Journal of Experimental Psychology:
Human Perception and Performance
2005, Vol. 31, No. 6, 1439–1448

0096-1523/05/$12.00 DOI: 10.1037/0096-1523.31.6.1439

Hyperspecificity in Visual Implicit Learning: Learning of Spatial Layout Is Contingent on Item Identity

Yuhong Jiang and Joo-Hyun Song
Harvard University

Humans conduct visual search faster when the same display is presented for a 2nd time, showing implicit learning of repeated displays. This study examines whether learning of a spatial layout transfers to other layouts that are occupied by items of new shapes or colors. The authors show that spatial context learning is sometimes contingent on item identity. For example, when the training session included some trials with black items and other trials with white items, learning of the spatial layout became specific to the trained color—no transfer was seen when items were in a new color during testing. However, when the training session included only trials in black (or white), learning transferred to displays with a new color. Similar results held when items changed shapes after training. The authors conclude that implicit visual learning is sensitive to trial context and that spatial context learning can be identity contingent.

Keywords: visual attention, visual search, implicit learning, spatial context learning

Research in the past decade has shown that although the human visual system is surprisingly limited in some respects, such as in detecting large-scale changes in natural viewing (Simons & Levin, 1998), it is also remarkably powerful in other respects. Humans are quick at extracting the gist of a scene (Potter, 1976), accurate at detecting specific objects within the scene (Thorpe, Fixe, & Marlot, 1996), and reliable at retaining such information for a relatively long time (Hollingworth & Henderson, 2002). Such rapid and effortless perception is further enhanced by a learning system that extracts regularities in visual displays. A single glimpse is sufficient for people to estimate the mean size or velocity of an array of dots (Ariely, 2001; Atchley & Andersen, 1995; Chong & Treisman, 2003). In addition, a few repetitions are enough for humans to learn the consistent association between object pairs (Chun & Jiang, 1999; Fiser & Aslin, 2001), motion trajectories (Chun & Jiang, 1999), temporal sequences (Howard, Howard, Dennis, Yankovich, & Vaidya, 2004; Olson & Chun, 2001), and spatial locations (Chun & Jiang, 1998, 2003; Peterson & Kramer, 2001). The current study focuses on spatial context learning.

Contextual Cuing

In an earlier study, Chun and Jiang (1998) showed that when, searching for a *T* target among *L* distractors, participants found the target faster on occasionally repeated displays than on novel displays. Such improvement is a form of associative learning: Because the distractor layout is consistently associated with a given target location, the spatial layout becomes predictive of the target. No improvement was seen on repeated layouts that had random target locations (Chun & Jiang, 1998; Wolfe, Klempen, & Dahlen, 2000). Such learning is known as contextual cuing. This learning is acquired after merely five or six repetitions of a search display, and it has a high capacity, allowing at least 60 repeated displays to be learned (Jiang, Song, & Rigas, 2005). Moreover, it is long lasting: A memory trace for repeated displays persists for at least a week (Chun & Jiang, 2003; Jiang et al., 2005).

The present study investigates how humans represent a repeated spatial context. Do we represent it as an abstract spatial layout, independent of what exact shape or color occupies the layout during training? Alternatively, do we represent it as a specific pattern, with the trained item identity integrated into the layout? Consider a real-world spatial learning analogy. As a man repeatedly navigates through the city of Boston, does he extract the spatial layout of the city as if each landmark is a shapeless and colorless mass, or does he integrate specific landmark shapes with the layout?

According to the identity-independent account, spatial information is acquired independently of the shape or color of items that form a spatial layout. This account is consistent with the general division of labor in human vision: Spatial information can typically be represented independently of identity information (Jiang, Olson, & Chun, 2000; Mishkin, Ungerleider, & Macko, 1983). It also provides functional significance: One does not need to relearn the same spatial context when the items change color, luminance, or shape.

Conversely, according to the identity-contingent account, when we learn a repeated spatial layout we also learn that the layout is composed of, for example, black letter *L*s. Next time a display is presented, we try to match it with the memory trace laid down by the trained display (Logan, 1988). If that memory trace includes identity information, a difference in color or shape of items may lead the visual system to reject the current display as not matching

Yuhong Jiang and Joo-Hyun Song, Department of Psychology, Harvard University.

This study was supported by National Institutes of Health Grant MH071788 and Army Research Office Grant 46926-LS. We thank Mark Auckland, Matthew Peterson, Amanda Rigas, Rachel Sussman, and Jeremy Wolfe for comments and suggestions and Sidney Burks for data collection.

Correspondence concerning this article should be addressed to Yuhong Jiang, Department of Psychology, Harvard University, 33 Kirkland Street, WJH 820, Cambridge, MA 02138. E-mail: yuhong@wjh.harvard.edu

a trained display, even though the two share the same spatial layout. Although such an identity-contingent learning mechanism may result in misses, it has the advantage of reducing unnecessary memory comparisons. Given that an infinite number of memory traces exist in one's mind, most of which are irrelevant to an incoming visual display, reducing unnecessary comparisons on the basis of identity mismatching can be advantageous.

Despite the clear theoretical significance, few studies have examined the contingency of spatial context learning on identity information. Two studies are relevant, and they have provided strong support for the identity-independent account. Chun and Jiang (1998) first trained their participants to search for a *2* or a *5* target among ▭ shaped distractors. After participants had acquired contextual cuing, the distractors changed into new shapes (rotated *2*s and *5*s). An immediate transfer of learning was seen, suggesting that spatial context learning is independent of the identity of individual items. In an even more dramatic identity-change experiment, Nabeta, Ono, and Kawahara (2003) trained participants to learn a repeated visual layout and then tested them with a tactile search task. Nabeta et al. were amazed to find that participants again searched the previously learned spatial layout more quickly for the target, even though it was presented in a different sensory modality. These findings suggest that contextual memory for spatial information can be abstracted from the exact shape (or input modality) that occupies the layout.

In this study, we ask whether spatial context learning is always identity independent. We reason that, under some conditions, the visual system may take advantage of identity information to help deploy spatial attention. Consider the following analogy. Suppose every time one sees the city of Boston, it is sunny and bright, whereas every time one sees New York City, it is dull and drab. One's visual system may become sensitized to the luminance cue, such that next time it encounters a drab scene, it will match that scene with New York City but not with Boston. The advantage of such identity-contingent learning is that an incoming display only needs to be matched with a subset of one's memory, which potentially increases the speed of finding a match. The disadvantage, of course, is that one's spatial memory for a sunny Boston will not help one navigate the city in rain. Given that identity-contingent learning affords both advantages and disadvantages, it may only be revealed under certain conditions. This study tests conditions under which spatial context learning becomes identity contingent.

Experiment 1: Shape-Independent Spatial Context Learning

This experiment introduces the paradigm used in this study to investigate identity-contingent learning of spatial contexts. Participants took part in a training phase and a testing phase. Twenty-four unique visual search displays were repeated several times during training. After training, we either maintained or changed the trained item's identity while keeping the spatial layout the same. If spatial context learning is independent of item identity, it should transfer to same-layout displays regardless of the identity of items. Conversely, if spatial context learning is contingent on learned item identity, it should not transfer to same-layout displays composed of a new item identity.

In Experiment 1, participants searched for a *T* target among *L* distractors. There were two types of *L* distractors (*L*1 and *L*2), which had a subtle shape difference. Both *L*1 and *L*2 were composed of two line segments, but the offset at the junction of the segments was small (0.1°) for *L*1 and large (0.3°) for *L*2. This offset difference made *L*2 more similar to the target *T*, leading to more inefficient search for the target among *L*2 than among *L*1. Figure 1 shows two displays: one involving *L*1 and the other *L*2.

During the training phase, one group of participants was trained on 24 spatial configurations containing *L*1 distractors, whereas another group of participants was trained on 24 spatial configurations containing *L*2 distractors. During the testing phase, the configuration that previously contained *L*1 (or *L*2, for the second group of participants) might contain either *L*1 distractors or *L*2 distractors. New displays not trained before were also included as a baseline control condition. Of interest is whether the mismatch in the shape of distractors between the training and testing phases affects spatial contextual cuing.

Method

Participants. Twenty students from the Boston area participated for payment, 9 in Experiment 1A and 11 in Experiment 1B. They were 18 to 29 years old; all had normal or corrected-to-normal visual acuity.

Stimuli. Twelve items (1.3° × 1.3°) were presented at randomly chosen locations from an invisible 10 × 10 matrix (24° × 24°). One item was a target *T* rotated 90° to the left or to the right; the others were distractor *L*s rotated 0°, 90°, 180°, or 270°. The background was gray. Participants searched for the *T* target and pressed the left or the right key to report its direction.

Phases. Each participant completed a practice phase (2 blocks of 100–120 trials, on nonrepeated displays), a training phase (24 blocks of 24 trials), and a testing phase (1 block of 96 trials). The two versions of the experiment (1A and 1B) differed in the training phase.

Training. Each training block included 24 different displays. They contained *L*1 distractors in Experiment 1A and *L*2 distractors in Experiment 1B. *L*1 displays contained the target *T* plus 11 *L*s with a small offset at the junction (0.1°; Figure 1A), whereas *L*2 displays contained the target *T* plus 11 *L*s with a larger offset (0.3°; Figure 1B) at the junction. These displays were presented at randomly selected locations. All 24 displays were unique. They were presented again in Blocks 2 through 24, amounting to a total of 24 times.

Testing. Half of the 96 trials in the testing phase involved trained layouts, and the other half involved new layouts. Each previously trained *L*1 display (Experiment 1A) was presented twice in the testing phase, once containing *L*1 distractors again (*L1–L1 repetition*) and once containing *L*2 distractors (*L1–L2 repetition*). Similarly, each previously trained *L*2 display (Experiment 1B) was presented twice in the testing phase, once containing *L*2 distractors again (*L2–L2 repetition*) and once containing *L*1 distractors (*L2–L1 repetition*). Thus, each layout presented during training appeared twice in the testing phase: once with the same type of distractor, and once with the other type. The order of these two presentations in the testing phase was randomly determined. To assess contextual cuing, we compared the repeated displays with new displays that matched the repeated ones in target location and distractor shapes. That is, each repeated display had a yoked control with the same target location and *L* shapes, but the spatial layout was novel.

Trial sequences. Participants pressed the space bar to initiate each block. A fixation point (400 ms) was followed by the search display that was presented until participants made the response. Accuracy feedback immediately followed each response. The training and testing phases were run continuously without any special instructions in between. Participants

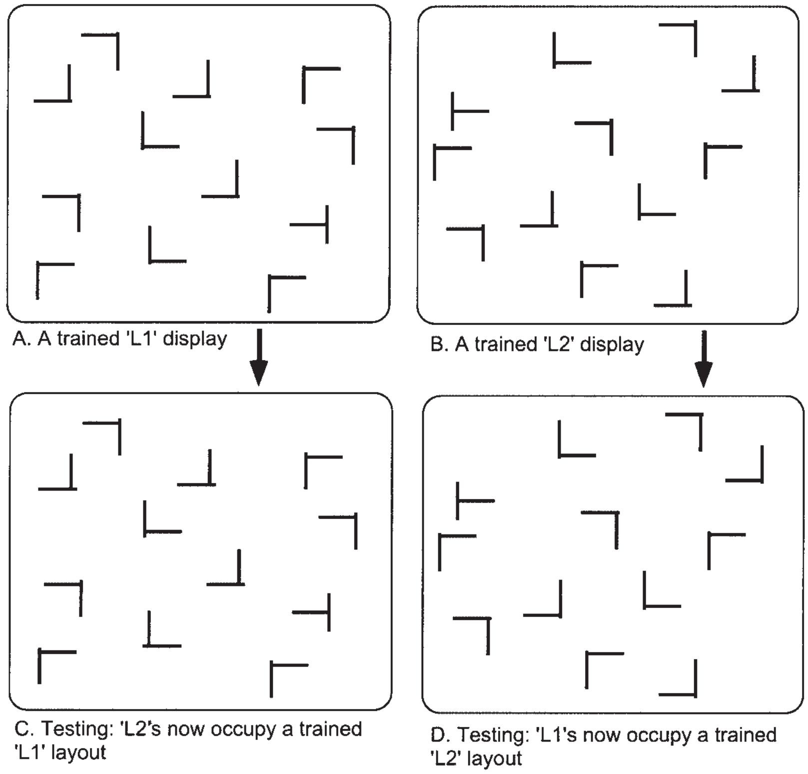

Figure 1. A schematic illustration of two different kinds of *L* distractors used in the experiment. A subtle difference in the offset of the *L*s differentiated *L*1 (Figure 1A) and *L*2 distractors (Figure 1B). After training, the shape of the *L*s might be the same as previously trained or different (as shown in Figures 1C and 1D). Two different groups of participants were trained on *L*1 and *L*2 displays.

were not informed of the nature of the design; they were simply instructed to conduct a visual search as accurately and as quickly as possible.

Results

The overall accuracy in this experiment (and all later experiments) was high (95% and above) and was not affected by any experimental factors (all ps $> .10$). We report only reaction time (RT) data from correct trials.

Training phase. Figure 2 shows mean RT in the training phase of Experiment 1. An analysis of variance (ANOVA) on distractor type (*L*1 or *L*2) as a between-subjects factor and training block as a within-subject factor showed a significant main effect of distractor type, with slower RTs for participants who searched through *L*2 distractors, $F(1, 18) = 19.22$, $p < .001$. The main effect of block was significant, $F(23, 414) = 3.82$, $p < .001$, showing improvement in speed as training progressed. The interaction between block and distractor type was also significant, $F(23, 414) = 2.54$, $p < .001$, and this was accounted for by a larger RT improvement for participants who received the harder task (*L*2 displays).

Testing phase. Figure 3 shows data from the testing phase of Experiment 1. The size of contextual cuing (new layout–trained layout) in different conditions is shown in Table 1. An ANOVA on layout repetition (repeated vs. new), testing distractor type (*L*1 vs. *L*2), and experiment (trained in *L*1 vs. *L*2, between-subjects) revealed a significant main effect of layout repetition, $F(1, 18) = 8.04$, $p < .011$, showing contextual cuing. There was also a significant main effect of testing distractor type, with slower RTs to *L*2 displays than to *L*1 displays, $F(1, 18) = 86.46$, $p < .001$. The main effect of experiment was not significant ($F < 1$), nor were any interaction effects (all Fs < 1.10, ps $> .30$). Thus, learning of

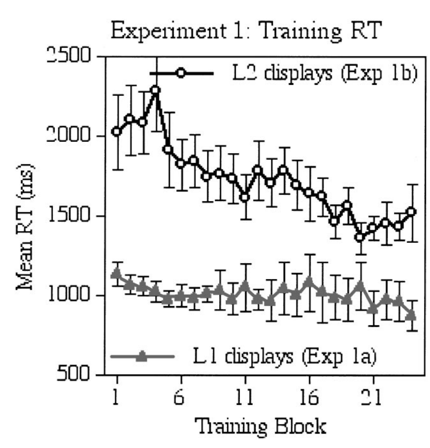

Figure 2. Experiment (Exp) 1: Training phase results. *L*1 and *L*2 displays differed in the shape of the *L* distractors. They were tested on two different groups of participants, each group trained on only one type of distractor. Error bars show standard error of the between-subjects variability. RT = reaction time.

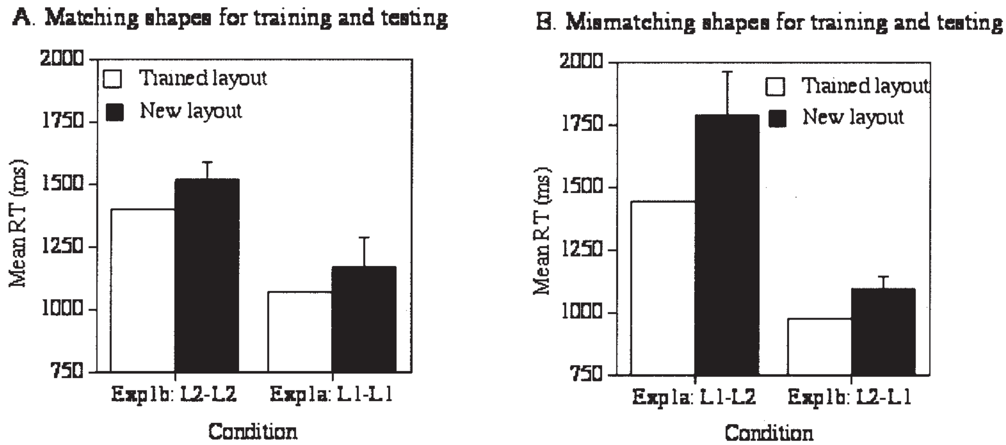

Figure 3. Experiment (Exp) 1: Results from the testing phase. A significant contextual cuing effect was observed, whether the distractor shapes in the testing phase and training phase matched (Figure 3A) or mismatched (Figure 3B). The error bars show standard error of the difference between trained layouts and new layouts. RT = reaction time.

the spatial layout transferred when distractors changed their shapes.

Discussion

Experiment 1 shows that when a visual search display was repeatedly presented, participants were able to extract the spatial layout independently of the item shapes. When the same layout was shown later, occupied by new distractor shapes, learning transferred to the new display. Such identity-independent learning was shown even though search difficulty changed considerably when distractor shapes changed. It is important to note that our training session has provided an opportunity to acquire identity-contingent learning: A spatial layout always contains one type of distractor shape during training. Yet the visual system had no difficulty in generalizing spatial context learning to another kind of distractor shape. In other words, after seeing many city scenes, all sunny and bright, one has no trouble generalizing spatial memory to drab scenes. These findings replicate previous studies that have found an insensitivity of spatial contextual cuing to the identity of individual items (e.g., Chun & Jiang, 1998).

Table 1
The Size of Contextual Cuing (New–Trained Layouts) in the Testing Session of Each Experiment

	Training and testing match (ms)		Training and testing mismatch (ms)	
Experiment	*L*1–*L*1	*L*2–*L*2	*L*1–*L*2	*L*2–*L*1
1 (trained in one shape)	110*	127*	351*	128*
2 (trained in two shapes)	122*	414*	43	5
3 (trained in one color)	185*		229*	
4 (trained in two colors)	213*		50	

* $p < .05$

Experiment 2: Shape-Contingent Spatial Context Learning

In Experiment 1 we found that spatial context learning was insensitive to the exact shape of items. Even though a spatial layout always contained *L*1 (or *L*2) distractors during training, contextual cuing transferred to displays that contained *L*2 (or *L*1) distractors. One interpretation of such results is that spatial context learning is always identity independent. Alternatively, participants might have ignored the identity information in Experiment 1 only because all trained displays contained the same type of distractors. Learning that a display contains *L*1 does not allow one to differentiate the current display from the other displays. What will happen when some spatial layouts contain *L*1 distractors and others contain *L*2 distractors during training? Under such conditions, if participants continue to ignore identity information, they will have to match each incoming display with all trained displays. If, instead, participants become sensitized to the identity information, they will only need to match an incoming display with some trained displays.

To find out whether spatial context learning is always independent of item shapes, we tested participants in a similar design as in Experiment 1, except that each participant was trained on both *L*1 and *L*2 displays. Half of the trials in the training phase contained *L*1 distractors, whereas the other half contained *L*2 distractors. These two types of trials were randomly intermixed. If spatial context learning is always identity independent, then learning should transfer to same-layout displays that contain new distractor shapes. Conversely, if spatial context learning becomes identity contingent, then learning should not transfer to same-layout displays that contain new distractor shapes.

Method

Participants. Fifteen new participants completed this experiment.

Design. This experiment was similar to Experiment 1, except that each participant was trained on both types of distractor shapes. In the training phase, each training block included 24 unique displays: Half were *L*1 displays, and the other half were *L*2 displays. These displays were randomly mixed in a block. All displays were presented for a second time in Block 2 and then again in Blocks 3 through 24. The type of object (*L*1 or *L*2) used for all repetitions of a display was held constant.

In the testing phase, half of the 96 trials involved trained layouts, and the other half involved new layouts. Each previously trained *L*1 display was presented twice in the testing phase, once containing *L*1 distractors (*L1–L1 repetition*), and once containing *L*2 distractors (*L1–L2 repetition*). Similarly, each previously trained *L*2 display was presented twice in the testing phase, once containing *L*2 distractors again (*L2–L2 repetition*), and once containing *L*1 distractors (*L2–L1 repetition*). Yoked controls involving the same target location and *L* shapes but new distractor layouts were also tested. All other aspects of the experiment were the same as in Experiment 1.

Results

Training. Figure 4 shows results from the training phase. An ANOVA on distractor shape (*L*1 vs. *L*2) and block (1–24) revealed a significant main effect of distractor shape, $F(1, 14) = 113.27$, $p < .001$, with much slower RTs for *L*2 displays. This was expected because *L*2 distractors are more similar to the target *T*, leading to more inefficient search (Duncan & Humphreys, 1989). The main effect of block was significant, $F(23, 322) = 2.55$, $p < .001$, showing decreased RTs as training progressed. Such reduction was statistically similar for *L*1 and *L*2 displays, reflected by the lack of an interaction effect between distractor shape and block ($F < 1$). Thus, search RT improved for both *L*1 and *L*2 displays as training progressed. Because all displays were repeatedly presented across blocks, the improvement reflected a combination of general learning of the visual search procedure and specific learning of the repeated displays.

Testing phase. We calculated mean RT separately for the four repeated conditions and their respective control conditions. Figure 5 shows the group mean. An ANOVA on layout repetition (repeated vs. new), testing distractor shape (*L*1 vs. *L*2), and shape matching (matched vs. mismatched) revealed a significant main effect of layout repetition, with faster RTs for trained layouts than for new layouts, $F(1, 14) = 7.86$, $p < .014$. Search RTs were faster for *L*1 than for *L*2 testing displays, $F(1, 14) = 31.92$, $p < .001$. The main effect of shape matching was not significant ($F < 1$). However, there was a significant interaction between layout repetition and shape matching, $F(1, 14) = 12.72$, $p < .003$. No other interaction effects were significant (all *p* values $> .15$). Planned contrast showed that layout repetition facilitated RT when the training and testing phases matched in distractor shapes, $F(1, 14) = 19.08$, $p < .001$, but layout repetition had no effect on RTs when the training and testing phases used different distractor shapes ($F < 1$).

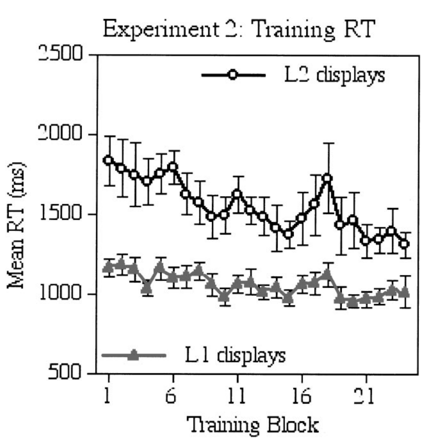

Figure 4. Mean reaction time (RT) for correct trials in the training session of Experiment 2. *L*1 and *L*2 displays differed in the shape of the distractors (see Figures 1A and 1B). They were randomly intermixed within a block. Error bars show standard error of the between-subjects variability.

Discussion

By training participants to conduct visual search from repeated spatial layouts and testing their spatial contextual cuing with displays that either matched or mismatched trained layouts in the shape of distractors, we have shown here, for the first time, evidence for the identity-contingent learning account. When participants were trained to conduct search from *L*1 distractors, spatial context learning did not transfer to displays that now contained *L*2 distractors, and vice versa. This is a surprising observation given that previous studies have found that spatial context learning can be abstracted from the exact item shapes, or even input modality, that occupied the layout (Chun & Jiang, 1998; Nabeta et al., 2003). What makes such results even more unexpected is that the two types of distractors—*L*1 and *L*2—were quite similar in shape. Even though this manipulation had a large effect on search RT, the shape difference between *L*1 and *L*2 was subtle and not immediately obvious when participants were busy searching for the target. The contingency of spatial context learning on such subtle shape differences suggests that contextual cuing can become hyperspecific to the identity of distractor items.

Results from Experiment 2 can be contrasted with those of Experiment 1, in which participants transferred their learning of spatial context to displays involving different distractor shapes. To us, the most important difference between the two experiments lies in the training phase. Whereas participants were trained only on *L*1 or *L*2 displays in Experiment 1, they were trained on both *L*1 and *L*2 displays in Experiment 2. Strictly speaking, the identity of items was irrelevant to spatial layout learning in both experiments: Whether a display contained *L*1 or *L*2 distractors, there was always a consistent association between the distractor layout and the target's location. Still, participants acquired identity-contingent learning when trained on both kinds of distractors but not when trained on only one kind of distractor.

These results suggest that acquiring the contingency between a distractor shape and a spatial layout is advantageous to mixed training: Relying on distractor shape information reliably narrows the matching space from 24 possible target locations to 12. Even though participants eventually have to rely on the spatial layout information to cue attention to the exact target's location, identity matching has already simplified that process. These findings suggest that the visual system is sensitive to the trial context within which a repeated display is presented and makes an intelligent decision to retain identity information when it is useful.

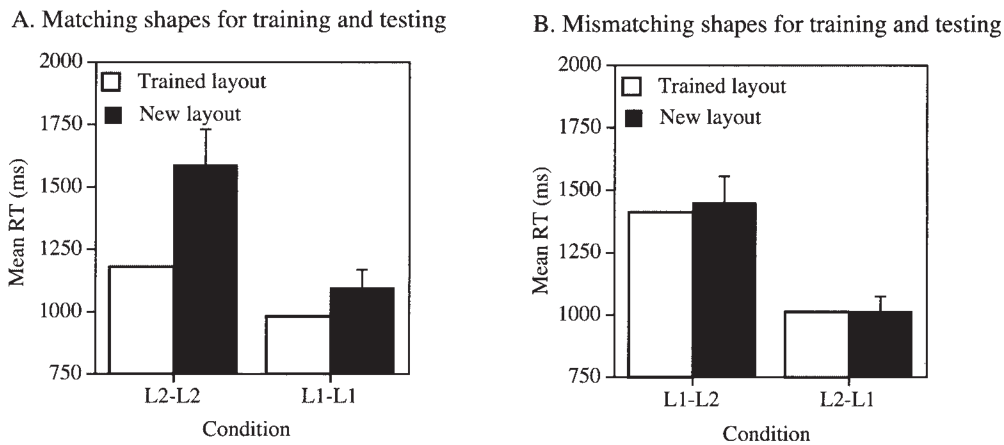

Figure 5. Experiment 2: Results from the testing phase. When the distractor shapes in the testing phase matched those during training, a significant contextual cuing effect was observed (Figure 5A). However, when the distractor shapes in training and testing mismatched, no contextual cuing was observed (Figure 5B). The error bars show standard error of the difference between trained layouts and new layouts. RT = reaction time.

Experiment 3: Color-Independent Spatial Context Learning

The previous two experiments suggest that spatial context learning can become identity-contingent under some conditions. In particular, mixing two types of distractor shapes during training sensitizes participants to identity information, leading to shape-contingent spatial learning. These results support the *sensitization* hypothesis. According to this hypothesis, when two types of distractors are trained and are consistently associated with certain repeated layouts, the visual system becomes sensitized to their identity difference, and learning of the spatial layout becomes contingent on identity. This hypothesis should hold whether the two types of distractors differ in shape or in other respects. The shape difference tested in Experiments 1–2 is special in that a subtle perceptual difference in the *L* shapes dramatically changes search difficulty. In Experiments 3–4, we wish to generalize the sensitization hypothesis to cases when the two distractor types differ in color. Whether a display contains black or white items is perceptually conspicuous, but this does not affect the difficulty of visual search. A replication of results found in Experiments 1–2 under such conditions will further strengthen the sensitization hypothesis.

In Experiment 3, participants were trained on displays with white items only (or black items only). During the testing session, half of the displays contained white items, the other half black items. The sensitization hypothesis predicts that because only one color is used for all trained displays, the visual system will disregard item colors. In turn, learning should transfer to same-layout displays that now contain a new distractor color. In Experiment 4, participants were trained on white items on half of the trials and black items on the other half. Because half of the trained displays were in one color and the other half in another, the sensitization hypothesis predicted that spatial context learning would become contingent on color. In turn, learning should not transfer to same-layout displays that now contain distractors in a different color.

Method

Participants. Ten new observers were tested in Experiment 3.

Training phase. Participants were trained in 24 blocks of 24 trials; all trials included white items for half of the participants and black items for the other half. Items were presented against a gray background throughout the experiment. All trials contained a target *T* and 11 *L*2 distractors. The target and distractors were always in the same color. The trials were presented once per block for 24 times.

Testing phase. There were four blocks, each including 96 trials. Each of the trained displays was presented twice in the testing phase, once in the trained color and once in the other color. These trials were contrasted with control trials involving a new spatial layout. To increase statistical power, we repeated the testing block of all 96 trials three times, which produced four testing blocks of identical trials. All other aspects of this experiment were the same as in Experiment 1.

Results and Discussion

Training phase. Because the exact color—black or white—of a given display did not affect RT, we collapsed data from participants who were trained in white and those who were trained in black displays. RTs improved from 1,573 ms in Block 1 to 1,363 ms in Block 24. The main effect of block was significant, $F(23, 207) = 2.17$, $p < .002$.

Testing phase. We collapsed data from all testing blocks (as they produced similar patterns of data) and both colors to produce *color match* and *color mismatch* trials. Figure 6 shows the group mean. An ANOVA on layout repetition (repeated vs. new layouts) and color matching (match vs. mismatch) revealed a significant main effect of layout repetition, $F(1, 9) = 5.26$, $p < .05$, showing a contextual cuing effect. The main effect of color matching was not significant, $F(1, 9) < 1$, nor was the interaction between training and color matching significant, $F(1, 9) = 1.01$, $p > .30$. Contextual cuing was significant even when the training and transfer colors did not match, $t(9) = 2.73$, $p < .023$. This finding is analogous to that of Experiment 1. Both suggest that spatial

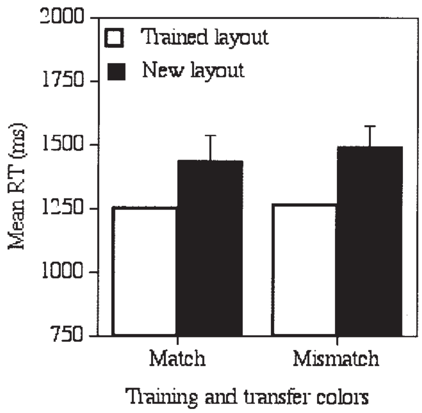

Figure 6. Experiment 3: Results from the testing phase. Contextual cuing was not color specific when participants were trained in only one color and tested in two colors. RT = reaction time.

context learning is independent of the identity (shape or color) of trained distractor items.

Experiment 4: Color-Contingent Spatial Context Learning

This experiment was similar to Experiment 3, except that half of the trained displays contained white items and the other half black items.

Method

Participants. Ten new participants were tested in this experiment.

Training phase. Participants took part in 24 blocks of training, including 24 trials each. All trials contained a target *T* and 11 *L*2 distractors. The target and distractors were always in the same color. Half of the displays contained white items and the other half black items. A given spatial layout was always presented in a given color in the training phase. Black and white displays were randomly intermixed within a single block.

Testing phase. Participants completed four testing blocks, with 96 trials each. These included 48 trials of repeated layout, evenly and randomly divided into *white–white repetition*, *black–black repetition*, *white–black repetition*, and *black–white repetition*. The other trials contained new distractor layouts, matched with the repeated displays in color and in target location. To increase statistical power, we repeated all 96 trials—including the new trials—once in each of Testing Blocks 2–4. All other aspects of the experiment were the same as in Experiment 1.

Results and Discussion

Training phase. Because RT for black displays was comparable to RT for white displays, we averaged across the two colors. Training RT improved from 2,001 ms in Block 1 to 1,560 ms in Block 24. The main effect of block was significant, $F(23, 207) = 2.24$, $p < .002$, showing a significant RT improvement.

Testing phase. Because color (black or white) had no effect on RT, we pooled across white–white and black–black repetitions and across white–black and black–white repetitions to produce color-matching and color-mismatching conditions. In addition, results from the four testing blocks showed a statistically similar pattern, so we averaged across all four blocks. Figure 7 shows results from the testing phase.

An ANOVA on layout repetition (repeated vs. new layouts) and color matching (match vs. mismatch) revealed a significant main effect of training, $F(1, 9) = 8.14$, $p < .019$, showing a contextual cuing effect. The main effect of color matching was significant, $F(1, 9) = 12.10$, $p < .007$, as was the interaction between layout repetition and color matching, $F(1, 9) = 6.17$, $p < .035$.

In particular, when items matched their colors between training and testing, there was a significant contextual cuing effect, $t(9) = 5.85$, $p < .001$. However, when the colors did not match, no contextual cuing was observed, $t(9) = 0.70$, $p > .45$. Thus, spatial contextual memory can become contingent on item color. This is consistent with results observed in Experiment 1.

Experiments 3 versus 4. To directly compare results from Experiments 3 and 4, we conducted an ANOVA using experiment as a between-subjects factor and layout repetition (repeated vs. new layouts) and color matching (match vs. mismatch) as within-subject factors. This analysis revealed a significant three-way interaction, $F(1, 18) = 6.88$, $p < .017$. This confirmed that although contextual cuing was color specific in Experiment 4, it did not show such specificity in Experiment 3, consistent with the sensitization hypothesis.

General Discussion

When a repeated visual search display is presented, do we represent an abstract spatial layout independent of item identity, or do we integrate identity information with the spatial layout? The identity-independent account postulates that, after learning a repeated display, we can transfer learning to the same layout containing a new item identity (Chun & Jiang, 1998). Representing the abstract spatial layout independently of item identity saves us the trouble of relearning the same display again. In contrast, the identity-contingent account predicts that spatial context learning is specific to the shape or color of the trained items, so learning should not transfer when a new display matches the trained one in layout but mismatches it in item color or shape. Such representation simplifies the memory matching process, such that an incoming display need not be compared with a vast array of memory displays.

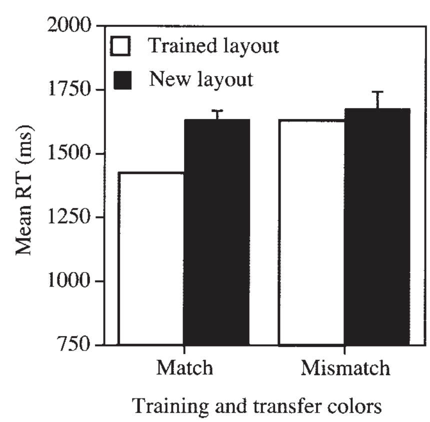

Figure 7. Experiment 4: Results from the testing phase. Participants were trained on black displays and white displays. Contextual cuing was color specific, showing up only when the training and testing colors matched. RT = reaction time.

The present study suggests that under some conditions spatial context learning is identity independent, whereas under other conditions it is identity contingent. Experiments 1 and 3 show that when participants are trained on only one distractor type, learning of repeated layouts is independent of the item identity. In contrast, Experiments 2 and 4 show that when participants are trained on one distractor type on some layouts and on another distractor type on other layouts, they become sensitized to distractor identity. Learning of a layout becomes contingent on the identity of the items, so no transfer is seen when the old layout contains new shapes or colors. These observations are consistent with the sensitization hypothesis. They suggest that the visual system is sensitive to the trial context within which a repeated display is presented and that it makes an intelligent decision by retaining useful statistical regularity.

Sensitivity to Trial Context

Laboratory tasks are often arbitrarily set up to contain discrete trial events, with no relation between one trial and the next. Yet, in the real world, visual events rarely appear in temporal isolation; they are always embedded in the preceding and following events. Driving requires the constant updating of information gathered from many instances, crossing a busy street requires memory for what is on one's left and right, and even the simplest social interactions depend on retaining information across space and time. Therefore, perhaps it is not surprising that the visual system is quite sensitive to frequency information (Bacon & Egeth, 1994; Hasher & Zacks, 1984) and intertrial associations.

One example of intertrial association is the negative priming effect (Tipper, 1985), in which RT to a target is slower if the target was a distractor in a preceding trial. In spatial context learning, if the spatial layout presented on one trial is predictive of the target's location on the next trial, RT to the next trial is facilitated, showing intertrial contextual cuing (Ono, Jiang, & Kawahara, 2005).

Sensitivity to trial context is clearly revealed in the current study. Representation of a given spatial display can either be abstract or be contingent on item identity, depending on what other trials within the block contain. Although it is still in its early stages of research, such a trial-context effect can have far-reaching influence on visual implicit learning. For example, suppose most of the trials in a block contain a consistent association between distractor shape and target shape. This biases the visual system to focus on identity and not spatial information. It may affect what can later be learned, when spatial information becomes predictive (Endo & Takeda, 2004). Such a biasing effect is analogous to the perception of ambiguous figures: When one has seen a clear image of a young woman, an ambiguous figure that can be interpreted either as a young woman or as an old woman often leads to a percept of a young woman.

Researchers can change trial context in many different ways—for example, by changing the proportion of repeated and nonrepeated trials, the kinds of associations on other trials, or the color and shape of items on other trials. Such variability makes it difficult to specify an overarching theory of trial context effects. Future studies that examine the effects of preceding and subsequent trials on a given visual event will significantly enhance our understanding of visual implicit learning.

Visual Contextual Learning Is Intelligent

In hindsight, it makes sense that spatial context learning becomes identity contingent when two types of distractors are used during training. Sensitivity to distractor identity allows the visual system to narrow down its necessary matching space: An incoming display only needs to be compared with half of the previously learned displays. It also makes sense that spatial context learning is identity independent when only one type of distractor is trained: The only information that is predictive of the target's location is the spatial layout, whereas the identity of distractors is irrelevant. In this sense, the visual system is quite smart: It keeps identity information when that is useful during training and discards identity information when it is uninformative. We did not explicitly predict such results to hold true prior to collecting data, even though our implicit visual learning system showed such a pattern of results: Our implicit visual learning system was smarter than our explicit theorizing.

By ascribing intelligence to the spatial context learning system, we certainly do not mean that explicit top-down knowledge and conscious decisions determine what information is retained. As Pylyshyn (1999) has nicely summarized, part of the visual system—what he called *early vision*[1]—is modular. It is informationally encapsulated from an individual's background knowledge. Yet early vision can act intelligently because of intrinsic constraints or built-in assumptions. Along similar veins, we believe that the intelligence revealed in this study is an intrinsic property of the visual statistical learning mechanism: It operates by maximizing statistical predictability. One way to achieve maximal predictability is to remain sensitive to trial context to extract what is constant or predictive across multiple trials.

Relation With Animal Learning Mechanisms

As a new paradigm, contextual cuing has ignited new research interest in human visual learning. Contextual cuing is mainly based on associative learning between one stimulus (the spatial layout) and another (the target location), which also has been an important mechanism in traditional animal learning research. Thus, there are significant parallels between contextual cuing and animal learning studies.

For example, Endo and Takeda (2004) observed overshadowing (Kamin, 1969; Pavlov, 1927/1960), in which a stronger stimulus dominates learning when both a weak and a strong stimulus are associated with the same target information. Endo and Takeda found that people tended to learn distractors' location rather than their shape in contextual cuing because location information is more salient in visual search. Another parallel between contextual cuing and animal learning studies is reflected in latent learning, proposed initially by Tolman and Honzik (1930). In contextual cuing, ignored spatial context can result in latent learning (Jiang & Leung, 2005).

Furthermore, mechanisms that affect discriminative learning and stimulus generalization in animals also apply to contextual

[1] The word *early* in *early vision* is interpreted here at a functional level; it does not necessarily map onto early visual areas in the brain. The correspondence between function and anatomy for early vision and late vision is not entirely straightforward (see, e.g., Hochstein & Ahissar's, 2002, reverse hierarchy theory).

cuing, an observation that our current study addresses. *Stimulus generalization* refers to the observation that a learned association between a conditioned stimulus and a conditioned response will generalize to another stimulus that resembles the conditioned stimulus (Watson & Rayner, 1920). The more similar the new stimulus is to the learned conditioned stimulus, the greater is the generalization response. In addition, whether generalization occurs depends partly on the initial learning procedure. Lashley and Wade (1946) proposed that because participants do not learn all aspects of their environment, learning relies on selective attention. If a pigeon is trained to peck a red light, it will later generalize to peck a yellow light. But if the pigeon is trained to peck only a red light and not a green light, then learning will not generalize to a yellow light. The current study is consistent with these findings, although the kind of association acquired in our study is much subtler than the kind commonly acquired in animal learning studies, which should increase the difficulty of discriminative learning.

The animal learning literature can also help to answer an important question: Will the training procedure in Experiments 1 and 3, in which a single distractor shape or color was used, always generalize to displays with the same spatial layout?[2] Given that the current study did not exhaust all possible displays with the trained layout, it cannot directly answer this question. However, studies on animals' discriminative learning have suggested that generalization is a graded function of the similarity between the new stimulus and the learned stimulus, even when animals were trained with a single stimulus (Riley & Leuin, 1971; Switalski, Lyons, & Thomas, 1966). When the transfer display is very different from the trained display, the generalization of learning may be reduced in strength (see also Jiang & Song, in press).

Conclusion

In summary, this study has shown, for the first time, evidence that spatial context learning can become identity contingent. We have also clarified one factor that determines whether spatial context learning is identity independent or identity contingent. In particular, when the training session includes multiple types of distractor displays, the visual system becomes sensitized to the distractor identity, which leads to identity-contingent representation of the spatial layout. Conversely, when the training session includes a single type of distractor identity, an abstract spatial layout is extracted and retained. We note that many other factors can potentially influence the contingency of spatial context learning on item identity. Although we did not intend to provide an exhaustive list of these factors, we believe that future studies focusing on intertrial context effects will generate fruitful results.

[2] We thank Jeremy Wolfe for raising this question.

References

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12,* 157–162.

Atchley, P., & Andersen, G. (1995). Discrimination of speed distributions: Sensitivity to statistical properties. *Vision Research, 35,* 3131–3144.

Bacon, W. F., & Egeth, H. W. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics, 55,* 485–496.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43,* 393–404.

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

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science, 10,* 360–365.

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

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Endo, N., & Takeda, Y. (2004). Selective learning of spatial configuration and object identity in visual search. *Perception & Psychophysics, 66,* 293–302.

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

Hasher, L., & Zacks, R. T. (1984). The automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist, 39,* 1372–1388.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron, 36,* 791–804.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 113–136.

Howard, J. H., Howard, D. V., Dennis, N. A., Yankovich, H., & Vaidya, C. J. (2004). Implicit spatial contextual learning in healthy aging. *Neuropsychology, 18,* 124–134.

Jiang, Y., & Leung, A. W. (2005). Implicit learning of ignored visual context. *Psychonomic Bulletin & Review, 12,* 100–106.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 683–702.

Jiang, Y., & Song, J.-H. (in press). Spatial context learning in visual search and change detection. *Perception & Psychophysics.*

Jiang, Y., Song, J.-H., & Rigas, R. (2005). High-capacity spatial contextual memory. *Psychonomic Bulletin & Review, 12,* 524–529.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive control* (pp. 279–296). New York: Appleton-Century-Crofts.

Lashley, K. S., & Wade, M. (1946). The Pavlovian theory of generalization. *Psychological Review, 53,* 72–87.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences, 6,* 414–417.

Nabeta, T., Ono, F., & Kawahara, J. (2003). Transfer of spatial context from visual to haptic search. *Perception, 32,* 1351–1358.

Olson, I. R., & Chun, M. M. (2001). Temporal contextual cuing of visual attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1299–1313.

Ono, F., Jiang, Y., & Kawahara, J. (2005). Intertrial contextual cuing: Association across successive visual search trials guides spatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 703–712.

Pavlov, I. P. (1960). *Conditioned reflexes* (G. V. Anrep, Trans.). New York: Dover. (Original work published 1927)

Peterson, M. S., & Kramer, A. F. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics, 63,* 1239–1249.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology, 2,* 509–522.

Pylyshyn, Z. W. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22,* 341–423.

Riley, D. A., & Leuin, T. C. (1971). Stimulus-generalization gradients in chickens reared in monochromatic light and tested with single wavelength value. *Journal of Comparative and Physiological Psychology, 75,* 399–402.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real world interaction. *Psychonomic Bulletin & Review, 5,* 644–649.

Switalski, R. W., Lyons, J., & Thomas, D. R. (1966). Effects of interdimensional training on stimulus generalization. *Journal of Experimental Psychology, 72,* 661–666.

Thorpe, S., Fixe, D., & Marlot, C. (1996, June 6). Speed of processing in the human visual system. *Nature, 381,* 520–522.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 37*(A), 571–590.

Tolman, E. C., & Honzik, C. H. (1930). Introduction and removal of reward, and maze performance in rats. *University of California Publications in Psychology, 4,* 257–275.

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1–14.

Wolfe, J. M., Klempen, N., & Dahlen, K. (2000). Postattentive vision. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 693–716.

Received April 21, 2004
Revision received September 3, 2004
Accepted May 24, 2005 ■