Representations of scene layout can consist of independent, functional pieces

THOMAS SANOCKI, KIMBERLY MICHELET, ERIC SELLERS, and JOSEPH REYNOLDS
University of South Florida, Tampa, Florida

A scene prime can induce a mental representation of layout that is functional in the sense that it facilitates the processing of depth relations in a subsequent same-scene target. Five experiments indicated that the representation can consist of separate and independent functional regions. In each experiment, primes with as many as four unrelated regions facilitated spatial processing within each region. The prime representations were functional despite structural discontinuity at region borders. The results indicate a limitation in the importance of structural constraint in representations of scene layout. However, when structural disruption occurred within regions that were perceived (Experiment 5), spatial processing was slowed. The results suggest that scene representation is more top down than is scene perception; the effects of structural disruption were overcome within representations, but not within perception.

In what ways might representations of a scene contribute to the ongoing perception of a scene? Potential functions include integration and prediction of scene structure (Henderson & Hollingworth, 1998; Hochberg, 1978; Intraub, 2002), facilitation of object processing (Biederman, 1981; Boyce, Pollatsek, & Rayner, 1989; Davenport & Potter, 2004), and navigation (e.g., Andersen, Hahn, & Saidpour, 2001; Vishton & Cutting, 1995). Another possible function, studied in the present research, is facilitating the processing of depth relations in familiar scenes (Sanocki, 2003; Sanocki & Epstein, 1997). Our concern is with representations of scenes that are familiar to viewers and with the role that structural constraints might play within those representations (Sanocki, 2003). Structural constraints are based on regularities of the environment, such as continuity of edges and surfaces. Organizational processes based on structural constraints are essential within many theories of perception (e.g., Biederman, 1987; Marr, 1982; Palmer, 1999; Palmer & Rock, 1994; Shepard, 1984; Wertheimer, 1924/1950; Witkin & Tenenbaum, 1983). We examined whether structural constraints are also critical for scene representation (e.g., Grossberg & Mingolla, 1985; Hummel & Biederman, 1992; Sanocki, 2003; Shepard, 1984) by exploring their role within functional representations of scene layout.

**Scene Priming and Representations
of Scene Layout**

The claim that layout representations are functional is based on scene-priming effects. A scene prime (see, e.g.,

Figure 1A) is presented for 1 sec, to activate the scene representation, or a control prime is presented (Figure 1B). The prime is followed by a brief blank interval and then a target picture that may be the same as the scene prime, except for two uniquely colored probes marking locations (e.g., two red probes; Figure 1C). Observers indicate which probe is closer to viewpoint (camera point) in the depicted space. Reaction time (RT) is the main measure (accuracy is usually high). The scene-priming effect is that RTs are shorter for spatial probes following same-scene primes than following control primes or different-scene primes (Sanocki, 2003; Sanocki & Epstein, 1997). The interpretation is that the scene prime activated a representation of the scene's layout, which facilitated the subsequent processing of spatial relations in the scene.

Previous results rule out many alternative explanations. The results indicate that the facilitation effect depends on information about the layout of objects and surfaces, rather than on the mere presence of objects and surfaces (Sanocki, 2003, Experiment 1). Furthermore, of particular interest here is the finding that the scene representations are functional across a broad scenic scope. That is, scene primes facilitated spatial processing across complex scenes, involving eight or more objects and major surfaces and many different spatial relations (Sanocki, 2003, Experiment 3). The scope of processing facilitated by a given prime can be illustrated with a probe activity map, which shows the set of spatial relations (probe pairs within the targets) that can follow a prime (Figure 2). Typically, each probe pair requires a different spatial interpretation, as defined by the objects or surfaces and their spatial relation (near vs. far by left vs. right). The probe set in Figure 2 involved 12 different interpretations within the scene. In a subsequent experiment, the scene representations facilitated processing of 20 relations, in foreground regions,

A. Scene Prime

B. Control

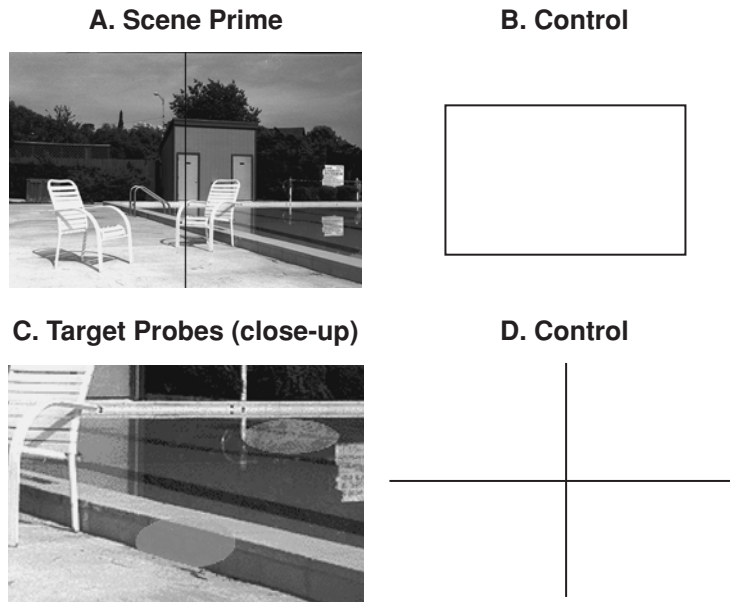**C. Target Probes (close-up)**

D. Control

Figure 1. Grayscale examples of stimuli. (A) Intact scene prime, Experiment 1. (B) Control prime, Experiments 1 and 2. (C) Close-up of target probes. (D) Control prime, Experiments 3 and 4. Photographic images appeared in color.

background regions, and unattended areas (Sanocki, 2003, Experiment 4). Other results are inconsistent with the possibility that facilitation results from persistence of sensory information from prime to target[1] or sudden onsets produced by the appearance of the target probes.[2]

Much previous work has been conducted with familiar scenes—that is, scenes viewed multiple times by the observers (Sanocki, 2003; Sanocki & Epstein, 1997). This practice is consistent with our emphasis on *representations*, which are most likely to exist and be functional after they have become established through experience. The reuse of scenes is also consistent with the fact that most human behavior takes place in familiar environments. Scene representations may become established quickly,

Figure 2. Probe activity map from Sanocki (2003, Experiment 3, large scope condition). Probes are gray ovals, and probe pairs are connected by lines. Some probes participated in two pairs.

however; Gottesman (2002) found facilitatory scene layout priming effects with novel scenes on each trial.

Explaining Scene-Priming Effects

The wide scope of the facilitation effects implicates a system that efficiently processes and represents spatial information distributed across the visual field. Evidence of efficient, distributed spatial perception has been provided by Enns and Rensink (1991), who found that certain basic spatial relations could be processed in parallel across many-element visual search displays (see also, e.g., Chun & Jiang, 1998). Enns and Rensink proposed an early vision system that processes typical, ecologically important spatial relations in an efficient manner across the visual field. In the present paradigm, the spatial-processing system would produce representations of much of the prime scene. The prime-based representation could provide spatial information (landmarks, depth cues, and spatial gradients) across much of the scene. When the target is presented, the spatial probes presumably attract attention to the probed regions in the target, and information available from relevant aspects of the prime representation is combined with target information to determine the correct response. Thus, the scene prime allows a head start on spatial processing, relative to control prime conditions, resulting in facilitation effects.

Role for Structural Constraints?

Although previous research has supported the idea of parallel spatial processing across the visual field, previous research does not tell us how a scene representation is maintained over time. In particular, how is a *broad-scope*
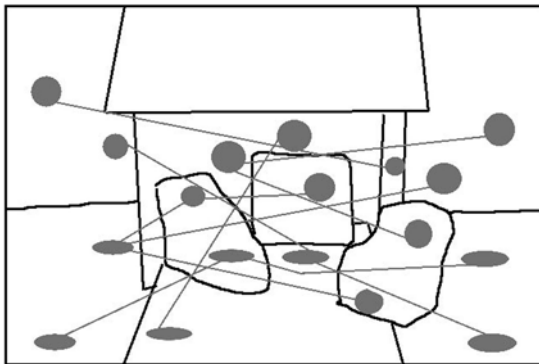
layout representation maintained over a short period of time? A hypothesis follows from the principle of structural organization. Structural organization has long been assumed to be necessary in theories of scene *perception*, because information contained in receptive fields and retinal images greatly underdetermines the interpretation of a scene. Therefore, stimulus information must be constrained during perceptual processing by structural relationships (see, e.g., Biederman, 1987; Marr, 1982; Palmer, 1999; Wertheimer, 1924/1950; Witkin & Tenenbaum, 1983). This is the process of perceptual organization, and the structural constraints include gestalt relations, such as parallelism and continuity, and higher order relations, such as same surface, cotermination, and same object (e.g., Biederman, Hilton, & Hummel, 1991; Palmer & Rock, 1994).

Perceptual processing produces representations that embody the structural constraints. Thus, structural constraints are integral to models of representations of forms and objects (Marr, 1982; Palmer, 1999; Witkin & Tenenbaum, 1983) and of surfaces and scenes (e.g., Marr, 1982; Sanocki, 2003). An analogous emphasis on structure holds in the extensive word perception literature, where structure is defined by co-occurrence in the language and other linguistic factors learned over years of reading (e.g., McClelland & Rumelhart, 1981; Rueckl, Mikolinski, Raveh, Miner, & Mars, 1997). Models of representation often instantiate structural constraints with connectionist principles (e.g., Grossberg & Mingolla, 1985; Hummel & Biederman, 1992; McClelland & Rumelhart, 1981). Thus, structurally consistent relations between elements in the representation activate excitatory connections, which establish and maintain the representation, whereas inconsistent relations inhibit the representation. These models have been supported by several literatures indicating that structural coherence leads to more efficient processing and representation than does disorganization. For example, in the word perception literature, four letters that together make a word are represented and processed more efficiently than letters that do not (McClelland & Rumelhart, 1981). Similarly, in object perception, structurally consistent object features activate an (implicit) memory representation, whereas structurally inconsistent features do not (e.g., Penney, Mecklinger, Hilton, & Cooper, 2000; Schacter, Cooper, Delaney, & Peterson, 1991).

We hypothesized that structural organization contributes to the integrity of broad-scope scene representations, thereby supporting their maintenance over time. Thus, better organized scenes should be represented more effectively than less organized scenes and should result in stronger facilitatory priming effects. We tested this hypothesis by disrupting global organization, by breaking images into pieces and rearranging the pieces (see, e.g., Figure 3A). Rearrangement disrupts global organization of the scene, violating numerous organizational relations, including parallelism, symmetry, and continuity of edges, objects, and surfaces. Each piece of the scene, however, was intact. In most cases, we measured spatial processing within the pieces; that is, both probes in a given target were contained within the same piece (as in Figures 1C and 3A). We examined the influence of global disruptions on the ability of pieces of scene representation to facilitate processing within their corresponding regions of the target. The main measure was the size of the facilitation effect caused by rearranged and intact scene primes. If structural organization is important throughout a scene representation, as was suggested above, rearranged scene primes should not be represented effectively and should produce less facilitation than do intact scene primes.
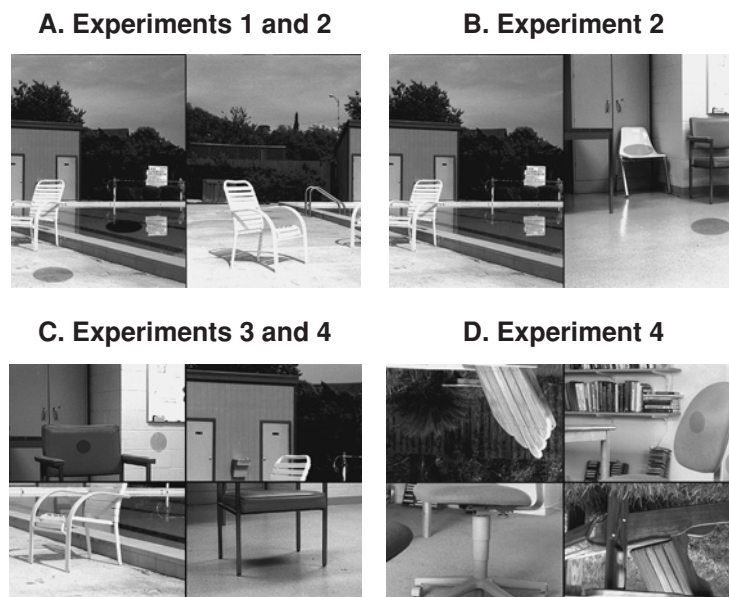
A. Experiments 1 and 2   **B. Experiment 2**

C. Experiments 3 and 4   **D. Experiment 4**

Figure 3. Grayscale examples of targets. Note that the images appeared in color in Experiments 1, 3, and 4.

The Present Experiments

The prediction was examined in the first four experiments. In Experiment 1, scenes were intact or rearranged by the interchanging of halves. As has been noted, rearrangement breaks numerous structural relations, including parallelism, symmetry, and continuity of edges, objects, and surfaces. In Experiment 2, scenes were rearranged semantically, as well as structurally, by mixing halves between scenes. In Experiments 3 and 4, rearrangement was further increased. To anticipate the results, we found that rearranged scene primes facilitated target processing as much as did intact scene primes. Thus, global structural coherence did not influence the functionality of the scene representations; each of the separate pieces of the representation was functional. Given this result, Experiment 5 changed the focus of the manipulations to begin examining whether structural organization is important at all in the representation or in the perception of scene layout.

EXPERIMENT 1

The purpose of Experiment 1 was to contrast the effectiveness of intact scene primes (Figure 1A) and scene primes that have been rearranged to disrupt overall organization (as in Figure 3A). The critical issue was how well the prime could be represented and used. This was measured by how much it facilitated processing in an identical target, relative to the control prime. Because the interest was in the functionality of the pieces, probe pairs were contained within one piece (one side) or the other, as was mentioned previously.

A preliminary study, also reported, indicated that the rearrangement manipulation did not interfere with *target* processing. When only control primes were used, rearranged and normal targets were responded to equally quickly. The results imply that red probes (the only red entities in the scenes) attracted attention immediately upon target onset and that processing was subsequently restricted to the probed side of the target. Therefore, the structural disruptions did not interfere with target processing. The similar efficiency of intact and rearranged target processing allows us to focus on the issue of how rearrangement influences the representation and use of the prime.

Method

Participants. The data in the main experiment were from 22 students from the University of South Florida, who participated for extra credit: 7 (6 of them female) in the *intact* condition and 15 (12 of them female) in the *rearranged* condition. Fewer observers were run in the intact conditions because it was very similar to standard conditions in many previous studies (Sanocki, 2003; Sanocki & Epstein, 1997). The data were discarded for 2 additional observers in the rearranged condition with low overall accuracy levels (<90% correct).[3] All the participants reported good or corrected vision. An additional 21 students participated in the preliminary study.

Design and Stimuli. Intact or rearranged stimuli were presented to separate groups of the observers because we wanted the observers to be able to maximize their strategies for the intact or rearranged primes. We were concerned that presenting a mixture of prime types might cause the observers to adopt a strategy that was optimal for one type, but not for the other.

The stimuli were derived from four color pictures taken from those used by Sanocki and Epstein (1997; a pool, front yard, basement, and study). The pictures were presented on a 13-in. color monitor in 256 colors and were 620×480 pixels ($23.5º \times 16.1º$ of visual angle at the 60-cm viewing distance). For the rearranged condition, the stimuli were split approximately in half vertically, and then the right and left halves were interchanged. In the intact condition, a black line was added where the scene was split for the rearranged condition, to produce local discontinuities (see, e.g., Figure 1A). The four intact and four rearranged pictures served as the scene primes.

From each of the four intact scene primes, 12 targets were created by superimposing pairs of red-toned ovals on scenic surfaces (i.e., redness was increased within the ovals). The pairs were positioned so as to be at different distances from viewpoint in the scene and to differ laterally (one probe more leftward than the other). Pairs were restricted so as to occur within one half or the other half of the split line, with six pairs occurring within each half. The probes were the only red-toned objects in the scenes, with the consequence that they should be easy to detect. Some probes were used twice, but in opposite spatial relations (i.e., close vs. far) in order to discourage responses based on a single probe. Also, some closer probes (four pairs, or 8.3%) were positioned higher in the picture in order to reduce validity of the visual cue *height-in-visual-field*.

Rearranged targets were made by interchanging the halves of the intact targets (see, e.g., Figure 3A). Thus, the spatial probes were identical across the arrangement condition, except for the change in side. The control stimulus is shown in Figure 1B.

Procedure. The observers participated individually in sessions lasting approximately 45 min. The 13-in. video monitor was enclosed in a large black frame, and the observers looked through a large rectangular cutout. This fixed viewing distance (at 60 cm). The observers were instructed to indicate which probe (left or right) was closer to viewpoint by pressing a left or a right key ("1" or "2" on the number pad). The observers were instructed to respond quickly but accurately, using fingers on their dominant hand. Each scene was shown once, as well as the control prime, and the observers were told to use all of these stimuli to "get ready." There were 480 test trials for each group, produced by combining the 4 scenes, 12 targets, 2 prime conditions, and 5 repetitions of each target. Test trials were preceded by an additional 32 practice trials randomly selected from the test trial conditions.

Each trial began as the prime was presented, along with a brief auditory beep. The prime duration was 1 sec, followed by an 84-msec blank (white) interval and then the target until a response was made. The computer changed consecutive images within 17 msec. After each response, auditory feedback was presented; a nice beep signaled correctness, whereas incorrect responses were followed by redisplay of the target, along with a gong sound.

Preliminary study. The speed of processing intact and rearranged targets was compared in a within-participants design by presenting only the control prime before intact and rearranged targets (in a randomized order). Mean correct RTs were 944 msec with intact targets and 940 msec with rearranged targets (the difference of 4 msec was not reliable; the standard error of the difference [*SED*] was 9 msec). Accuracy levels were similar (*M* = 97.4% for intact and 97.5% for rearranged targets). This indicates that intact and rearranged targets are responded to equally quickly when a prior scene representation has not been induced. The rearrangement does not disrupt processing of the target—presumably, because observers immediately shift their attention to the red probes when the target appears.

Results and Discussion

The critical issue is how the scene primes influenced processing, relative to control prime conditions. In the intact condition with the control prime, the mean RT for correct responses was 889 msec. When the targets were preceded by scene primes, responses were 44 msec faster

(*SED* = 10 msec). This difference (facilitation effect) was reliable [$F(1,6) = 20.98$, $p < .01$]. Accuracy was high overall and similar in the two conditions (*M*s = 96.6% for the scene prime and 96.5% for the control; $F < 1$). This is the standard facilitation effect found previously (Sanocki, 2003; Sanocki & Epstein, 1997).

In the rearranged condition, the control mean was 868 msec. With rearranged scene primes, responses were 45 msec faster [*SED* = 9 msec; $F(1,14) = 24.37$, $p < .01$]. Every participant showed a facilitatory priming effect. Accuracy was high overall and similar in the two conditions (*M* = 98.6% for the scene prime and 98.3% for the control; $F < 1$). Thus, scene primes were as effective when they had inconsistent overall organization as when they had consistent overall organization. It appears that each piece of the prime representation was functional, independently of the other piece.

Although the comparison between intact and rearranged priming was between observers, the relatively small standard errors indicate that the within-observer measures of priming were fairly sensitive. The intact facilitation effect is also comparable in size to that in previous experiments (Sanocki, 2003; Sanocki & Epstein, 1997), and the subsequent experiments provided further evidence on the size of rearranged priming effects.

To examine the effects of experience with the scenes and targets, the test session was divided into quarters (see Table 1). The average number of trials with each scene (counting practice trials) is shown in the left column of the table. For example, in the first quarter of the session, each scene was experienced a total of 30 times (on average), on Trials 9–38 with the scene (Trials 1–8 were practice). The actual experience varied with the target (probe pair), and each target was experienced once every 12 trials with a scene (an average between 2 and 3 times in the first quarter). Thus, the stimuli were familiar to the observers but were not highly overlearned in the first quarter. As can be seen, priming was generally constant across the session. ANOVAs for each group replicated the main effects of prime condition (reported above) and provided no evidence of an interaction with quarter of the experiment [intact, $F(3,18) < 1$; rearranged, $F(3,42) < 1$]. The substantial facilitation effects in the early quarters indicate that facilitatory priming does not require extensive practice with the scenes.

Could the overall priming effects be explained by the fact that the primes provided knowledge of the target scene, rather than its actual spatial layout? Relevant data were provided by an additional study with 16 observers. The rearranged targets in Experiment 1 (see, e.g., Figure 3A) were presented after the control prime, after identical rearranged primes, or after the intact scenic primes (see, e.g., Figure 1A). The intact scene primes provided information about the scene and its content, but this information did not map directly onto the layout of the rearranged targets (the sides switched between the prime and the target). Responses were faster with the identical, rearranged primes than with the control primes (by 57 msec; *SED* = 13). However, there was no reliable advantage for the intact primes (9-msec advantage over control; *SED* = 10, n.s.). This reinforces the idea that correspondence of spatial layout between the prime and the target is necessary for a facilitation effect and provides another baseline for comparison (see also Sanocki, 2003).

EXPERIMENT 2

The similar facilitation effects for intact and rearranged primes suggest that the two pieces of rearranged primes were independently functional; that is, each piece was represented and functional, regardless of the structure of the other piece. This finding was rather striking to us; it contrasts with findings that structural consistency contributes to the integrity of a representation (e.g., McClelland & Rumelhart, 1981; Penney et al., 2000; Schacter et al., 1991). However, the rearrangement manipulation can be strengthened in a way that increases the independence of (the differences between) the pieces of the primes. In Experiment 2, inconsistency was increased by disrupting semantics as well. Scene halves were mixed between scenes, creating semantically and structurally inconsistent scenes (e.g., the pool/basement in Figure 3B). There is a well-established literature on the importance of semantic organization in scene representation; a central theoretical construct is the idea of a scene schema that integrates semantic and structural information about a scene (e.g., Biederman, 1981). Because scene schemas are based on single coherent scenes, a semantically inconsistent scene should be represented less well than a semantically consistent scene (see, e.g., Biederman, 1981).

Table 1
Control Means and Facilitation Effects (in Milliseconds)
in Each Quarter of Experiment 1

Quarter (Trials With Scenes)	Intact			Rearranged		
	Control	*M*	*SED*	Control	*M*	*SED*
1 (9–38)	1,026	35	19	1,060	59	12
2 (39–68)	949	59	17	890	48	19
3 (69–98)	863	16	21	840	56	16
4 (99–128)	808	59	28	786	35	14

Note—*SED*, standard error of the difference.

A preliminary study, also reported below, contrasted the semantically consistent and inconsistent targets within the control conditions. It indicated that there was no difference in how fast such targets are responded to. Thus, the main issue of Experiment 2 is how the two types of rearrangement influence the representation and use of prime information.

Method

In the main experiment, the rearranged condition from Experiment 1 was replicated and termed the *rearranged, semantically consistent* condition. The second condition—the *rearranged, semantically inconsistent* condition—was created by the interchanging of halves between scenes, to produce four mixed scenes (e.g., Figure 3B). Because of salient color differences between scenes, the stimuli were converted to grayscale, and new target probes were created by darkening ovals or (in a few cases) by lightening ovals. There were 384 test trials (four repetitions of each target), preceded by 16 randomly determined practice trials. The data are from 30 students, 15 in each semantic condition (11 of them female in the consistent condition, 13 of them female in the inconsistent condition). An additional 4 participants were dropped from the analyses: 1 participant in the consistent condition for low accuracy (<90%) and 3 in the inconsistent condition for near-chance performance (<60% accuracy).

In the preliminary study with 18 additional observers, we compared RTs for the semantically consistent and inconsistent targets directly and within observers. RTs were 973 and 976 msec for the consistent and inconsistent targets, respectively. The 2-msec difference was not reliable ($SED = 5$ msec); accuracy levels were 98.1% and 98.4%, respectively ($F < 1$). This indicates that the processing of targets was unaffected by semantic inconsistency.

Results

In the rearranged, semantically consistent condition, the facilitatory priming effect in Experiment 1 was replicated; the control mean was 934 msec, and responses with the scene primes were 50 msec faster [$SED = 10$ msec; $F(1,14) = 26.74, p < .001$]. Accuracy was high overall and similar between conditions (Ms = 98.2% for the scene prime and 98.0% for the control; $F < 1$).

In the rearranged, semantically inconsistent condition, the control mean was 1,086 msec, and responses with scene primes were 64 msec faster [$SED = 14$ msec; $F(1,14) = 20.50, p < .001$]. Every participant showed a priming effect. Accuracy was high overall and similar between conditions (Ms = 98.2% correct; $F < 1$).

There was an apparent difference in RTs between semantic groups (152 msec in the control conditions), and this was examined with an ANOVA, with group and prime condition as factors. The main effect of group was not reli-

able [$F(1,28) = 2.80, p = .10$]. The sizable difference between groups appears to reflect the slowness of a few observers. The preliminary study provided a more sensitive, within-observers comparison of semantically consistent and inconsistent stimuli. As was reported in the Method section, the study indicated that there was no difference in the speed with which semantically consistent and inconsistent targets were processed in the control conditions.

The priming effects across the session are shown in Table 2. The test session was divided into thirds because it was shorter than that in Experiment 1. As can be seen, priming effects were generally constant across session. In an ANOVA (with factors of consistency group, prime, and third), there were no interactions (Fs < 1).

EXPERIMENT 3

Experiments 1 and 2 suggest that a representation of scene layout can have two separate pieces, each of which is functional. Can the idea of separate, independent pieces of representation be pushed further? Experiment 3 measured the effectiveness of scene primes cut into four separate pieces. We began with the semantically inconsistent stimuli in Experiment 3, cut each half horizontally into two pieces, and then mixed the pieces diagonally (see Figure 3C). Scene primes and targets always contained four pieces, and the observers could not predict which piece of the target would be probed.

Method

For each of the 4 four-piece scenes, 16 targets were created, with 4 of the targets having probe pairs within each quadrant. Because the previous experiments had indicated that pieces of the image were represented independently, we used color versions of the stimuli, with the effect that each four-piece image contained two somewhat different color schemes, one from each contributing scene. There were 384 test trials (three repetitions of each target), preceded by 96 practice trials. Practice was increased because of the unusual stimulus arrangement. Only the four-piece rearranged condition was run, with 12 observers (11 of them female). Data from 2 additional observers were dropped for low accuracy (<90% for one and <60% for the other).

Results

The control mean for the four-piece stimuli was 1,235 msec. With scene primes, responses were 54 msec faster [$SED = 13$ msec; $F(1,11) = 17.10, p < .01$]. Every participant save 1 showed the facilitatory priming effect. Accuracy was high overall and marginally better in the

Table 2
Control Means and Facilitation Effects (in Milliseconds)
in Each Third of the Test Session in Experiment 2

Third (Trials With Scenes)	Semantically Consistent			Semantically Inconsistent		
		Facilitation			Facilitation	
	Control	*M*	*SED*	Control	*M*	*SED*
1 (5–37)	1,048	56	15	1,199	45	31
2 (38–69)	922	68	13	1,072	79	22
3 (70–101)	834	24	15	997	62	18

Note—*SED*, standard error of the difference.

Table 3
Control Means and Facilitation Effects (in Milliseconds)
in Each Third of the Test Session in Experiment 3

| Third (Trials With Scenes) | Control | Facilitation | |
		M	SED
1 (25–56)	1,371	81	26
2 (57–88)	1,202	42	16
3 (89–120)	1,144	52	12

Note—*SED*, standard error of the difference.

scene prime than in the control condition [97.4% vs. 96.5%; $F(1,11) = 4.64$, $p = .05$]. Therefore, the observers were able to encode the four-piece scene primes and use them to facilitate spatial processing for each piece of the target. Each of four pieces of the representation was functional, despite the independence (lack of structural or semantic relation) between pieces.

The priming effects were generally constant across thirds of the session (see Table 3). In an ANOVA (with third and prime condition as factors), third did not interact with prime [$F(2,22) < 1$].

EXPERIMENT 4

In Experiment 4, we tried to push the idea of independence between regions of scene layout even further. Four-piece scenes were used again. However, in the critical mixed condition, two of the regions were transformed radically, to further disrupt global scene organization. The regions were inverted, as in Figure 3D. The ground, or gravitational basis, in scenes is a critical organizational factor for visually presented scenes. When the visual field is inverted optically, perception and action are severely disrupted initially (e.g., Schöne, 1984). The importance of orientation was confirmed in the present experiment. When all four pieces of the scenic images were inverted and control primes were used, responses took 278 msec longer than in the normally oriented control prime condition. Thus, inverting the scene disrupts spatial processing in the present task.

The critical issue in the experiment was the effectiveness of scene primes containing both upright and inverted pieces (see Figure 3D). Could each piece be represented and functionally effective, despite the absence of overall organization?

Method

Because of the increasing evidence of independence within scene representations, the present experiment was run entirely within participants. There were three orientation conditions, each involving 4 "scenes": the 4 four-piece upright scenes of Experiment 3 (*all upright*), 4 scenes in which all four pieces were inverted (*all inverted*), and 4 mixed scenes in which two pieces were inverted and two were not (the critical *mixed* condition; see, e.g., Figure 3D). As in Figure 3D, the inverted pieces were diagonally opposite in the mixed scenes. Both diagonals were used for inverted pieces, in different scenes. The probe pairs were the same as those in Experiment 3. (However, when the scenes are inverted, the closest probe switches sides [e.g., the "right" probe is now on the left]. Therefore, the correct responses were reversed with inversion.) There were 384

test trials (one repetition of each target within each of the two prime conditions), preceded by 96 practice trials. The data came from 22 university students. Data for 3 additional observers were discarded (accuracy was <90% for one and <60% for another; 1 observer's mean RT was more than 3 *SD*s above the group mean).

Results

The results are summarized in Table 4. For the all-upright stimuli, the control mean was 1,255 msec, and responses were 46 msec faster with scene primes than with the control prime [$SED = 19$ msec; $F(1,21) = 5.68$, $p = .03$]. Accuracy was high overall and similar between conditions ($M = 97.3\%$ scene prime, 96.7% control, $F < 1$). This replicates the priming of four separate pieces in Experiment 3.

The results from the all-inverted stimuli indicate that inverted primes can be represented and used effectively. Responses were 73 msec faster with the scene primes than with the control prime [$SED = 19$ msec; $F(1,21) = 13.84$, $p < .01$]. The control mean for the all-inverted stimuli was 1,533 msec, which was longer than that in the all-upright condition [$SED = 32$ msec; $F(1,21) = 73.46$, $p < .001$]. Inversion reduced accuracy somewhat, relative to all-upright stimuli ($M = 92.6\%$ for the scene prime and 92.4% for the control) but was similar between the two prime conditions ($F < 1$). In summary, spatial processing was slower and less accurate when all four pieces were inverted. However, the inverted primes were functional in that they facilitated processing.

Of most interest was the mixed condition. The control mean was 1,375 msec, intermediate between those for the other two stimulus conditions. Responses were 58 msec faster with the scene primes than with the control prime [$SED = 24$ msec; $F(1,21) = 5.82$, $p = .02$]. Accuracy was between that for the upright and that for the inverted stimuli ($M = 93.8\%$ for the scene prime and 94.6% for the control) and was similar between prime conditions ($F < 1$). This indicates that the observers were able to encode (represent) and use the mixed scene prime. In a further analysis, the data for probe pairs occurring in inverted and upright pieces were separated. The facilitation effect averaged 81 msec when upright pieces were probed and 34 msec when inverted pieces were probed. The difference between these two priming effects was not reliable [$SED = 29$ msec; $F(1,21) = 2.59$, $p > .10$]. Thus, both upright and inverted pieces of the scene layout representation were functional.

An ANOVA was used to examine effects of time within the session (third of session). There were no reliable interactions involving time.

Table 4
Control Means and Facilitation Effects (in Milliseconds)
in Experiment 4

| Condition | Control | Facilitation | |
		M	SED
All upright	1,255	46	19
All inverted	1,533	73	19
Mixed	1,375	58	24

Note—*SED*, standard error of the difference.

Discussion of Experiments 1–4

Experiments 1–4 indicate that representations of scene layout can be composed of structurally separate and independent pieces, or regions of representation. Each region is functional, facilitating spatial processing within that region of the target. These findings contrast with a basic result in object recognition, that coherent objects are more effectively represented and processed than are less coherent objects (e.g., McClelland & Rumelhart, 1981; Schacter et al., 1991). Furthermore, the principle of structural organization is essential in many theories of perception (e.g., Biederman et al., 1991; Marr, 1982; Palmer, 1999; Palmer & Rock, 1994; Wertheimer, 1924/1950; Witkin & Tenenbaum, 1983). This raises the question of why structural organization was not important for the scene representations examined here. As a step toward answering this question, we examined a more limited question in the next and last experiment: *Can* structural organization be important in the present task?

EXPERIMENT 5

In Experiment 5, we modified the methods in order to directly measure the effects of structural disruption on the perception of targets. We added *global* probe pair targets, in which a spatial probe appeared on each side of the picture, with disruptions falling in between (see, e.g., Figure 4B). The effect of structural disruption on target perception was measured most directly by performance in the control prime conditions, because there was no influence of scene prime information. On the basis of considerable prior research on the importance of structure for perception, we expected that the disruptions would slow the processing of rearranged targets, relative to the corresponding intact targets, in the control conditions.

If structural disruptions slow the processing of rearranged global probes in control conditions, will they also influence the scene representations? Note that global spatial relations—relations that span the two halves of the scene—were now relevant to the observers' probe task. With global relations now directly relevant to the observers' goals, it was possible that the scene representations would now be sensitive to structural disorganization. This would result in smaller priming effects for the global probe pairs and, perhaps, also for the local probe pairs in the previous experiments. Alternatively, the previous experiments suggested that structural constraint was *not* critical to the representation of scene layout, leading to the prediction that rearranged scenes could still be represented effectively in the present experiment, causing substantial facilitation effects with global probe pairs, as well as with local probe pairs.

Method

Because of the evidence of independence in the previous experiments, a totally within-participants design was used. We used the stimuli from Experiment 1, together with a new set of global probe targets. Each global probe pair had one member on each side of the scene. When the sides were interchanged in the rearranged condition, the absolute distances to the probes were preserved, along with the differences between probes. For each scene, six targets from Experiment 1 were used, now termed *local* probe pair targets. We also created six new global probe pair targets. Each combination of scene (four intact and four rearranged), prime condition (scene and control), and target (six global and six local) appeared twice during the test trials. There were 32 practice trials. The data came from 29 university students. Data for 3 additional observers were discarded (two with <90% accuracy, 1 with a mean RT more than three *SD*s above the group mean).

Results

Main results. The accuracy rates were generally high in all the conditions (>97%) and will be reported later. Control RTs provide the purest measure of how efficiently targets were perceived (without influences of scene primes) and are listed in Figure 4C. Consider first local probe pair targets. As in the pilot study for Experiment 1, there was no difference in control RTs for intact and rearranged targets ($F < 1$). Thus, rearrangement did not influence the speed with which local targets were processed. With rearranged global probe pair targets, however, structural disruptions occurred between the probes. Control RTs were 60 msec longer with the rearranged global targets than with the corresponding intact targets [$F(1,28) = 10.85$, $p < .001$]. Thus, structural disruption slowed processing in the present task when the disruptions occurred within the probed scene regions.

How effectively can the local and global scene relations be represented? The effects of the scene primes are shown in Figure 4D, together with indicators of their reliability in simple comparisons. For local probes, both intact and rearranged primes were reliably effective (see Figure 4D). However, the magnitude of facilitation was greater with intact than with rearranged scene primes [$F(1,28) = 4.38$, $p = .046$]. Also, in a direct comparison of RTs for local intact and rearranged probes following scene primes, there was a reliable difference [$F(1,28) = 4.82$, $p = .04$]. This suggests that intact layout information was represented more effectively than rearranged layout information in the present case.

With global probe targets, the facilitation effect was strong with the intact scenes (77 msec). There was also a reliable facilitation effect with the rearranged scenes (39 msec). Thus, rearranged global relations were effectively represented. The difference in magnitudes of facilitation for the intact and the rearranged global probes approached reliability [$F(1,28) = 3.83$, $p = .06$]. Thus, with both global and local probe targets, there was some evidence that layout information was represented more effectively with intact scenes than with rearranged scenes. However, the present priming results interacted with time within the session.

Changes within the session. The results for each third of the session are shown within parentheses in Figures 4C (control means) and 4D (facilitation effects). As can be seen in Figure 4D, for intact scenes, facilitation effects were generally strong throughout the session. However, for the rearranged scenes, facilitation effects were nonexistent in the first third of the session; the mean facilitation effect for global and local probes was 0 msec. These

Experiment 5

Intact Stimuli	Rearranged Stimuli

A. Primes

B. Targets

Global	Local	Global	Local

C. Control RT (msec)

832	899	892	892
(916, 809, 770)	(987, 875, 836)	(976, 878, 821)	(957, 868, 850)

D. Facilitation (msec)

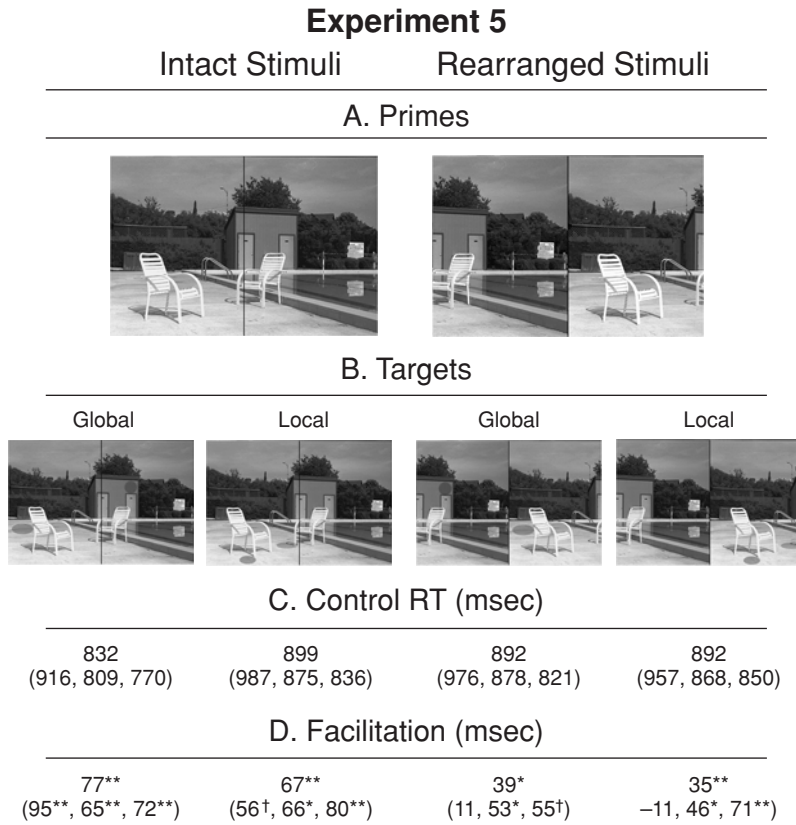| 77**	67**	39*	35**
(95**, 65**, 72**)	(56†, 66*, 80**)	(11, 53*, 55†)	(−11, 46*, 71**)

$*p < .05, **p < .01, †.05 < p < .10$

Figure 4. Stimuli and results in Experiment 5. The results for each prime and target condition appear below the target illustration. The results for each third appear in parentheses.

were Trials 5–20 with each intact or rearranged scene. The facilitation effects became strong in the remainder of the session, averaging 56 msec (Trials 21–52 per scene). The change in facilitation across the session for rearranged stimuli was reliable [$F(1,28) = 5.42, p = .03$, for the contrast of prime condition $\times$ 1st third vs. 2nd plus 3rd third]. Thus, in the present experiment, there were strong learning effects with the rearranged stimuli: The participants appeared to learn to represent the information necessary for facilitating processing of rearranged scenes. In the later portion of the experiment, the facilitation effects for the rearranged stimuli were as strong as those for the intact stimuli.

In each of the previous experiments, the observers presumably learned to use the rearranged information quickly during the practice sessions, because priming effects were as strong early in the sessions as later in the sessions.[4] The interaction in the present experiment may arise from the additional stimulus types within the design, as will be discussed later.

At the same time, note that in the *control* conditions, practice had little effect on the disadvantages for rearranged global targets, relative to intact global targets (which were 60, 69, and 51 msec in the respective thirds;

see Figure 4C). The structural disruptions slowed target processing throughout the session. What was learned during the session appears to have been the ability to use rearranged primes effectively.

Percentages correct. As has been noted, accuracy was generally high in all the conditions. An ANOVA was used, with all the relevant factors (intact vs. rearranged scene $\times$ prime condition $\times$ local vs. global target $\times$ third of session). There was a small main effect advantage for intact over rearranged scenes [98.1% vs. 97.4% correct; $F(1,28) = 8.36, p < .01$]. There was also a main effect in which accuracy improved over the session [97.0%, 97.8%, 98.5%; $F(2,56) = 7.31, p < .01$]. Finally, there was a small interaction of scene type and local versus global target, as is reported in Table 5 [$F(1,28) = 12.61, p < .01$]. Accuracy was somewhat lower with rearranged global targets, consistent with the disadvantage in baseline (control) RT.

Discussion

This experiment was designed to examine the effects of structural disruptions on the perception of layout information in targets, as well as the representation of the prime. Processing was slower for rearranged global targets than for intact global targets in control conditions.

Table 5
Interaction of Stimulus and Target Type (Percentages Correct)
in Experiment 5

	Global	Local
Intact	98.9	97.4
Rearranged	97.2	97.6

This indicates that structural organization is critical within processed regions; when structure is disrupted between the spatial probes, processing is slower and slightly less accurate. This negative effect of structural disruption was strong throughout the session.

At the same time, structural disruptions had only a modest impact on the representation of scene layout. In the main analysis, there were reliable facilitation effects for rearranged stimuli, with both global and local probe targets. The priming effects were smaller than those for intact stimuli but were positive and reliable nevertheless. Analyses within the session suggest that after a brief learning period, the rearranged facilitation effects were as strong as those for the intact stimuli. Facilitation increased from an average of 0 msec in the first third of the session to 56 msec in the rest of the session. The learning period appears to be necessary only when there is a mixture of prime types and probed relations, because there was no evidence of learning with rearranged primes in the first four experiments. Those experiments had more homogeneous stimuli (i.e., each group of observers saw only one prime type). Perhaps the mixture of stimulus types in Experiment 5, including more "normal" intact stimuli, reduced the learning of the less normal rearranged stimuli. In summary, the present results are consistent with those in the previous experiments in suggesting that disrupted spatial relations can be represented and used functionally to speed spatial perception.

Some caution should be taken in comparing the magnitudes of global probe facilitation effects between intact and rearranged scenes. Because the rearranged global probes were processed more slowly in the control conditions, there was somewhat more "room" for facilitatory effects of primes. This might inflate the size of the effect, relative to that with faster (intact) probes. For this reason, and because of the potential learning effects, precise estimates of the relative sizes of these facilitation effects will require further research.

One other caveat is that the disadvantage for rearranged global targets can be interpreted in at least two ways. First, it can be attributed to structural disruption. Second, it can be attributed to the disruption of depth cues caused by rearrangement. If it is possible to separate structural disruption from disruption of depth cues, this issue could be investigated in future research.

In summary, the present experiments provide clear evidence that structural constraint influences the perception of target information. However, structurally disrupted scenes can be represented and functional even when the task-relevant spatial relations are global in nature.

GENERAL DISCUSSION

The main finding of the present experiments is the rather striking functional independence for separate pieces of scene prime representations. Representations with two unrelated pieces were as effective in facilitating spatial perception as were unitary, intact counterparts in Experiments 1 and 2. In Experiments 3 and 4, scenes were bisected vertically and horizontally by severe structural inconsistency, and yet each piece was represented effectively and functionally. Thus, structural disruptions did not reduce the facilitatory effects of primes. This is contrary to the idea that structural organization may help bind scene representations together (Sanocki, 2003). More generally, this finding contrasts with a basic idea in models of perceptual representation that structural organization serves to bind the representations together. The importance of organization has been implied by a basic finding in object recognition, which is that structurally coherent objects are more effectively represented and processed than are less coherent objects (e.g., McClelland & Rumelhart, 1981; Schacter et al., 1991).

Experiment 5 indicated that structural organization can influence the perception of targets—when the disruptions occur between probed locations. However, rearranged scene primes were still represented in a functionally effective manner, facilitating the processing of both global and local target relations. Taken together, the results suggest that there are some limitations to the importance of structural constraints in representations of scene layout. Scene representations may contrast, in this regard, with object representations, for which organization is more likely to be necessary (see Biederman et al., 1991; Hummel & Biederman, 1992).

Limits and Possibilities of Structural Constraints

Organizational structure is thought to be essential to perception (see, e.g., Biederman, 1987; Marr, 1982; Palmer, 1999; Wertheimer, 1924/1950; Witkin & Tenenbaum, 1983). How can this principle be reconciled with the apparent nonnecessity for representations of layout? A potential answer to this question begins with the idea that representations are at least partly a top-down construction, abstracted from experience. Functional representations of scene layout are abstract with respect to lighting, mode of depiction, and retinal position (Sanocki, 2003; Sanocki & Epstein, 1997). Perhaps scene representations can consist of separate, structurally good pieces that are linked by abstract, nonvisual relations, such as *part of same (strange) picture* or *part of same view*. Perhaps the representations contain "snapshots" of where observers have attended in the picture and links between snapshots (cf. Hochberg, 1978). Thus, structural inconsistencies need not be represented literally; they could be replaced by links between structurally coherent regions, as when the rearranged global relations were represented in Experiment 5. The importance of linking separate views to create representations has been emphasized by Hochberg,

including the consequence that the representations can consist of functionally separate pieces (e.g., Hochberg & Peterson, 1987). These pieces might be treated as separate views within a single scene or, perhaps, a concatenation of separate pictures. The links might be as arbitrary as *contained within a single view*. Flexibility may be necessary to represent the full variety of layouts a person might view, which could include layouts with separate indoor and outdoor regions adjoined by doors or windows.

Further evidence of representational pieces linked together comes from the literature on spatial memory, where the idea of construction of representations from separate experiences is well documented (e.g., Tversky, Kim, & Cohen, 1999). Spatial memory can link together disparate pieces, such as landmarks, gradients, images, and even information learned from narratives. The ability to link together separate pieces would seem to be adaptive in human evolution because of the importance of developing spatial representation from separate experiences (Hochberg, 1978). For example, a person listening to navigational directions may integrate current views of the scene, remembered images or layouts, and verbal descriptions of future layouts or actions.

Thus, scene-priming effects with disrupted scenes can be explained by assuming that structural organization is important within pieces of the representation, but not across the entire representation. A scene representation can include links between functional separate regions. Such an explanation begins to emphasize top-down processes in the construction of a layout representation. At the same time, structural consistency remains a bottom-up constraint on scene *perception*, as was demonstrated in Experiment 5.

Independence and Two-System Theory

An alternative perspective on the present results comes from two-system theory (e.g., Milner & Goodale, 1995; Norman, 2002). The theory posits two somewhat independent, anatomically separate perceptual-processing streams (e.g., Milner & Goodale, 1995; Norman, 2002). Much of the spatial processing observed in scene-priming experiments could be a function of the dorsal stream, which processes spatial properties related to action. The spatial properties include the egocentric property of closeness to the observer. The system contrasts with a ventrally based system for processing object properties, identities, and perhaps, allocentric spatial relations (e.g., Milner & Goodale, 1995; Norman, 2002).

In applying two-system theory to the present situation, several assumptions about the functioning of the dorsal spatial action system must be made. First, priming implies that there is a memory for spatial properties that functions over short durations (see, e.g., Milner & Goodale, 1995; Norman, 2002). Second, the wide scope of scene priming (Sanocki, 2003) implies that the representation is wide in scope, which could be seen as a logical design feature of a system for action in three-dimensional space.

An important theme in two-system theory is the idea of functional independence. In particular, the spatial action

system can process input–output streams separately from the ventral system (e.g., Milner & Goodale, 1995; Norman, 2002). A system for guiding action within a three-dimensional world could exhibit a second type of independence as well: independence between separate regions of a scene representation, as was found in the present experiments. Consider an observer approaching two adjacent but distinct areas, such as a home entryway and a garden path, both of which are within the field of view. The present results suggest that both regions could be represented in parallel and that both regions of representation could be functional—that is, able to facilitate spatial processing. Such a system seems to be adaptive in a world with multiple regions for action. However, two-system theory assumes there would be no learning (long-term memory) of spatial relations in the dorsal system. The theory may have problems with the preliminary finding of learning effects in Experiment 5.

Null Effects of Semantic Inconsistency

There were no effects of semantic inconsistency in the experiments. This suggests that layout can be represented independently of meaning or, at least, coherence.

Familiar Versus Unfamiliar Scenes

We have focused on the representation of familiar scenes for two reasons. First, most human behavior occurs within familiar scenes. Second, familiar scenes are most likely to be represented, and our interest is in representations. This leaves open the question of how novel scenes might be represented. As Experiment 5 indicates, structurally disrupted scenes are more difficult to perceive in control conditions. Because they are difficult to perceive, novel disrupted scenes may also be represented less effectively, at least initially. This implication is consistent with the lack of facilitation for rearranged scene primes early in Experiment 5. However, the finding of strong early priming effects for rearranged scene primes in Experiments 1 and 2 suggests that learning can occur quickly for disrupted novel scenes, producing functionally effective representations.

Scope of Scene Layout Representations

The present results reinforce the claim that layout representations can be broad in scope, facilitating spatial processing across a full scene (Sanocki, 2003). Not only were the present representations broad in scope, but also they included multiple regions that were separate and complex. Scope and complexity can be illustrated with a probe activity map (see, e.g., Figure 2). Figure 5 shows the probe activity map for a representative scene from the present Experiment 3 (the image is shown in Figure 3C). As can be seen, the spatial relations probed in each region were fairly complicated and cannot be reduced to a single object or a single depth relation. The scope of priming raises questions that remain unanswered: How can a broad scope representation be maintained? Is the ability to maintain multiple regions of functional representation an operating characteristic of the human brain?
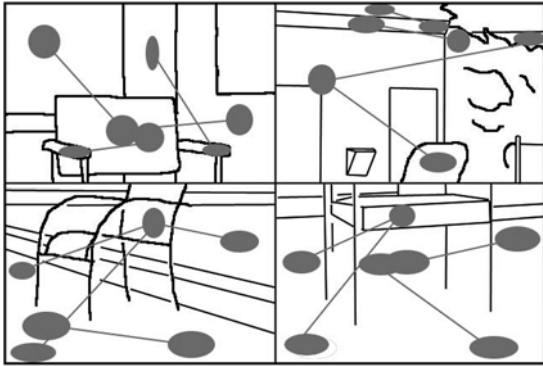
Figure 5. Probe activity map from Experiment 3. Probes are gray ovals, and probe pairs are connected by lines. Some probes participate in two pairs.

REFERENCES

ANDERSEN, G. J., HAHN, S., & SAIDPOUR, A. (2001). Static scene information and the perception of locomotion [Abstract]. *Journal of Vision*, **1**, 2a. Available at journalofvision.org/1/3/2, DOI 10.1167/1.3.2.

BIEDERMAN, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 213-253). Hillsdale, NJ: Erlbaum.

BIEDERMAN, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, **94**, 115-147.

BIEDERMAN, I., HILTON, H. J., & HUMMEL, J. E. (1991). Pattern goodness and pattern recognition. In G. R. Lockhead & J. R. Pomerantz (Eds.), *The perception of structure: Essays in honor of Wendell R. Garner* (pp. 73-95). Washington, DC: American Psychological Association.

BOYCE, S. J., POLLATSEK, A., & RAYNER, K. (1989). Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 556-566.

CHUN, M. M., & JIANG, Y. (1998). Contextual cuing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, **36**, 28-71.

DAVENPORT, J. L., & POTTER, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, **15**, 559-564.

DI LOLLO, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General*, **109**, 75-97.

ENNS, J. T., & RENSINK, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review*, **98**, 335-351.

GOTTESMAN, C. V. (2002, November). *Mental extrapolations as primes for spatial layout processing*. Presented at the 43rd Annual Meeting of the Psychonomic Society, Kansas City, MO.

GROSSBERG, S., & MINGOLLA, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, **92**, 173-211.

HENDERSON, J. M., & HOLLINGWORTH, A. (1998). Eye movements during scene viewing: An overview. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 269-293). Amsterdam: Elsevier.

HOCHBERG, J. E. (1978). *Perception* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

HOCHBERG, J. [E.], & PETERSON, M. A. (1987). Piecemeal organization and cognitive components in object perception: Perceptually coupled responses to moving objects. *Journal of Experimental Psychology: General*, **116**, 370-380.

HUMMEL, J. E., & BIEDERMAN, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, **99**, 480-517.

INTRAUB, H. (2002). Anticipatory spatial representation of natural scenes: Momentum without movement? *Visual Cognition*, **9**, 93-119.

MARR, D. (1982). *Vision*. San Francisco: Freeman.

MCCLELLAND, J. L., & RUMELHART, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, **88**, 375-407.

MILNER, A. D., & GOODALE, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

NORMAN, J. (2002). Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches. *Behavioral & Brain Sciences*, **25**, 73-96.

PALMER, S. E. (1999). *Vision science*: *Photons to phenomenology*. Cambridge, MA: MIT Press.

PALMER, S. [E.], & ROCK, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review*, **1**, 29-55.

PENNEY, T. B., MECKLINGER, A., HILTON, H. J., & COOPER, L. A. (2000). Priming and recognition of novel 3D objects: Guidance from event-related potentials. *Cognitive Science Quarterly*, **1**, 67-90.

RUECKL, J. G., MIKOLINSKI, M., RAVEH, M., MINER, C. S., & MARS, F. (1997). Morphological priming, fragment completion, and connectionist networks. *Journal of Memory & Language*, **36**, 382-405.

SANOCKI, T. (2003). Representation and perception of scenic layout. *Cognitive Psychology*, **47**, 43-86.

SANOCKI, T., & EPSTEIN, W. (1997). Priming spatial layout of scenes. *Psychological Science*, **8**, 374-378.

SCHACTER, D. L., COOPER, L. A., DELANEY, S. M., & PETERSON, M. A. (1991). Implicit memory for possible and impossible objects: Constraints on the construction of structural descriptions. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 3-19.

SCHÖNE, H. (1984). *Spatial orientation: The spatial control of behavior in animals and man* (C. Strausfeld, Trans.). Princeton, NJ: Princeton University Press.

SHEPARD, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, **91**, 417-447.

TVERSKY, B., KIM, J., & COHEN, A. (1999). Mental models of spatial relations and transformations from language. In G. Rickheit & C. Habel (Eds.), *Mental models in discourse processing and reasoning* (pp. 239-258). Amsterdam: Elsevier.

VISHTON, P. M., & CUTTING, J. E. (1995). Wayfinding, displacements, and mental maps: Velocity fields are not typically used to determine one's aimpoint. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 978-995.

WERTHEIMER, M. (1950). Gestalt theory. In W. D. Ellis (Ed.), *A sourcebook of Gestalt psychology* (pp. 1-11). New York: Humanities Press. (Original work published 1924)

WITKIN, A. P., & TENENBAUM, J. M. (1983). On the role of structure in vision. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 481-543). New York: Academic Press.

NOTES

1. For example, scene-priming facilitation was invariant across changes in image position, which should prevent sensory integration of the prime and the target (Sanocki, 2003). Sanocki (2003, Appendix) measured sensory persistence directly, using a variant of the Di Lollo (1980) integration paradigm, and found no evidence of sensory persistence.

2. The probes in most experiments were changes in color or darkness (Figure 1C), which could, in theory, be more salient against the background of an identical scene prime than of a control prime. However, scene priming also occurs with drastic changes in image values between the scene prime and target, caused by dramatic changes in lighting (Sanocki, 2003) or by changing from a line drawing scene prime to a detailed color target (Sanocki & Epstein, 1997). Therefore, scene priming cannot be explained by assuming that probes provide the only salient onsets in the scenes.

3. A small portion of participants are unable to perform this task accurately, some despite long RTs. The nature of their problems has not been investigated.

4. There was a small difference between Experiment 5 and the previous ones in the amount of experience with each scene in the early phases of the sessions. Because of the more complex design, there was less

experience per scene (5–20 trials) in the early phase of Experiment 5. For comparison, we examined the initial portions of the first two experiments (up to the 20th trial per scene) to see whether the facilitatory effects for rearranged stimuli might be weaker then. Facilitation remained strong, averaging 54 msec. The effects in each experiment were the following: Experiment 1 rearranged, 81 msec; Experiment 2 semantically consistent, 46 msec; Experiment 2 semantically inconsistent, 36 msec. These data were from Trials 9–20 (Experiment 1) and 5–20 (Experiment 2) with each scene. Experiments 3 and 4 could not be analyzed in this way because they had longer practice periods (motivated by the more unusual stimulus arrangements).