

## **NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS**

The copyright law of the United States [Title 17, United States Code] governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that use may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgement, fulfillment of the order would involve violation of copyright law. No further reproduction and distribution of this copy is permitted by transmission or any other means.

Lehigh University ILL



ILLiad TN: 405540

**Borrower:** PMC

**Lending String:**

\*LYU,OBE,TFW,OLP,TEU,LS1,CTW,YQU,MWR,C  
UY,VGM,LRU,VRC,TJC,OKS

**Patron:**

**Journal Title:** Dynamic cognitive processes /

**Volume:** Issue:

**Month/Year:** Pages: 11-35

**Article Author:**

**Article Title:**

**ILL Number:** 166800611



**Call #:** 153 D997

**Location:** FM 2nd Floor North

This Material may be Protected by Copyright  
Law (Title 17 U.S. Code)

**Charge**

**Maxcost:** 85.00IFM

**Shipping Address:**

Carnegie Mellon University-Hunt Library-ILL  
Office

4909 Frew St.

Hunt Library Room 305

Pittsburgh, Pennsylvania 15213

United States

**Fax:** 412-268-6945

**Odyssey:** 128.2.20.146

**Email:** hl82@andrew.cmu.edu

**Borrower:**

**Call #:**

# Acquisition of Long-Term Visual Representations: Psychological and Neural Mechanisms

Marlene Behrmann, Joy Geng, and Chris Baker

Carnegie Mellon University, Pittsburgh, USA

**Summary.** How do we so rapidly achieve an organized, coherent visual percept of our superficially chaotic world? One way of reducing the complexity of the input is to take advantage of the statistical regularities and regular co-occurrences between aspects of objects and between objects and their spatial locations. In this chapter, converging data obtained from normal and brain-damaged individuals, as well as from single unit recording studies in monkeys, are presented, all of which address the psychological and neural mechanisms associated with statistical learning. The first section deals with learning regularities associated with particular spatial locations, presumably a function of the dorsal 'where' stream and data from normal individuals and from patients with hemispatial neglect are presented. The second section reports the findings from human and monkey studies, which show how statistical contingencies of the visual environment are reflected in behavior and how neurons in monkey inferotemporal cortex, the ventral "what" stream, appear to mediate these statistical effects. Taken together, using data from a variety of methodologies, this work attests to the flexibility and robustness of the visual system and sheds light on the way in which perceptual organization operates to convert raw input into long-term visual representations.

**Key words.** Vision, perceptual organization, visual learning, neuropsychology, agnosia, neurophysiology

## Introduction

It is well established that human observers can learn statistically or probabilistically defined patterns (e.g., frequent co-occurrence) of both auditory stimuli (Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999) and visual stimuli (Chun, 2002; Edelman, Hiles, Yang, &

Intrator, 2002; Fiser & Aslin, 2001, 2002a). It is also the case that humans are sensitive to spatial regularities in the environment and are able to exploit the statistical contingencies that determine the location of a visual target (Chun & Jiang, 1998; Hoffman & Kunde, 1999; Lewicki, Czyzewska, & Hoffman, 1987; Maljkovic & Nakayama, 1996; Mayr, 1996). Such learning, often called "statistical learning," is commonly described as incidental or implicit, in that learning of new representations can occur automatically without instruction and without observers explicitly attending to and encoding the patterns. For example, statistical learning has been demonstrated both when the stimuli are presented passively to observers in the absence of any explicit task (Fiser & Aslin, 2001), and when observers are attending to and performing a separate, unrelated task (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). Indeed, even if participants demonstrate no explicit awareness of the relation between targets and distractor context, they nevertheless respond faster to a visual target that appears in a repeated distractor configuration compared to one that appears in a novel configuration (Chun & Jiang, 1998).

Statistical learning is so ubiquitous that it has also been observed in human infants 9 months old or younger (Fiser & Aslin, 2002b; Kirkham, Slemmer, & Johnson, 2002; Saffran et al., 1996) as well as in naïve non-human primates (Baker, Behrmann, & Olson, 2002; Hauser, Newport, & Aslin, 2001). For example, one might think of the A-not-B error exhibited by infants—looking in the location where the previous stimulus appeared rather than in the location of the current stimulus—as an inability to inhibit the most probabilistic response associated with reward. But Smith, Thelen, Titzer, and McLin (1999) have argued that this type of error is caused by a directional bias in motor planning due in part to the history of looking and reaching to the A location during the preceding trial(s). Because infants have immature control systems, a brief visual input signaling the current B location is too weak to overcome the motor bias. However, if the visual stimulus at B is salient, it can pull the motor response toward that location (Smith et al., 1999). This suggests that whereas mechanisms supporting simple matching behaviors may be modulated by spatial orientation, they may also operate independently.

Indeed, the ability to track statistical probabilities linking behavior to reward appears to be widespread in animal species. The *matching law* characterizes the absolute rate of response as a linear function of the frequency of reinforcement. In one of the earliest examples, Herrnstein (1961) demonstrated that the frequency with which pigeons pecked at each of two response keys was commensurate with the reinforcement schedule at each key. Although different reinforcement paradigms result in under- or over-matching, the ranking of responses in correspondence with the

available reinforcement hierarchy is well established (Baum, 1979; Greggers & Mauelshagen, 1997).

The central question is what psychological and neural mechanisms mediate statistical learning and, specifically, for the current purposes, visual statistical learning? This chapter examines this issue by reviewing a series of recent studies we have done addressing both learning of spatial regularities and learning of shape regularities in the visual modality. Each of the two sets of studies tracks the acquisition of these regularities in one of the two visual cortical streams, with the former set associated with parietal cortex and the latter with temporal cortex. In the course of conducting these studies, we have exploited a range of methodologies, including psychophysical studies with normal participants, behavioral investigations with individuals who have sustained brain damage, and single unit recording studies in awake, behaving monkeys. Through this convergence of methods, our goal is to elucidate the processes whereby visual representations are acquired as a function of the statistics of the input.

## **Learning and acquisition of spatial regularities**

As alluded to previously, studies with humans and animals have shown that the sensitivity to the spatial location of a stimulus is contingent on the regularities in the input; for example, people respond faster to targets that appeared in the same distractor configuration compared to those that appear in novel configurations (Chun & Jiang, 1998). In fact, people appear to be sensitive to repetitions in target location over approximately 5-8 intervening trials even when there is no probability manipulation (Maljkovic & Nakayama, 1996). These findings are consistent with prior data suggesting that probabilistic distributions in target location are related to performance optimization (Shaw & Shaw, 1977).

Taken together, the results from human and nonhuman species implicate an evolutionarily primitive mechanism that is sensitive to environmental regularities that result in behavioral success. To explore this behavioral sensitivity further, we have conducted studies examining whether adult human participants match their behaviors to implicit regularities in target location during a visual search task (Geng & Behrmann, 2002). In particular, we (Joy Geng and Marlene Behrmann) were interested in investigating whether, in a visual search task, target discrimination is facilitated when targets are more likely to appear in locations on one half of the display compared to the other half of the display. This issue was investigated both in normal individuals and in patients with hemispatial neglect.

T			E		
	F				T
		T		E	
C1	C2	C3	C4	C5	C6

Fig. 1. Schematic depiction of the visual search display used to investigate facilitation of spatial regularities on target discrimination. Note that the grid was not visible to the participants; it is shown here for purposes of illustration. Six columns and three rows were used. Six stimuli appeared on each trial, a target and five distractors. The location of the target was systematically manipulated to appear with high probability in the same location on the uneven condition trials.

### Visual search and spatial regularities

To investigate the facilitation of target discrimination as a function of statistical frequency, we used a visual display containing 6 items, one target (either F or L) and five distractors (T or E). There were 18 possible locations, formed by a grid of 6 vertical columns by 3 horizontal rows (see Figure 1), although this grid was not visible to the participant. Six letters appeared on each trial, one in each column. Participants responded by pressing one button for "L" targets and another for "F" targets. The probability manipulation was implemented across two conditions that appeared sequentially in separate blocks of trials. In the *baseline* condition, targets were equally likely to appear in any of the six columns. In the *uneven* condition, targets appeared on one half of the screen (e.g., columns 1-3) with 80% probability and on the other half (e.g., columns 4-6) with 20% probability. The target was equally likely to appear in all possible locations within the selected screen side. The screen side containing 80% of the targets was counterbalanced (left or right). No mention of the prob-

ability manipulation was made at the beginning of the experiment; participants were simply instructed to indicate which target was present as quickly and accurately as possible.

To ensure that subjects maintained central fixation, prior to the onset of the visual search display they reported a digit from 1-9 that appeared centered over fixation. To gauge their awareness of the probability manipulation, at the end of the experiment participants were asked, "Did you feel that the target was more likely to appear in one location or region, or did you feel that targets were well distributed?" and the response was recorded.

We tested two groups of normal subjects, a young group consisting of college students, to ascertain the extent of the probability effect, and then a group of elderly subjects (to serve as controls for the patients with hemispatial neglect, see below). All participants were right handed and had normal or corrected-to-normal vision. Subjects made very few errors (1% and 2.5% for the young and elderly subjects) and so the analyses were performed on reaction time (RT). Because the side of space on which the probability was manipulated did not influence performance, the location factor was collapsed for the analysis. Also, because both young and elderly subjects were influenced by the probability manipulation to an equivalent extent, the age factor was collapsed for the analysis. The critical finding was that significant differences between the probability conditions were observed in all columns except 5, which showed the same numerical trend (see Figure 2). However, because the two probability conditions were always run in sequence, with the baseline first so as to prevent contamination from the altered probability distribution, it was difficult to determine whether decreases in RT were due to general practice effects and/or to the probability manipulation. A comparison of the RT difference between columns *within* each probability condition provided further answers. Pairwise comparisons revealed the following: In the *baseline* condition, none of the column pairs differed significantly from each other. In the *uneven* condition, however, columns 1 and 2 were significantly different from columns 6 and 5, respectively. This comparison confirmed that targets in the left-most columns were detected more quickly than targets in the right-most columns but only when the statistics governing target location were biased towards the left side of the screen.

Twelve of the sixteen undergraduate participants and all of the elderly participants reported having no awareness of the probability distribution at the end of the experiment; re-running the analysis excluding the four who reported some awareness did not alter the findings. That participants report no obvious tracking of the probability suggests that the results are not due to explicit anticipation of target location at the beginning of each trial but

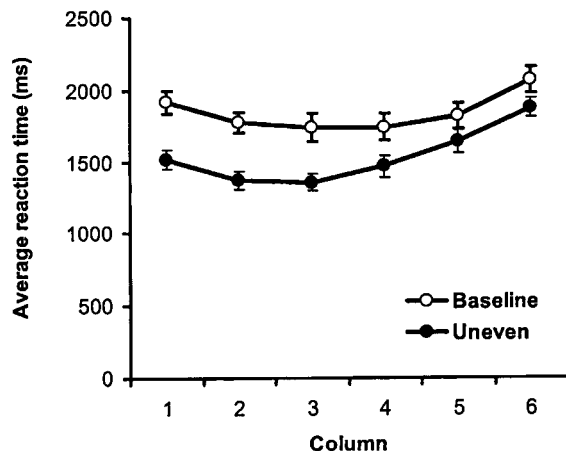


Fig. 2. Mean reaction time as a function of screen column for control participants for baseline and uneven conditions. Note that the data from the young and elderly control participants are collapsed here. Error bars indicate standard error.

rather that the spatial contingencies are being coded implicitly. From these data, we conclude that young and elderly normal participants are sensitive to the probability distribution of target objects, even when the distribution is over a region that includes several locations and when there is no explicit awareness of the contingency. We take these results to be a clear demonstration of how a flexible and adaptive orienting system may direct attention optimally in response to statistical contingencies in the visual field in normal individuals.

To explore further the consequences of the probabilistic distribution of targets on spatial representation, we conducted the same experiment in individuals with hemispatial neglect. Hemispatial neglect (or neglect, for short) is a deficit in representing contralesional space, acquired following a brain lesion. Individuals with this disorder fail to notice or report information on the side of space opposite the lesion, despite intact sensory and motor processes (Bartolomeo & Chokron, 2001; Bisiach & Vallar, 2000). Thus, for example, patients with a right hemisphere lesion fail to copy features on the left of a display despite incorporating the corresponding features on the ipsilesional right. The same individual may eat from only the right side of their plate or dress only the right side of their body. The deficit may affect all sensory modalities, including contralateral visual, auditory, somatosensory, and olfactory inputs. The presence of neglect may also adversely affect manual and oculomotor behavior in that these patients

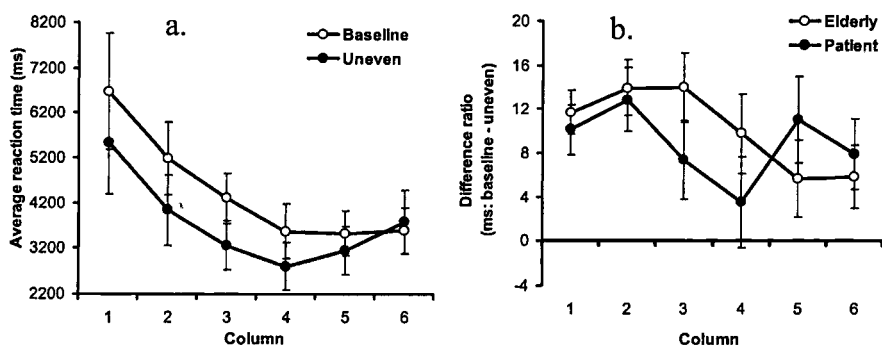


often are impaired at directing their eyes and/or hand to the contralateral side, even in the absence of visual input (Gore, Rodriguez, & Baylis, 2001/2002; Hornak, 1992; Mattingley, Husain, Rorden, Kennard, & Driver, 1998). Finally, neglect can even affect the contralateral side of an internal representation in the absence of sensory input, and can be reflected in mental imagery, as so elegantly demonstrated in the seminal work by Bisiach and Luzzatti (1978).

Many studies have demonstrated a significant impairment in the visual search abilities of neglect patients (Aglioti, Smania, Barbieri, & Corbetta, 1997; Behrmann, Ebert, & Black, 2004; Eglin, Robertson, & Knight, 1989; Esterman, McGlinchey-Berroth, & Milberg, 2000; Riddoch & Humphreys, 1987). There is evidence, however, that patients with neglect are able to exploit explicit spatial cues such as arrows indicating the target location, or verbal instructions to orient leftwards (Halligan, Manning, & Marshall, 1991; Lin, Cermak, Kinsbourne, & Trombly, 1996; Riddoch & Humphreys, 1983). The critical question is whether these individuals with neglect are able to exploit spatial contingencies in the same way as normal participants. To assess this, we examined the impact of statistical regularities of target location as an orienting cue in relation to the spatial gradient of neglect behavior with seven individuals, all of whom had sustained a lesion to the right hemisphere and all of whom exhibited left-sided visual neglect on the Behavioral Inattention test (Wilson, Cockburn, & Halligan, 1987).

As is evident in Figure 3a, the critical finding was the presence of a significant interaction between condition (*baseline*, *uneven*) and column of target. Pairwise comparisons of corresponding columns in the two probability conditions revealed a significant reduction in RT in columns 1-3 in the *uneven* compared with the *baseline* condition but no significant change in columns 4-6. Results from the *baseline* condition were consistent with classic symptoms associated with hemispatial neglect: columns 1 and 2 were significantly slower than columns 6 and 5, whereas column 3 was not significantly different from column 4. In the *uneven* probability condition, however, discrimination performance in column 1 was still significantly slower than column 6. Importantly, there were no significant differences between columns 2 and 5 and columns 3 and 4.

To assess whether the effect of statistical cueing was qualitatively different for control vs for patient populations, difference ratios between the two conditions were calculated for patient and for control participants for each column (see Figure 3b). Note that the control subjects here are just the elderly subjects described above. Most notably, the population (control, patient)  $\times$  column difference ratio interaction was not significant, indicating that the change in performance between the *baseline* and *uneven* condi-



**Fig. 3.** Mean reaction time as a function of screen column (a, left panel) in patients with hemispatial neglect for baseline and uneven conditions and mean difference ratios as a function of screen column for elderly controls and neglect patients (b, right panel). Error bars indicate standard errors.

tions as a consequence of the contingency in target location was equivalent in elderly and patient participants. As with the normal participants, none of the patients reported having noticed the probability difference during the experiment, further suggesting that the results are not a consequence of explicit strategy formation.

In sum, our data demonstrate that neglect patients show an additive sensitivity to statistical contingencies governing the distribution of targets in the visual field, that their behavior is modulated without explicit knowledge, and that the relative decrease in RT between conditions is equivalent to that seen in elderly participants. The results indicate that the behaviors of both normal and patient populations reflect sensitivity to the statistical spatial contingencies. Importantly, despite the obvious ability to exploit these contingencies, participants report being unaware of the uneven distribution of target locations. These results provide an important demonstration of how the visual attention system may adapt to environmental statistics reflexively in order to maximize the efficiency of behavioral output.

What neural mechanism might track these statistical effects? At least three possibilities exist: the facilitation occurs during the perceptual encoding of the input, during the sensorimotor transformation between the visual input and motor response (Platt & Glimcher, 1999), or during the planning or execution of the response. Although we do not definitively know the answer, the results from the patients render the second possibility unlikely because the effect of the uneven cueing did not alter the attentional gradient of patients *per se*. Furthermore, because neglect is often thought of as resulting from damage to regions of the brain that implement sensorimotor

transformations (Andersen, Snyder, Bradley, & Xing, 1997; Behrmann, Ghiselli-Crippa, Sweeney, Dimatteo, & Kass, 2002), it is likely that the facilitation reported in patients occurs in the encoding or the response process. The notion that the uptake of the target information is facilitated by the contingencies is consistent with the finding that early ERP components such as lateral occipital P1 and N1 are larger for targets appearing in expected rather than unexpected locations (Handy, Green, Klein, & Mangun, 2001). It is also in agreement with fMRI evidence that top-down attentional effects can selectively enhance V1 activation with concomitant suppression in surrounding regions; selective enhancement may act to reduce competition when multiple stimuli are present (Fink, Driver, Rorden, Baldeweg, & Dolan, 2000; Sengpiel & Huebener, 1999).

These results suggest that attentional expectancies in this experiment (although implicit) may provide feedback to early visual areas to enhance processing of objects located in the most probable region. It is also possible, however, that the contingencies facilitate subject responses, in this case, the saccadic eye movements that are necessary for target discrimination. The facilitation in target discrimination may arise because neurons involved in coding saccadic eye movements to the more probable side of space are primed. For example, Basso and Wurtz (1998) recorded from buildup neurons in the superior colliculus of monkeys performing a saccadic eye-movement task. They found greater activation during the delay period prior to target selection when the target always appeared in the same location compared to when it appeared in different locations. Their findings suggest that presaccadic activation is modulated by increased target location probability.

At present, exactly what gives rise to this facilitated discrimination performance as a function of contingency remains to be determined. Some functional imaging studies on this topic are under way and promise to be illuminating. We now turn to examining statistical learning of shapes, where we are better able to suggest a neural mechanism that likely supports this form of learning.

## Learning and acquisition of shape regularities

Statistical learning has been demonstrated for frequently co-occurring shapes embedded in simple displays (Chun & Jiang, 1999; Edelman et al., 2002; Fiser & Aslin, 2001). In these studies, observers typically are presented with a series of visual displays containing multiple stimuli. For example, Fiser and Aslin (2001) used displays comprising a three-by-three

grid containing six simple shapes. Across such displays, the joint probability (probability of co-occurrence) or conditional probability (probability of stimulus A given the presence of stimulus B) of stimulus pairs varies systematically. Learning of these stimulus statistics has been shown with a number of different dependent measures. Edelman, Hiles, Yang, and Intrator (2002) measured statistical learning as the greater reduction in reaction time for frequent compared with infrequent stimulus pairs in a probe detection task. Fiser and Aslin (2001) found that, in a forced-choice familiarity judgment task, observers could discriminate between frequent and infrequent stimulus pairs, suggesting an explicit representation of stimulus statistics. However, Chun and Jiang (1999) have argued that the memory may be implicit. They found shorter reaction times for frequent over infrequent target-distractor pairings in a visual search task, in the apparent absence of explicit memory for the target-distractor associations, measured in forced-choice familiarity judgments.

Although there are now several elegant studies in the domain, a number of questions persist – What factors influence performance? Do bottom-up factors such as perceptual organization override statistical regularity? Is it necessary to attend to the entire display to obtain the statistical effects? Is the sensitivity to the statistics explicitly reportable by participants? What neural mechanism mediates the learning? To examine this form of learning in more detail, we (Chris Baker, Carl Olson, and Marlene Behrman) developed a paradigm that could be used with both humans and monkeys; we report these findings here. In the humans, we measured RT and we also tracked performance in a forced-choice familiarity judgment to provide a comparison with the studies above (Baker, Olson, & Behrmann, 2004).

## **Acquiring new visual shape representations in normal participants**

In the experiments, displays were composed of two simple geometric stimuli (Figure 4) either unconnected or connected by a vertical bar. The total height of the displays was approximately 2.5 degrees of visual angle and the distance between the two stimuli was approximately 1 degree. There were eight target stimuli and eight distractor stimuli. Displays were constructed by combining one target and one distractor, producing 32 different displays. Each target was associated with a given response (left or right). Each distractor was paired equally often with left and right targets, so distractors carried no information about the appropriate response.

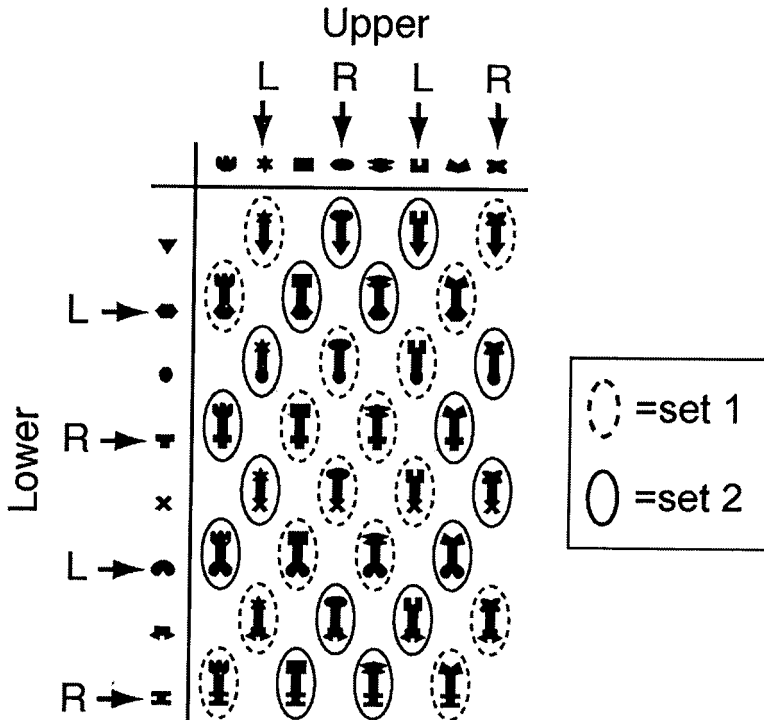


Fig. 4. Stimulus set for exploring visual pattern learning. The eight target parts are indicated by arrows and annotated with the designated response (R = right, L = left). The remaining eight parts are distractors and were equally associated with left and right responses. The solid (set 1) and dotted (set 2) ovals surrounding each stimulus indicate the frequency set to which the stimulus was assigned.

The critical manipulation employed in the experiments was the frequency of presentation of specific target-distractor combinations. Frequent combinations were presented four times as often as infrequent combinations. It was reasoned that if participants processed the target-distractor combinations they would respond faster and more accurately to the frequent than to the infrequent combinations. Displays were divided into two equal sets (set 1 and set 2, Figure 4). For a given participant, one set was designated "frequent" and the other set "infrequent," with set counterbalanced across participants. Each target was paired frequently with two distractors and infrequently with another two.

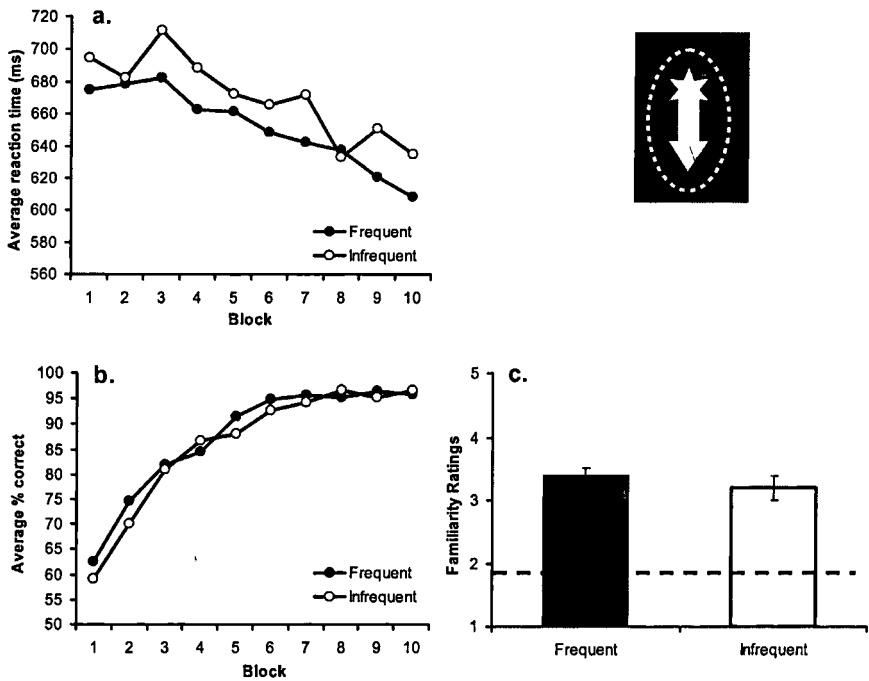
At the start of each trial, a fixation point was presented at the center of the screen and participants depressed two levers, one with each hand. The

fixation point turned red and remained so for 500 ms, until it was extinguished and a stimulus was presented for 100 ms. The participants had to release either the left or the right lever depending on the target present in the stimulus. On each trial, feedback was given: three short tones for a correct response and a large red circle flashing on the screen for an incorrect response. These parameters were chosen to be maximally similar to the monkey experiment to be reported next.

At the start of the experiment, participants were informed as to the nature of the task, but were not informed which targets instructed a left response and which a right response. They were allowed to perform practice trials on a separate set of stimuli before engaging in the main experiment. Subjects completed 10 blocks of trials; within every block, the frequent target-distractor combinations were presented four times each and the infrequent combinations once only. At the end of the experiment, participants were presented with 40 different stimuli and asked to rate them on a scale of 1-5 for familiarity (1 = least familiar, 5 = most familiar). All 32 of the experimental stimuli were presented, as well as 8 novel stimuli that had not been presented during the experiment. The novel stimuli were either target-target or distractor-distractor combinations.

In the first experiment, the upper and lower stimuli in the displays were connected. Four targets were presented at the upper location and four targets at the lower location (Figure 4). Participants were not informed which parts were targets and which distractors. Over the course of the 10 blocks, participants showed decreasing RTs and increasing accuracy for both frequent and infrequent target-distractor pairs. Critically, participants were faster and more accurate for frequent target-distractor pairings than for infrequent pairings (Figure 5a and b). Most relevant is the finding of a main effect of frequency on RT and on accuracy (although the latter was only marginally significant in this case). These results indicate statistical learning of the target-distractor combinations. Participants also rated frequent stimuli as significantly more familiar than infrequent stimuli (Figure 5c). This result confirms, in a test of explicit recall, statistical learning of the target-distractor combinations and suggests that the representations are explicit.

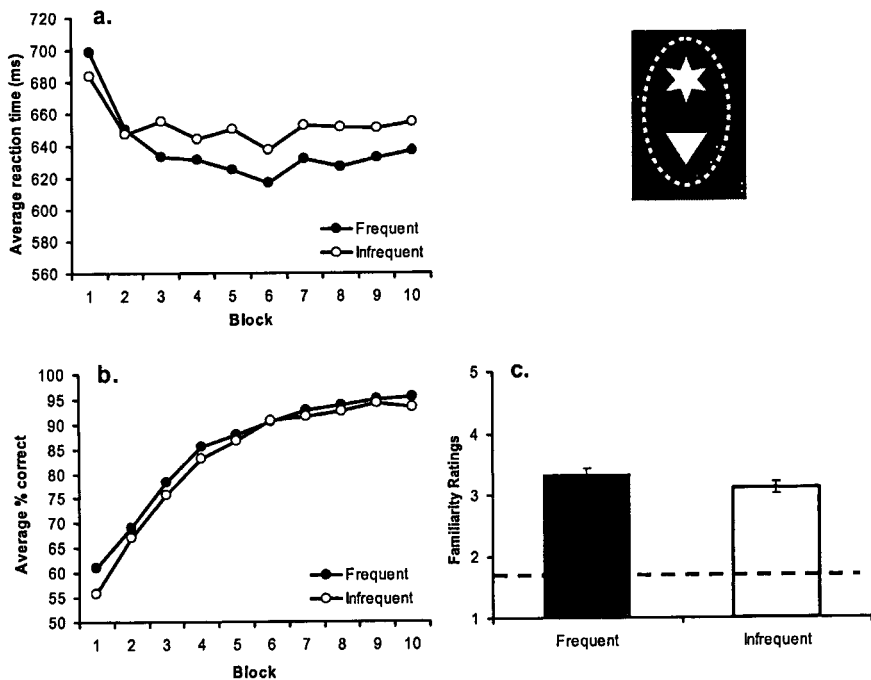
We conducted several other experiments to examine further the factors that influence this statistical learning. In one study, we examined whether participants would show the same sensitivity to the statistics if the parts were not connected (Figure 6). Both RT and accuracy showed the facilitation in performance for frequently occurring target-distractor combinations even if the top and bottom were separated, and this facilitation was as great as when they were connected. Participants also rated frequent stimuli as more familiar than infrequent stimuli (and to an equivalent degree whether



**Fig. 5.** Mean reaction times (a, top left panel); mean accuracy (b, bottom left panel); and mean familiarity ratings (c, bottom right panel) for subjects showing better performance on all measures for frequent target-distractor pairs over infrequent when subjects attend to targets at the bottom or top positions and when the parts are connected.

connected or not). That performance across the different dependent measures is equivalent independent of whether the parts are connected or not is surprising but suggests that learning is so robust that perceptual organization principles such as uniform connectedness (Palmer & Rock, 1994; Saiki & Hummel, 1998) are not necessary to bind the component parts together.

In the two experiments just described, targets were presented at both the upper and lower locations (unpredictably) so that participants had to attend to both locations. This attention might have served to bind the frequency pairings together. In further experiments, we asked whether the effect observed above was dependent on participants attending to both locations. We repeated the same two experiments with the upper and lower stimuli connected in one experiment (Figure 7) but not in a second (Figure 8), but in both cases, all eight targets were presented at one location only (either



**Fig. 6.** Mean reaction times (a, top left panel); mean accuracy (bottom left panel); and mean familiarity ratings (c, bottom right panel) for subjects showing better performance on all measures for frequent target-distractor pairs over infrequent when subjects attend to targets at the bottom or top positions but when the parts are not connected.

upper or lower, between subjects). Interestingly, even when participants are required to attend to one location only, they still show statistical learning of the target-distractor combinations regardless of whether they are attending to the upper location only or the lower location only when the stimuli are connected. This facilitation by frequency was also apparent in the subjects' explicit recall on the familiarity-rating task. However, when participants had to attend to one location only and the parts were unconnected, they no longer showed statistical learning of the target-distractor combinations. Under these conditions, participants were not faster nor were they more accurate for the frequent target-distractor combinations than for the infrequent combinations, and participants gave similar familiarity ratings to the frequent and infrequent stimuli. These results indicate that when participants have to attend to one location only and the parts are



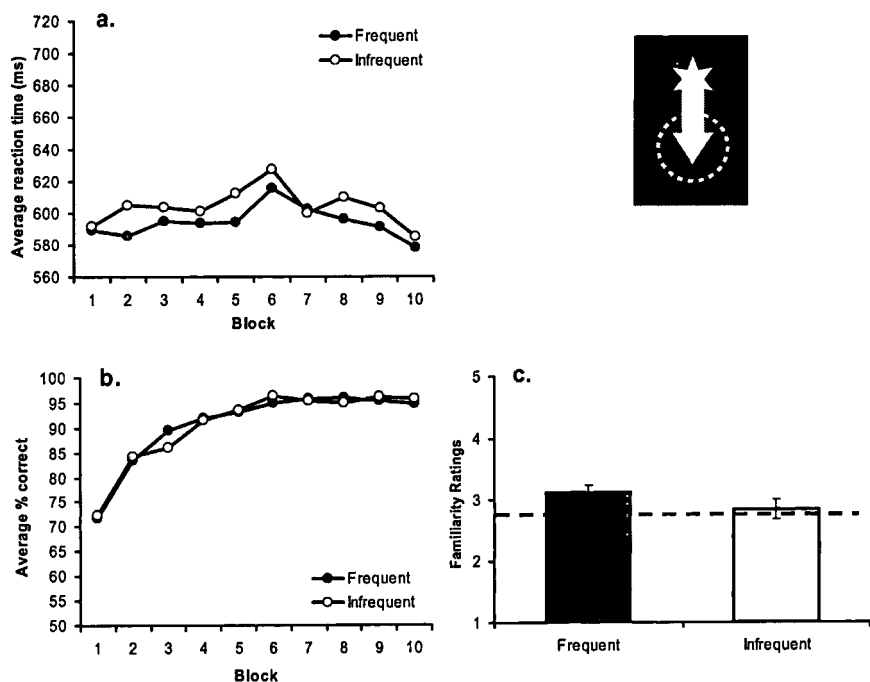
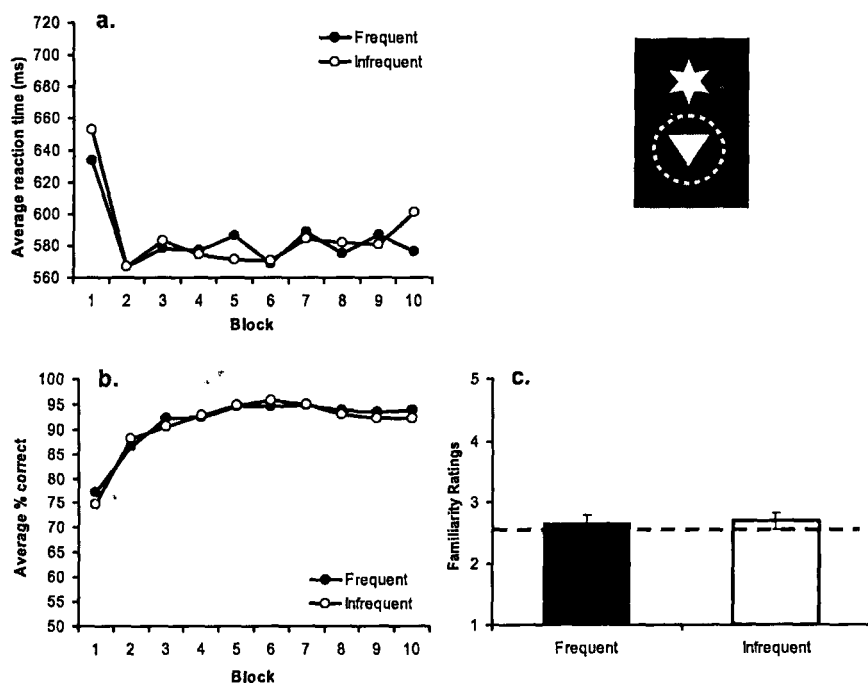


Fig. 7. Mean reaction times (a, top left panel); mean accuracy (bottom left panel); and mean familiarity ratings (c, bottom right panel) for subjects showing better performance for frequent target-distractor pairs over infrequent when subjects attend to targets at the bottom position and when the parts are connected.

unconnected, they no longer show statistical learning of the target-distractor combinations.

Taken together, these findings suggest that both attention and perceptual grouping can modulate visual statistical learning and that, importantly, these two factors interact. When participants attended to both stimulus locations, statistical learning was observed regardless of whether the stimuli were connected or not. When the participants attended to one location only, however, statistical learning was observed only when the stimuli were connected. Attention to the individual stimuli appears to be necessary, although this could be produced voluntarily by explicit direction of attention or involuntarily by the automatic spreading of attention induced by perceptual grouping. This suggests that in real world scenes, in the absence of explicit attentional control, statistical learning of feature combinations is much more likely for connected features than for unconnected features.



**Fig. 8.** Mean reaction times (a, top left panel); mean accuracy (b, bottom left panel); and mean familiarity ratings (c, bottom right panel) for subjects showing no difference in performance for frequent target-distractor pairs over infrequent when subjects attend to targets at the bottom position and the parts are not connected

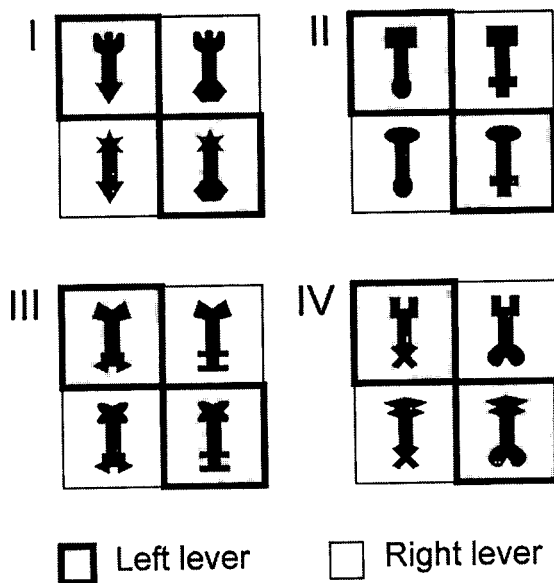
The statistical learning just described suggests that participants may be automatically encoding the upper and lower stimuli as a single whole object and forming unitary representations of the stimulus displays when attending to its entirety. Recent neurophysiological data from monkeys provide a possible neural mechanism by which such unitization could occur, as seen in the next section in which we describe an experiment in which we trained monkeys on a discrimination task using the same stimuli and found an increase in the number of neurons in inferotemporal cortex coding the conjunction of visual features.

## Acquiring new visual shape representations in non-human primates

The results just reported indicate that humans are sensitive to repeated combinations of top and bottom parts of an object even if one part is not relevant for the response, revealing considerable statistical learning. The question that arises is what neural system might possibly mediate this pattern of acquisition. One obvious candidate is inferotemporal (IT) cortex, given that neurons in this region have been shown to respond more similarly to visual stimuli that have been paired in prior training than to non-paired stimuli (Messinger, Squire, Zola, & Albright, 2001; Sakai & Miyashita, 1991). These results are also consistent with the data showing that lesions to IT impair pattern recognition. To examine the role of IT in pattern acquisition, we (Chris Baker, Carl Olson, and Marlene Behrmann) trained two monkeys in a discrimination task using the same stimuli that we had used for humans (bottom and top connected) (Baker et al., 2002). We then recorded from single neurons during a fixation task while the monkey fixated both learned and unlearned stimuli.

In this experiment, to ensure that there were no new emergent features when the top and bottom were juxtaposed (and that this single emergent feature was what the monkey learned, rather than the co-occurrence of the top and bottom part), we assembled tetrads of the stimuli by combining two top parts and two bottom parts orthogonally in a 2 x 2 design (see Figure 9). In the discrimination task, the monkey was taught to respond with a lever (right or left) depending on a particular combination of a top and bottom part. Because no part definitively indicated the response, to associate the stimuli with a response side, the monkey had to encode both the top and bottom part (see Figure 9). Each monkey was trained on two tetrads with the remaining two serving as the unlearned controls; learned tetrads for one monkey served as the unlearned controls for the second monkey and vice versa (see Figure 9 for examples). Once the monkeys performed the task well, we directly compared the neural responses elicited by learned versus unlearned stimuli. Recording sites, localized by magnetic resonance imaging, occupied the ventral aspect of the temporal lobe lateral to the anterior medial temporal sulcus, and thus were all in visual area TE of the inferior temporal lobe (see Figure 10 for recording site for monkey 1).

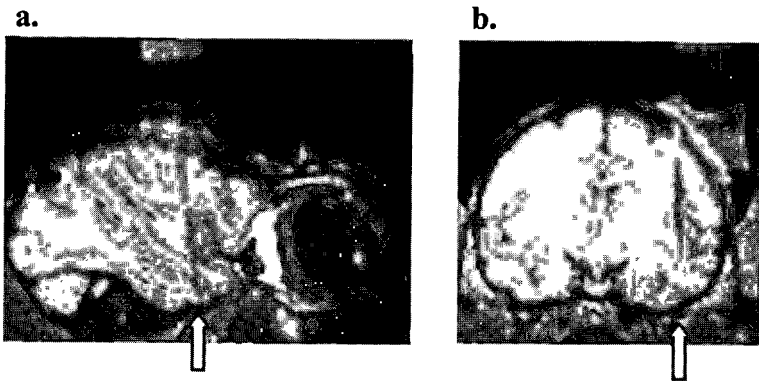
In the fixation condition, we assessed the neural response to all 16 stimuli and then chose one learned and one unlearned tetrad that evoked a neural response. Across the two monkeys, this was done 502 times (monkey 1: 331, monkey 2: 171). We were particularly interested in the number of



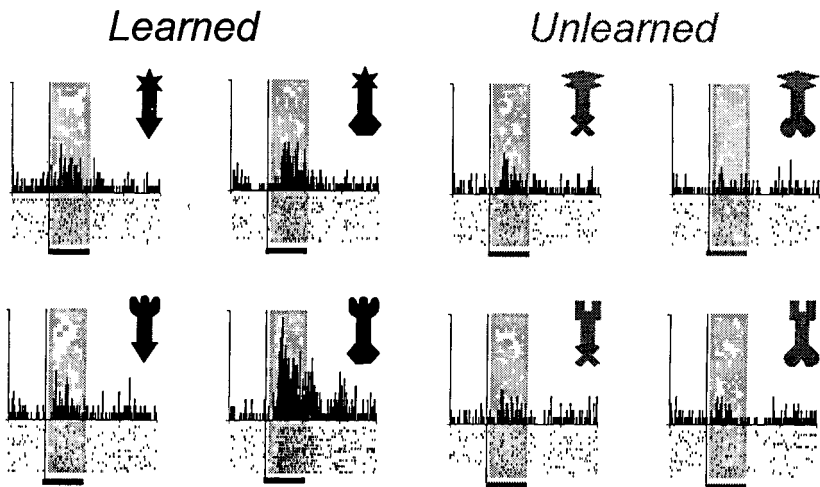
**Fig. 9.** Four tetrads of batons were used in discrimination training. Monkey 1 was trained on tetrads I and II and monkey 2 on tetrads III and IV. The batons that were used for training one monkey were also used as unlearned controls for the other monkey. Batons requiring left and right-lever responses are indicated by the different boxes around the stimuli. During the experiment, the boxes were not present.

neurons that responded selectively to a specific combination of a top and bottom part (see Figure 11 for an example), and to the extent that this selectivity differed for learned versus unlearned stimuli. To quantify the differential selectivity for the learned versus unlearned stimuli, for each neuron, we performed a two-way ANOVA with top and bottom part as factors, separately for learned and unlearned stimuli, and then counted up how many neurons showed no sensitivity to either the top or bottom part, one main effect, two main effects, or the interaction.

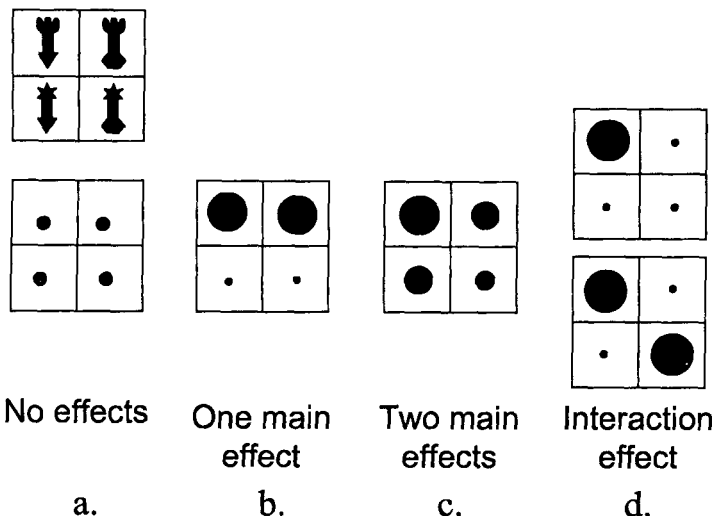
The schematic depiction of these various outcomes is shown in Figure 12, with the size of the circle showing the strength of the response. As is evident from Figure 13, in which we simply plot the number of neurons exhibiting a particular outcome on the ANOVA, there were no differences for the number of neurons that showed either one or two main effects for learned versus unlearned displays. There were, however, more frequent interactions of top and bottom factors for learned versus unlearned stimuli and fewer neurons showing no sensitivity for learned over unlearned stimuli.



**Fig. 10.** Magnetic resonance images showing the recording site in the ventromedial cortex in monkey 1: (a) sagittal and (b) coronal. The dark line running through the cortex is a shadow surrounding an electrode that was placed at the most medial recording site.



**Fig. 11.** Spike frequency histograms from a single neuron from Monkey 1 showing greater selectivity of response for combinations of particular top and bottom parts for learned over unlearned stimuli. Traces are aligned on the onset of the 500ms stimulus (vertical line). The duration represented by the entire horizontal axis is 2s.



**Fig. 12.** The two-way ANOVA with top and bottom part as factors could yield a number of possible outcomes, including selectivity (a) for neither part of the object, (b) for one part of the object, (c) for both parts of the object, independently, or (d) an interaction between top and bottom, such that one object produces a strong discharge (upper square) or objects sharing the same behavioral response but no parts in common produce a strong discharge. The interaction of interest here is the former, indicating a preferential response for a combination of a top and bottom part, especially in the learned case.

The important finding is that stimuli evoked a differential response from the neuron if the particular combination was learned. We note, however, that the statistical interaction of top and bottom factors need not necessarily emerge from the selective response in which we are interested (see Figure 12). For example, an interaction might also be observed if the two stimuli on the diagonal both elicited strong neural responses. Such a result emerges from an association of the stimuli with a particular response (recall that left and right responses are aligned with the stimuli on the diagonal) and is not of interest for this current investigation.

The final analysis consisted of ensuring that the increased number of neurons showing the top  $\times$  bottom interaction indeed displayed the super-additivity or nonlinearity of the component parts for one of the four stimuli. This was indeed the case, suggesting that, over the course of experience, the neurons came to represent not only the parts, as evident in the persisting main effect of parts, but also the combination of parts. A mechanism such as this might underlie the increased sensitivity to frequency observed in human experiments reported above, and demonstrates the ability

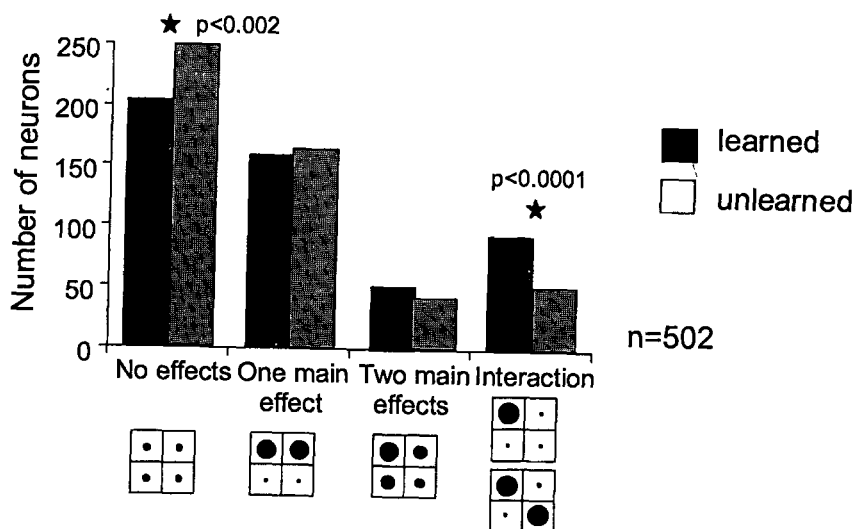


Fig. 13. The particular outcomes of the ANOVA from Figure 12 conducted separately for each neuron. The responses for each neuron appear in only one bar for learned and one for unlearned. The totals for the learned bars and the unlearned bars are 502 each. If a neuron showed a main effect and an interaction effect, then in this plot, it went into the interaction bar only. The major difference is the smaller number of neurons which showed no sensitivity for either part for the unlearned batons, and the greater number of neurons which showed the top x bottom interaction for the learned compared with the unlearned batons.

of inferotemporal (IT) cortex to acquire new visual representations as a function of experience.

## Conclusion

The focus of this chapter has been on the processes by which new visual representations are acquired. In particular, this was assessed in a variety of ways by examining the sensitivity of the visual system to the statistics of the input. Studies are described exploring the acquisition of spatial representations by virtue of sensitivity to spatial regularities in the environment. We showed that normal individuals and brain-damaged patients with hemispatial neglect are able to track the regularities, and that this facilitates target discrimination in a visual search task. We also showed that normal individuals are able to learn specific combinations of top and bottom parts of a stimulus and that this learning is so robust that it can survive a discon-

nection between the parts or attention to only one component. Monkeys trained with the same componential stimuli show neuronal responses that reflect the sensitivity to the particular combinations. Taken together, these findings illustrate how spatial and pattern representations are acquired (and maintained), and demonstrate that the visual system is exquisitely sensitive to the regularities of the input.

## Author Notes

The research reported here is supported by a grant from the National Institute of Mental Health (MH54246). Correspondence may be sent to Marlene Behrmann, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213-3890 USA. Alternatively, Marlene Behrmann can also be reached at behrmann@andrew.cmu.edu.

## References

- Aglioti, S., Smania, N., Barbieri, C., & Corbetta, M. (1997). Influence of stimulus salience and attentional demands on visual search patterns in hemispatial neglect. *Brain and Cognition*, 34, 388-403.
- Andersen, R. A., Snyder, L. H., Bradley, D. C., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience*, 20, 303-330.
- Baker, C., Behrmann, M., & Olson, C. (2002). Impact of visual discrimination training on the representation of parts and wholes in monkey inferotemporal cortex. *Nature Neuroscience*, 5, 1210-1216.
- Baker, C., Olson, C., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science*, 15, 460-466.
- Bartolomeo, P., & Chokron, S. (2001). Levels of impairment in unilateral neglect. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (Vol. 4, pp. 67-98). North-Holland: Elsevier Science.
- Basso, M. A., & Wurtz, R. H. (1998). Modulation of neuronal activity in superior colliculus by changes in target probability. *Journal of Neuroscience*, 18, 7519-7534.
- Baum, W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior*, 32, 269-281.
- Behrmann, M., Ebert, P., & Black, S. E. (2004). Hemispatial neglect and visual search: A large scale analysis from the Sunnybrook Stroke Study. *Cortex*, 40, 247-264.



- Behrmann, M., Ghiselli-Crippa, T., Sweeney, J., Dimatteo, I., & Kass, R. (2002). Mechanisms underlying spatial representation revealed through studies of hemispatial neglect. *Journal of Cognitive Neuroscience*, 14, 272-290.
- Bisiach, E., & Luzzatti, C. (1978). Unilateral neglect of representational space. *Cortex*, 14, 129-133.
- Bisiach, E., & Vallar, G. (2000). Unilateral neglect in humans. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (2<sup>nd</sup> ed., Vol. 1, pp. 459-502). North-Holland, Amsterdam: Elsevier Science.
- Chun, M. (2002). Contextual cueing of visual attention. *Trends in Cognitive Science*, 4, 170-178.
- Chun, M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, 36, 28-71.
- Chun, M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science*, 10, 360-365.
- Edelman, S., Hiles, B. P., Yang, H., & Intrator, N. (2002). Probabilistic principles in unsupervised learning of visual structure: Human data and a model. In T. G. Dietterich, S. Becker, & Z. Ghahramani (Eds.), *Advances in Neural Information Processing Systems* (Vol. 14, pp. 19-26). Cambridge, MA: MIT Press.
- Eglin, M., Robertson, L. C., & Knight, R. T. (1989). Visual search performance in the neglect syndrome. *Journal of Cognitive Neuroscience*, 1, 372-385.
- Esterman, M., McGlinchey-Berroth, R., & Milberg, W. (2000). Preattentive and attentive visual search in individuals with hemispatial neglect. *Neuropsychology*, 14, 599-611.
- Fink, G. R., Driver, J., Rorden, C., Baldeweg, T., & Dolan, R. J. (2000). Neural consequences of competing stimuli in both visual hemifields: A physiological basis for visual extinction. *Annals of Neurology*, 47, 440-446.
- 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. (2002a). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 458-467.
- Fiser, J., & Aslin, R. N. (2002b). Statistical learning of new visual feature combinations by infants. *Proceedings of National Academy of Sciences USA*, 99, 15822-15826.
- Geng, J. J., & Behrmann, M. (2002). Probability cueing of target location facilitates visual search implicitly in normal participants and patients with hemispatial neglect. *Psychological Science*, 13, 520-525.
- Gore, C. L., Rodriguez, D. P., & Baylis, G. C. (2001/2002). Deficits of motor intention following parietal lesions. *Behavioral Neurology*, 13, 29-37.
- Greggers, U., & Mauerlshagen, J. (1997). Matching behavior of honeybees in a multiple-choice situation: The differential effect of environmental stimuli on the choice process. *Animal Learning & Behavior*, 25, 458-472.
- Halligan, P. W., Manning, L., & Marshall, J. C. (1991). Hemispheric activations vs spatio-motor cueing in visual neglect: A case study. *Neuropsychologia*, 29, 165-176.

- Handy, T. C., Green, V., Klein, R. M., & Mangun, G. R. (2001). Combined expectancies: Event-related potentials reveal the early benefits of spatial attention that are obscured by reaction time measures. *Journal of Experimental Psychology: Human Perception & Performance*, 27, 303-317.
- Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton-top tamarins. *Cognition*, 78, B53-B64.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4, 267-272.
- Hoffman, J., & Kunde, W. (1999). Location-specific target expectancies in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1127-1141.
- Hornak, J. (1992). Ocular exploration in the dark by patients with visual neglect. *Neuropsychologia*, 30, 547-552.
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 83, B35-42.
- Lewicki, P., Czyzewska, M., & Hoffman, H. (1987). Unconscious acquisition of complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 13, 523-530.
- Lin, K. C., Cermak, S. A., Kinsbourne, M., & Trombly, C. A. (1996). Effects of left-sided movements on line bisection in unilateral neglect. *Journal of the International Neuropsychological Society*, 2, 404-411.
- Maljkovic, V., & Nakayama, K. (1996). Priming of popout II: The role of position. *Perception and Psychophysics*, 58, 977-991.
- Mattingley, J. B., Husain, M., Rorden, C., Kennard, C., & Driver, J. (1998). Motor role of human inferior parietal lobe revealed in unilateral neglect patients. *Nature*, 392, 179-182.
- Mayr, U. (1996). Spatial attention and implicit sequence learning: Evidence for independent learning of spatial and nonspatial sequences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 22, 350-364.
- Messinger, A., Squire, L. R., Zola, S. M., & Albright, T. D. (2001). Neuronal representations of stimulus associations develop in the temporal lobe during learning. *Proceedings of the National Academy of Sciences*, 98, 12239-12244.
- Palmer, S. E., & Rock, I. (1994). On the nature and order of organizational processing: A reply to Peterson. *Psychonomic Bulletin and Review*, 1, 515-519.
- Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, 400, 233-238.
- Riddoch, M. J., & Humphreys, G. W. (1983). The effect of cueing on unilateral neglect. *Neuropsychologia*, 21, 589-599.
- Riddoch, M. J., & Humphreys, G. W. (1987). Perceptual and action systems in unilateral visual neglect. In M. Jeannerod (Ed.), *Neurophysiological and neuropsychological aspects of spatial neglect* (pp. 151-181). New York: Elsevier.

- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27-52.
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, 8, 101-105.
- Saiki, J., & Hummel, J. E. (1998). Connectedness and the integration of parts with relations in shape perception. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 227-251.
- Sakai, K., & Miyashita, Y. (1991). Neural organization for the long-term memory of paired associates. *Nature*, 354, 152-155.
- Sengpiel, F., & Huebener, M. (1999). Spotlight on the primary visual cortex. *Current Biology*, 9, R318-R321.
- Shaw, M. L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception & Performance*, 3, 201-211.
- Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: The task dynamics of the A-not-B error. *Psychological Review*, 106, 235-260.
- Wilson, B., Cockburn, J., & Halligan, P. W. (1987). *Behavioral inattention test*. Suffolk, England: Thames Valley Test Company.