). How visual and semantic information influence learning in familiar contexts. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 1315–1327. doi:10.1037/a0028126.

Jiang, Y. V., & Swallow, K. M. (2013a). Body and head tilt reveals multiple frames of reference for spatial attention. *Journal of Vision, 13*(13):9, 1–11, http://www.journalofvision.org/content/13/13/9, doi:10.1167/13.13.9. [PubMed] [Article]

Jiang, Y. V., & Swallow, K. M. (2013b). Spatial reference frame of incidentally learned attention. *Cognition, 126*, 378–390. doi:10.1016/j.cognition.2012.10.011.

Jiang, Y. V., Swallow, K. M., & Capistrano, C. G. (2013). Visual search and location probability learning from variable perspectives. *Journal of Vision, 13*(6):13, 1–13, http://www.journalofvision.org/content/13/6/13, doi:10.1167/13.6.13. [PubMed] [Article]

Jiang, Y. V., Swallow, K. M., Rosenbaum, G. M., & Herzig, C. (2013). Rapid acquisition but slow extinction of an attentional bias in space. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 87–99. doi:10.1037/a0027611.

Jiang, Y. V., Swallow, K. M., & Sun, L. (2014). Egocentric coding of space for incidentally learned attention: Effects of scene context and task instructions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 233–250. doi:10.1037/a0033870.

Jiang, Y. V., Won, B.-Y., Swallow, K. M., & Mussack, D. M. (in press). Spatial reference frame of attention in a large outdoor environment. *Journal of Experimental Psychology: Human Perception and Performance*. doi:10.1037/a0036779.

Kunar, M., Flusberg, S., Horowitz, T., & Wolfe, J. (2007). Does contextual cuing guide the deployment of attention? *Journal of Experimental Psychology, 33*, 816–828. doi:10.1037/0096-1523.33.4.816.

Mathôt, S., & Theeuwes, J. (2010). Gradual remapping results in early retinotopic and late spatiotopic inhibition of return. *Psychological Science, 21*, 1793–1798. doi:10.1177/0956797610388813.

Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences, USA, 105*, 2745–2750. doi:10.1073/pnas.0708424105.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10*, 233–238. doi:10.1016/j.tics.2006.03.006.

Reber, A. S. (1993). *Implicit learning and tacit knowledge an essay on the cognitive unconscious.* New York: Oxford University Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.

Saygin, A. P., & Sereno, M. I. (2008). Retinotopy and

attention in human occipital, temporal, parietal, and frontal cortex. *Cerebral Cortex, 18*, 2158–2168. doi:10.1093/cercor/bhm242.

Smith, A. D., Hood, B. M., & Gilchrist, I. D. (2010). Probabilistic cuing in large-scale environmental search. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 605–618. doi: 10.1037/a0018280.

Stadler, M. A., & Frensch, P. A. (1998). *Handbook of implicit learning*. Thousand Oaks, CA: Sage.

Tsuchiai, T., Matsumiya, K., Kuriki, I., & Shioiri, S. (2012). Implicit learning of viewpoint-independent spatial layouts. *Frontiers in Psychology, 3*, 207. doi: 10.3389/fpsyg.2012.00207.

Turk-Browne, N. B. (2012). Statistical learning and its consequences. *Nebraska Symposium on Motivation, 59,* 117–146.

Umemoto, A., Scolari, M., Vogel, E. K., & Awh, E. (2010). Statistical learning induces discrete shifts in the allocation of working memory resources. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1419–1429. doi:10. 1037/a0019324.

Vickery, T. J., & Jiang, Y. V. (2009). Associative grouping: Perceptual grouping of shapes by association. *Attention, Perception & Psychophysics, 71*, 896–909. doi:10.3758/APP.71.4.896.

Wang, R. F., & Spelke, E. S. (2000). Updating egocentric representations in human navigation. *Cognition, 77*, 215–250.

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research, 48*(20), 2070–2089. doi: 10.1016/j.visres.2008.03.021.

Zhao, J., Al-Aidroos, N., & Turk-Browne, N. B. (2013). Attention is spontaneously biased toward regularities. *Psychological Science, 24*, 667–677. doi:10.1177/0956797612460407.