Priming of Simple and Complex Scene Layout: Rapid Function From the Intermediate Level

Thomas Sanocki and Noah Sulman
University of South Florida

Three experiments examined the time course of layout priming with photographic scenes varying in complexity (number of objects). Primes were presented for varying durations (800–50 ms) before a target scene with 2 spatial probes; observers indicated whether the left or right probe was closer to viewpoint. Reaction time was the main measure. Scene primes provided maximum benefits with 200 ms or less prime duration, indicating that scene priming is rapid enough to influence everyday distance perception. The time course of prime processing was similar for simple and complex scene primes and for upright and inverted primes, suggesting that the prime representation was intermediate level in nature.

Keywords: priming, layout, scene perception, intermediate-level processing, spatial processing

Human abilities such as navigation and scene identification are amazingly good compared to artificial systems. Of particular interest here is the representation and perception of spatial layout within familiar scenes, as examined with a scene priming paradigm. Our hypotheses begin with the assumption that observers extract a representation of a scene's layout over a brief learning period (as short as a few trials with normal scenes; see Sanocki, Michelet, Sellers, & Reynolds, 2006). This representation is assumed to be activated by brief exposure to the scene (a scene prime) and to facilitate distance perception when a similar target scene follows the prime (Sanocki, 2003; Sanocki & Epstein, 1997; Sanocki et al., 2006).

Speed of Processing and Scope of Processing

We view the benefits of scene primes as one facet of the efficient use of spatial information from the environment (see, e.g., Cutting & Vishton, 1995; Domini, Caudek, & Tassinari, 2006; Gibson, 1979; Ni, Braunstein, & Andersen, 2005; Sedgwick, 1986). Another facet of efficient information processing is speed. In the literature on identification of scene categories, the speed of identification is a provocative finding. In seminal experiments, Potter (1975, 1976) found that observers can pick out a prespecified scene category (e.g., beach) with high accuracy from a stream of different scenes, each presented for as little as 167 ms (see also, e.g., Evans & Treisman, 2005; Michod & Intraub, 2007). Using briefly presented low- and high-pass images, Oliva and Schyns (1997; Schyns & Oliva, 1994) found that scene categorization is facilitated by general layout information even when objects are obscured. And, in binary categorization tasks (e.g., Is an animal present?) with novel scenes on each trial, evidence of rapid categorization has been found (e.g., Rousselet, Fabre-Thorpe, & Thorpe, 2002; Rousselet, Joubert, & Fabre-Thorpe, 2005). This includes directional eye movement responses initiated in less than 120 ms from stimulus onset on many trials (Kirchner & Thorpe, 2006).

However, note that categorization requires only a single decision about each scene, limiting the scope of scene processing to that decision. Yet, a typical scene contains a complex of information distributed across space, including multiple spatial relations. Is scene processing efficient across the spatial extent of a scene? Here, we measured the speed of processing scene primes that represented a typical view and contained either a few or many spatial relations. We presented the scene primes for varying amounts of time and used their effects to infer how much they were processed. We asked the following question: How long must the prime be presented for its benefits for target processing to reach maximum levels?

Other seminal research in scene perception has demonstrated effects of semantic constraints and physical-layout constraints throughout much of the scenes (see, e.g., Biederman, 1972, 1981). These experiments are consistent with the hypothesis of rapid processing of an entire scene; however, the experiments were not designed to provide precise information about the time course of processing, including particular stages of processing.

Stages of Processing

Most theories of vision assume that processing proceeds through several stages, from lower to high levels. In the present experiments, we examined the stage or stages at which the prime-induced scene representations operate. At one extreme, scene primes could be processed to create high-level representations, such as a scene category or a construction of attention (e.g., Logan, 1995; Rensink, 2000). However, such representations are limited in scope, to no more than several arguments or entities; consequently, their usefulness for complex scene layouts would be limited. At the other

extreme, the representations could function at a low level, such as sensory copies of an initial fixation (see, e.g., McConkie & Rayner, 1976). However, sensory representations are also limited, because they are tied literally to a sensory snapshot. There are several types of evidence against explanations of scene priming benefits based on sensory representations, including direct measurements that found no evidence of sensory integration and findings of invariance across lighting and position (Sanocki, 2003; Sanocki & Epstein, 1997).

In contrast to low- and high-level representations, intermediate-level representations are wide in scope and could be general enough to be useful across the variegated challenges of everyday scene perception. The challenges include transformations of incoming information due to eye movements, attention shifts, and observer movement. In the literature, intermediate-level representations are usually assumed to be coarse, partially developed representations driven by stimulation efficiently processed through early vision. Although driven by stimulation, the representations are also somewhat abstract, which increases their generality of function. In different research contexts, intermediate representations have been conceived of as a 2.5D sketch (e.g., Marr, 1982), as early object files (e.g., Kahneman, Treisman, & Gibbs, 1992; Henderson & Anes, 1994), as a stew of proto-surfaces and proto-objects (e.g., Rensink, 2000; Wolfe & Bennet, 1997), and as inchoate structural networks (Sanocki, 2003). All of these conceptions share the assumptions that the representations change quickly with stimulation, are moderately abstract, and are inchoate rather than complete.

We hypothesized that scene primes would induce intermediate-level representations of layout and that these representations would facilitate subsequent processing. Previous findings of invariance of scene priming across position of the image on the eyes and lighting (Sanocki, 2003; Sanocki & Epstein, 1997) are consistent with the idea of general, intermediate-level representations of scene primes.

The intermediate-level construct is essential in most theories of vision, but the evidence on the construct is still relatively sparse. Prime-induced representations of scene layout might be an empirically tractable example of functional intermediate-level representations. If so, then our measurements of the speed of scene processing will pertain mainly to the processes that produce intermediate-level representations.

How Quickly Do Layout Representations Become Functional?

The ultimate task on each trial was a relative distance task (Sanocki & Epstein, 1997). A prime was presented on each trial, followed by a target picture that contained the response-relevant information—two markers (probes) indicating laterally separated locations in the scene. The probes appeared in a variety of locations throughout the scene. Observers indicated whether the left or right probe was closer to viewpoint (camera point) by pressing a spatially corresponding key. Reaction time was the main measure. Representations were manipulated by using differing primes—either the scene itself (sans probes) or a control prime. The scene priming benefit is that responses are faster following same-scene primes than following control primes and different-scene primes (Sanocki, 2003; Sanocki & Epstein, 1997).

To examine how quickly scene primes are processed, we varied how long the prime was present before the target—lead time. The shortest lead time was none (0 ms, or no prime), which provided one baseline measure. In this condition, only a general preparation stimulus (a cross) appeared during the 1,200-ms pretarget interval. There should be no prime-based benefits in this condition. The positive lead times, during which a prime stimulus was continually present, ranged from 50 ms to 800 ms in the experiments. The trial sequence is schematized in Figure 1A. As lead time (prime duration) increases (leftward from the 0-ms point; see vertical line in Figure 1A), there is more time for a prime scene to be processed and for the representation to become fully effective. As lead time increases with scene primes, reaction time to targets is expected to decrease. Eventually, reaction time should decrease to a minimum, which reflects the maximum prime-based preparation level. The shape of this function should be informative. Consider possible functions for scene primes that match the target scene. The function could be a steady reduction in reaction time as lead time gets longer, suggesting a gradual increase in functionality over the lead time period (see Figure 1B, dark line). Or the function could change abruptly, implying that the scene representations require a fixed amount of time to become functional (see Figure 1B, light line).

Additive Factors Logic

To examine the stage at which layout representations function, we analyzed performance in terms of stages of processing and applied additive factors logic (Sternberg, 1969, 1998). The assumptions of this logic are controversial, especially when close or overlapping stages of processing are contrasted; however, the method can be informative when distinct stages are contrasted and when additional constraints are available (see, e.g., McClelland, 1979; Miller, 1988; Sternberg, 1998). In applying the logic here, we made a broad distinction between early and late phases of visual processing. Such a distinction is supported by much work in visual cognition (e.g., Neisser, 1967; Treisman & Gelade, 1980). The early phase includes two stages of processing, the extraction of features and the integration of features into a coarse, intermediate-level representation. This phase is assumed to be rapid, efficient, and generally parallel across the visual field. The process appears to be primarily local and includes some basic relations (e.g., Enns & Rensink, 1991; Treisman & Gormican, 1988). Later processing takes as input the intermediate representation and resolves the representation in a task-relevant manner. This phase of processing is assumed to be flexible, task dependent, and limited in scope and attentional resources, and it may involve several stages.

The early–late distinction can be applied to our interpretation of processing in the present task. When a scene is first presented on a trial, we assume that early processes produce an intermediate-level representation of layout across the display. The representation may include coarse information about objects and object parts, as well as rough surface gradients and nonmetric relations between regions or proto-objects. If the first scenic stimulus was a scene prime, then the intermediate representation becomes available for subsequent processing. To be useful, the prime and target information must match at moderately abstract levels (Sanocki, 2003). When the scene prime matches the target, late processing can begin soon after the target appears. When the first stimulus is a control prime or different-scene prime, there is little or no match-
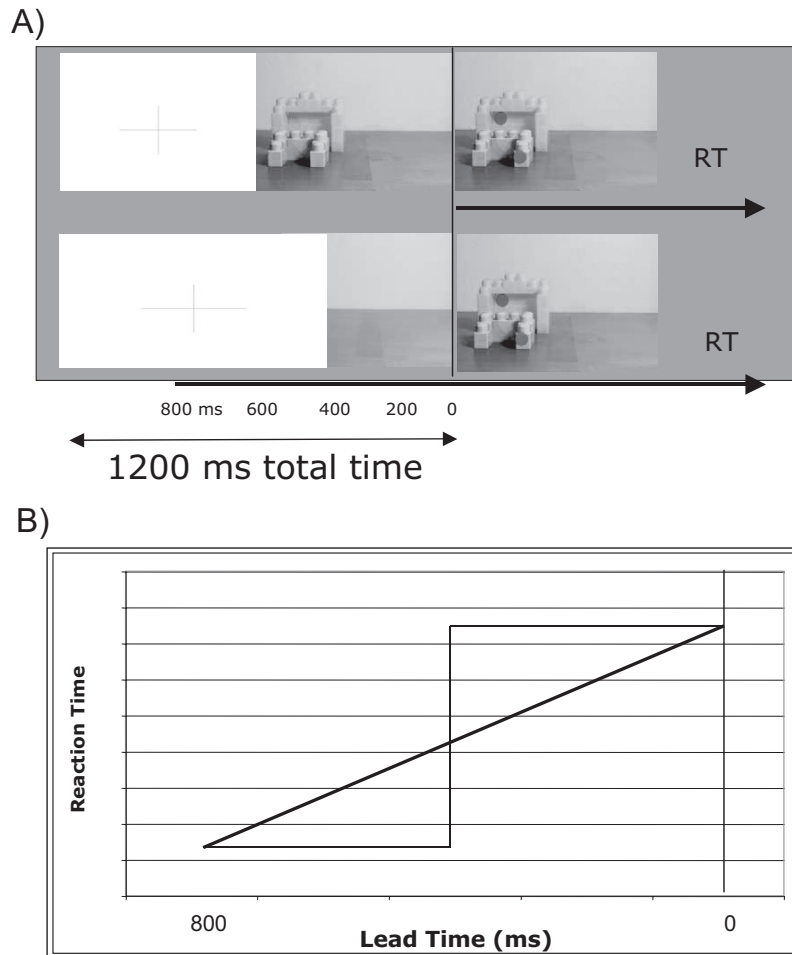
Figure 1. A: Sequence on scene prime and control prime trials (prime width analogous to time): fixation stimulus, prime (scene prime or control), and target until response. B: Possible time course functions. RT = reaction time.

ing information, and the onset of late processing is delayed until an appropriate intermediate representation of the target becomes sufficiently active—this produces the priming effect.

As the target is processed, the spatial probes attract attention because they are uniquely colored in the scene. In later processing, information from probe-relevant portions of the target is used to refine the intermediate representation, until there is sufficient resolution in the relevant regions. The duration of this phase will be affected by the difficulty of resolving the probes' locations and relation, among other factors. Thus, this later phase should vary with the difficulty of particular targets (probe pairs). The result of the late phase is a stable, locally complete representation that observers can respond to and become conscious of. This representation is achieved earlier on scene prime trials because of the head start provided by the intermediate representation of the prime.

This framework can be used to make predictions and inferences about phases of processing. Following additive factors logic, we can covary factors that would be expected to selectively influence early or late phases of processing, such as prime type and scene complexity. If the factors influence separate phases of processing, then the effects will be independent of each other and simply add

together. That is, the two factors would add a constant amount of time to each of two separate phases, and their combined effects would be additive. In contrast, if the factors influence the same phase of processing, then interactive effects may occur.

The Present Experiments

The main measure of priming effects was the contrast between scene primes that matched the target scene and a background (control) prime that contained the same ground and background surfaces as the target but no object information (Figure 1A shows both conditions). The background prime specified the scale and general reference frame of the target scene but no object-layout information. This information should be useful during target processing, but we expected that the object and object-layout information in the scene prime would provide benefits beyond those of the background prime.

The second independent variable was the complexity of the scene, defined in terms of number and complexity of objects present. This variable can be thought of as analogous to display size, which has been studied extensively (e.g., Enns & Rensink,

1991; Schneider & Shiffrin, 1977; Treisman & Gormican, 1988). The present scenes varied from one simple block object (1 part with 2 surfaces) to four complex block objects (14 or more parts in total and 42 surfaces; see Figure 2). Research indicates that both number of objects and complexity of the objects influences late processing (e.g., Alvarez & Cavanagh, 2004). Consistent with previous research and theory, we assumed that the early phase of processing would be efficient—generally parallel across the display and therefore minimally affected by scene complexity. In contrast, later phases of processing that resolve spatial relations should require additional time when the scene is more complex.

The following prediction can be made on the basis of additive factors logic. If scene priming influences early phases of processing, the priming effects should simply add to the effects of scene complexity, because complexity influences a distinct, later phase of processing. On the other hand, if scene primes influence later phases of processing, then interactive effects of prime and complexity may occur. Additional evidence from two additional factors will be brought to bear during the experiments.

General Method

Stimuli

Four color pictures varying in complexity served as the scenes and as scene primes (see Figure 2). Probe pairs were superimposed throughout the scenes to create 16 targets for each scene (e.g., Figure 3). We defined each probe pair by superimposing a red hue on two oval regions (e.g., see Figure 3C). The probe regions were at clearly different distances (from camera point) and differing lateral positions, and they were the only red regions in the scene. Targets were created in sets of four, one for each scene, with roughly similar probe positions within a set (one set is shown in

Figure 3A). This resulted in a total of 16 sets of targets for each scene; Figure 3B shows the 16 probe pairs used for one scene.

We defined probe regions by changing their hue in Adobe Photoshop to appear reddish against the background surface, while leaving texture and shading intact. The oval shape was varied depending on the depicted surface orientation so that the region shape looked like it was formed by a red flashlight pointed straight at (normal to) the scene surface (e.g., see Figure 3C). In half of the 16 targets, the left region was closer, and in the other half, the right region was closer. We used some individual probes in two targets but in different spatial relations (i.e., near once, far once) to discourage responses based on individual probes. In addition, some near probes were placed higher on the visual field than far probes, invalidating the depth cue of height in the visual field. In some targets, we had to modify luminance of the probe region(s) to obtain a red tone; however, in all cases the probe colors were kept consistent with scene lighting (no neon reds).

Procedure and Design

We ran participants individually through the experiment using a G4 Macintosh computer controlled by PsyScope (Cohen, Mac-Whinney, Flatt, & Provost, 1993). The stimuli were presented on a 17-in. (43.18-cm) cathode ray tube monitor in full color, and were 640×480 pixels (23.5×16.1 degrees of visual angle at the 60-cm viewing distance).

Each trial began with a gray cross (against white) that served as a fixation stimulus, followed by the prime. We held the total pretarget preparation time (fixation duration plus prime duration) constant by varying the fixation and prime durations inversely (e.g., Figure 1A shows two different pairs). The total preparation time was 1,200 ms in Experiments 1 and 3 (as in Figure 1). When the lead time interval elapsed, the target appeared within a refresh
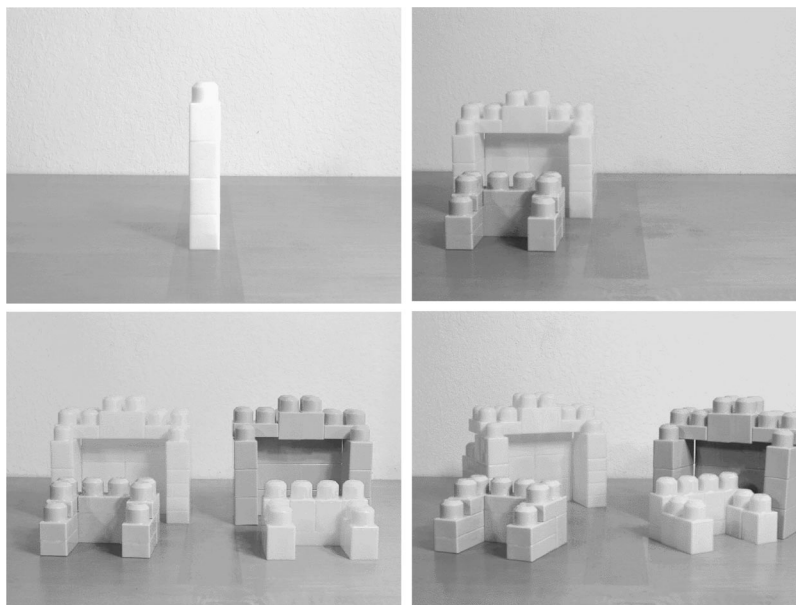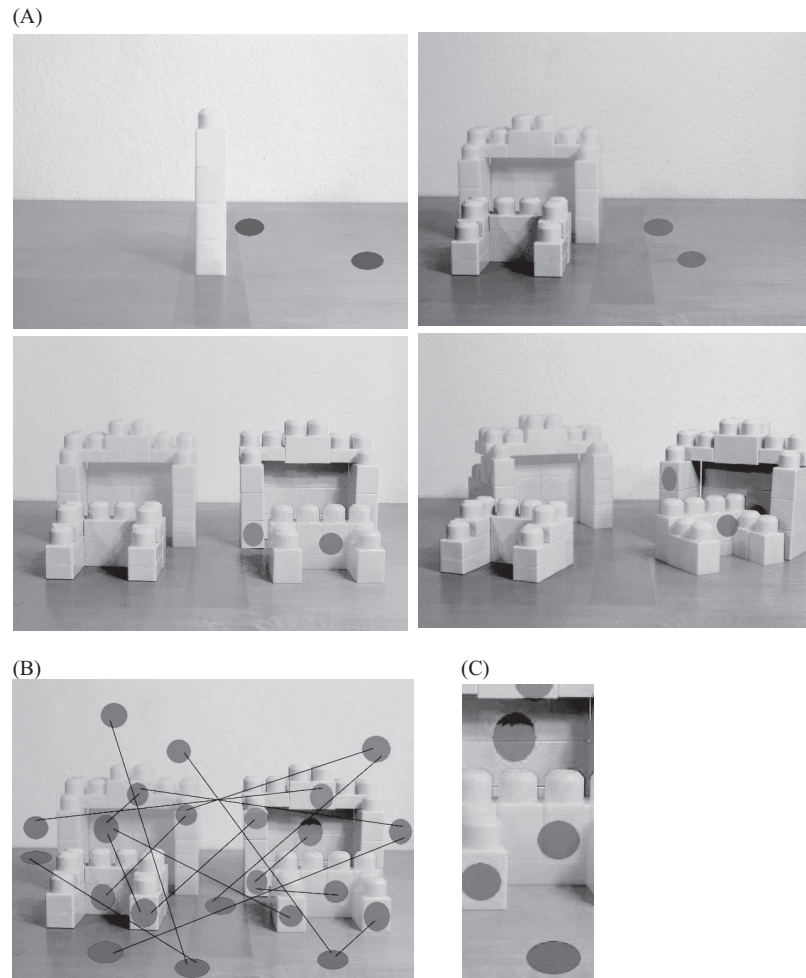

Figure 2. Simple scenes (top row) and complex scenes (bottom row) used in the experiments (Scenes S0, S2, S4, and S4O).

Figure 3. Target and probe examples. A: The four versions of a Target Number 13, one for each scene. B: For Scene S4, all of the target probes (connected by added lines). C: Close-up of some probes.

and remained visible until the observer pressed a left or right key. Note that abstract coding (left vs. right) was not necessary because probe position was mapped to key position. Observers were instructed to respond as to which probe was closer, as quickly and accurately as possible.

In each experiment, five lead time durations were crossed with two prime conditions, four scene conditions, and response (right or left closer). Scene primes were always the same scene as the target. These four factors created 80 combinations. On each trial, a combination was selected randomly, without replacement until all 80 combinations were used. The target number was selected randomly for each trial, without replacement. A particular target was repeated as many as nine times on average, during the practice and test portions of an experiment. The test trials were presented in blocks of 16 and initial practice trials in blocks of 8, with short breaks after each block.

Participants

The participants were students at the University of South Florida who volunteered for extra course credit. All participants reported normal or corrected vision. A proportion of participants has difficulty responding accurately in the present task, and analyses were restricted to participants with reasonably high accuracy ($>90\%$). Human subject protocols have not been developed for examining the reasons for the difficulty.

Experiment 1

Method

The five lead times were 800, 600, 400, 200, and 0 ms. There were 480 test trials in the experiment. They were preceded by a total of 92 practice trials. For the initial 32 practice trials, we used 1-s lead times to familiarize participants with the primes. The next 60 practice trials were the same as the test trials in composition. Twenty-eight participants contributed data (3 men); 6 additional participants failed to meet the accuracy criterion.

Results

The main analyses were conducted on mean reaction times for correct responses.

Lead time functions. The functions for the two prime types are shown in Figure 4, with the longest lead time beginning at the left. As can be seen, reaction time was lower with positive lead times, and the decrease was greater with the scene primes than with the background control prime. The overall difference in lead time functions was confirmed by analysis of variance (ANOVA), with prime type and lead time as factors; $F(4, 108) = 3.05$, $p = .02$, $MSE = 7,730$, $\eta_p^2 = .101$, for the interaction of prime and lead time.

Benefits for the primes were measured in two ways. First, we measured absolute benefits for both background and scene primes by comparing the average of the four positive lead times (800–200 ms) to the corresponding 0-ms lead time mean for that prime condition. For the background prime, the benefit of background information relative to no information averaged 24 ms. The overall benefit was reliable in a planned contrast, $t(27) = 2.73$, $p = .01$, $SE_{difference} = 9$ ms. For the scene primes, the absolute benefit averaged 52 ms across the four positive lead times and reached a maximum of 73 ms at the 200-ms lead time. The overall benefit was highly reliable, $t(27) = 5.27$, $p < .001$, $SE_{difference} = 10$ ms. The scene prime benefit reflects the combined contribution from background information and objects and their relations in the scenes.

The main measure of scene prime benefit was the relative benefit for the scene primes compared to the background prime. We also used this analysis to examine the time course of priming effects across the four positive lead times. The object and object-relational information in the scene primes produced responses that averaged 30 ms faster than with the background prime, $F(1, 27) = 26.26$, $p < .001$, $MSE = 7,600$, $\eta_p^2 = .493$, for the main effect. The results in the figure suggest that the advantage for the scene primes was especially strong at the 200-ms lead time (a 56-ms relative benefit). However, the interaction of lead time and prime type was only marginally reliable, $F(3, 81) = 2.16$, $p = .10$, $MSE = 9,096$, $\eta_p^2 = .074$. Nevertheless, note that benefits were largest at the 200-ms lead time; additional lead time did not increase the benefits of the scene representation. Therefore, the scene representations became functional rapidly, with a maximum benefit occurring at 200 ms.

The full set of benefit values is summarized in Table 1 for this experiment and subsequent experiments. In this table, the two 0-ms lead times have been averaged together.[1] This makes the
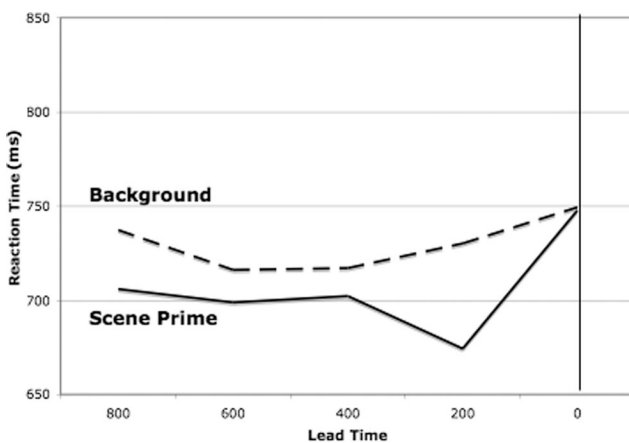

Figure 4. Lead time functions for background prime and scene primes in Experiment 1 (reaction time as a function of lead time).

Table 1

Benefits (in Milliseconds) in Each Experiment and Baseline Reaction Time (RT)

Experiment and scene	Background absolute	Scene relative	Scene absolute	Baseline RT (*SE*)
Experiment 1	23	30	53	749 (28)
S0 and S4Ob	28	41	70	
Experiment 2	13	26	39	821 (34)
S0 and S4Ob	8	26	34	
Experiment 3				
Inverted	40	64	103	1,114 (37)
Upright	49	46	95	968 (29)

Note. Also shown are benefits for the two scenes used in Experiment 3.

relations between the measures more clear: The absolute benefit for the scene prime (third column) is equal to the absolute benefit for the background prime plus the relative benefit for scene primes over the background prime.

Scene complexity. Total reaction times increased by 92 ms from the simplest to the most complex scene, $F(3, 81) = 41.30$, $p < .001$, $MSE = 12,820$, $\eta_p^2 = .605$. The scene means, from simplest to most complex, were 656, 720, 748, and 748 ms, respectively. The increase in reaction time can be explained by assuming that the later stages of spatial processing are more difficult with more complex scenes.

If scene priming and scene complexity influence different phases of processing, then the effects of scene primes should add to those of scene complexity. In the overall ANOVA, the variables were scene (complexity), prime type, and lead time (five levels). As reported, there was a main effect of scene complexity. However, complexity did not interact with prime type, $F(3, 81) = 1.30$, $p > .10$, $MSE = 7,864$, $\eta_p^2 = .046$, or lead time (all $ps > .10$). Thus, the benefits for scene primes did not differ systematically across different levels of scene complexity. Furthermore, the time course of priming was similar across scene complexity, as indicated by the lack of a Prime × Complexity × Lead Time interaction. This result can be illustrated by the lead time functions for scene primes, shown in Figure 5. In this figure, we grouped the two simplest and most complex scenes to reduce noise. As can be seen, there is a sizable difference in overall reaction times caused by complexity but no indication of a reduced or delayed priming process with more complex scenes. Thus, priming affected a phase of processing that was separate from and presumably prior to the phase of processing affected by complexity.

Target difficulty. To obtain converging evidence on phases of processing, we also analyzed target difficulty, which should have fairly specific effects. When the probe pairs were in more difficult positions (smaller depth separations or structurally complex interprobe relations), the resolution of target details should take longer. We assumed that this effect would occur primarily in late processing. The interesting question was as follows: How would the effects of target difficulty combine with the other main variables?

[1] Results at the two 0-ms lead time conditions, for scene primes and background primes, were slightly but not significantly different. The separate values were used in the statistical analyses but are combined in Table 1 for simplicity.
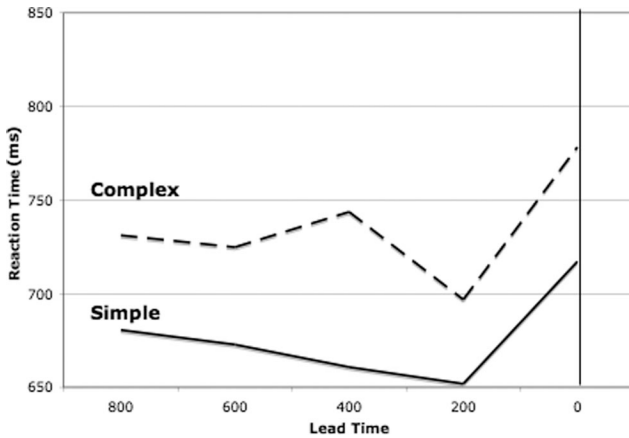
Figure 5. Lead time functions for simple and complex scene primes in Experiment 1.

As described in the *Method* section, targets were generated in sets coordinated across scenes—for each target number (from 1 to 16), the two probes were placed in roughly comparable positions across the four scenes (e.g., see Figure 3A). To measure target difficulty, we calculated the mean reaction time for each target number, from all of the data, collapsing across all other variables. Reaction time differed by 280 ms between the easiest and most difficult target numbers. We grouped the target numbers into quartiles, from fastest to slowest. We then took only the data from the four positive lead times (the times when primes were effective) and used an ANOVA to examine the joint effects of difficulty with prime type and scene complexity (two levels of complexity, high and low). If prime type influences early processes, then the effects should be additive with target difficulty. In contrast, the late factor of scene complexity should interact with the late factor of target difficulty.

The results are plotted in Figures 6A and 6B, with target difficulty along the abscissa. As would be expected, the effect of target difficulty was highly reliable, $F(3, 81) = 47.49, p < .001, \eta_p^2 = .638$. As can be seen in Figure 6A, the effects of prime type and target difficulty were generally additive; $F(3, 81) < 1, \eta_p^2 = .021$, for the interaction. This is further evidence that prime type influences relatively early processes, separate from later stages of processing the probes. There was an interaction of the two late variables, scene complexity and target difficulty (see Figure 6B), $F(3, 81) = 18.56, p < .001, \eta_p^2 = .407$. Reaction times were especially long for the most difficult probe pairs, when the scenes were most complex. This is converging evidence for the idea that complexity and target difficulty both influence relatively late stages of processing.

Accuracy. The percentage correct was above 95% for all 10 of the Lead Time × Prime conditions, as shown in Table 2. In an ANOVA parallel to the main reaction time analysis, there was a main effect of scene complexity, $F(3, 81) = 21.10, p < .001, MSE = 46.50, \eta_p^2 = .439$. Mean percentages for each scene were 98.5%, 95.8%, 94.2%, 97.2%, respectively, from simplest to most complex.

Discussion

The effects of lead time reveal a rapid time course with which scene representations become functional. The background information in the control prime provided some benefit, and more

benefits occurred with the additional object and object-location information in scene primes. For scene primes, the representations reached maximum effectiveness within the 200-ms lead time. Additional lead time reduced the benefits somewhat, if anything. At the longest lead time (800 ms), there is a slight upturn in the functions for both prime types. Observers might have become less attentive as lead time exceeded 600 ms. Alternatively, as suggested by a reviewer, the longer lead times may allow observers to begin developing a later representation that is selective and idiosyncratic; this representation may not match the complete target as well as an earlier prime representation.

The lead time functions were similar in shape with simple and complex scene primes, as shown in Figure 5; there were no effects of scene complexity on the overall size of the benefits or on the time courses with which benefits developed. These results are consistent with the idea that scene priming influenced early stages of processing and that prime processing was efficient, parallel across the visual field, and unaffected by scene complexity. Converging support for these conclusions comes from the additive effects of prime type and target difficulty.

Scene complexity had a strong effect on overall reaction time. However, the additivity of the complexity effect and the priming effects imply that complexity influenced a separate, and presumably late, phase of processing, such as the resolving of spatial relations. This same late stage of processing would be most likely
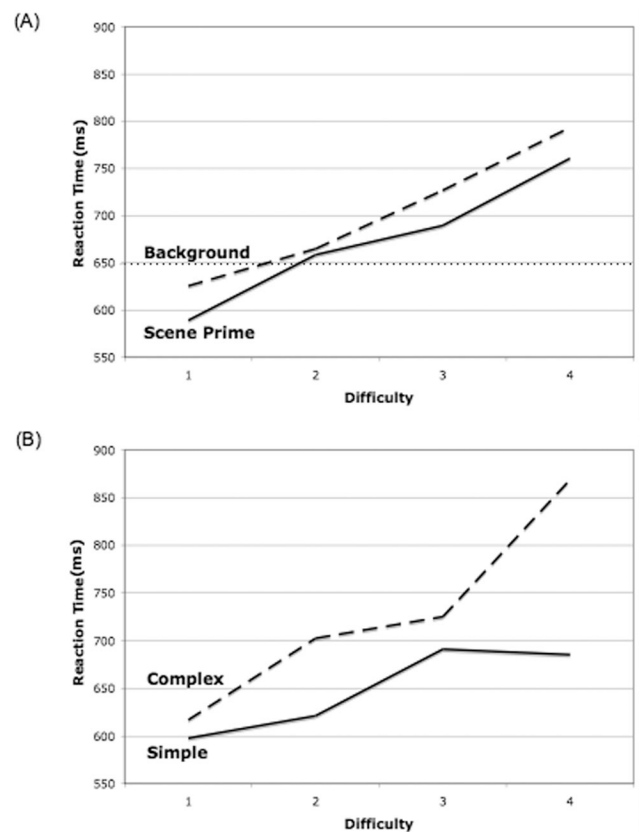

Figure 6. Effects of target difficulty with prime type (A) and scene complexity (B) in Experiment 1.

	Lead time				
Prime	800	600	400	200	0
Scene	96.9	95.9	96.2	96.6	96.6
Background	96.0	96.3	96.7	96.5	96.5

to be affected by target difficulty, and consistent with this idea, there was an interaction of these two factors.

Because the magnitude of priming effects was maximal at the shortest positive lead time, it is not clear when scene representations first become functional. Experiment 2 examined a shorter set of durations.

Experiment 2

Method

In order to focus more power at shorter lead times, we varied the five lead times in an approximately logarithmic manner: 0 ms, 50 ms, 150 ms, 300 ms, and 500 ms. We varied fixation duration from 900 ms to 400 ms, respectively, to produce a total of 900 ms before each target. There were 64 initial practice trials with 1-s lead times, followed by 480 test trials. Twenty-six participants participated (4 men; an additional 6 participants failed to maintain 90% accuracy).

Results

Lead time functions. The lead time functions for the two prime types are shown in Figure 7, plotted on the same scale as in Experiment 1. The overall pattern of reaction times was generally similar to Experiment 1, although the reaction times were somewhat longer overall, and benefits were smaller in magnitude. The interaction of prime type and lead time was reliable, $F(4, 100) = 2.98$, $p = .02$, $MSE = 7,098$, $\eta_p^2 = .106$.

The absolute benefit for the background prime relative to its 0-ms lead time was only 7 ms and not reliable ($t < 1$; $SE_{\text{difference}} = 10$ ms). For scene primes, the absolute benefit averaged 46 ms and

was reliable, $t(25) = 3.73$, $p = .001$, $SE_{\text{difference}} = 12$ ms. The benefit accrued quickly; the absolute benefit was sizable at the shortest positive lead time (46 ms).

The relative benefit for objects in the scenes primes, compared to the background prime, was 26 ms, $F(1, 25) = 13.91$, $p < .001$, $MSE = 9,842$, $\eta_p^2 = .357$. There was no interaction with lead time ($F < 1$). Nevertheless, to measure the reliability of the scene prime advantage at individual lead times, we conducted separate t tests. The advantage was reliable at the shortest lead time of 50 ms (27-ms relative benefit; $t[25] = 2.24$, $p = .03$) but only marginal at 150 ms (17-ms difference, $t[25] = 1.83$, $p = .08$). It was also reliable at the intermediate 300-ms lead time (42-ms benefit; $t[25] = 2.60$, $p = .02$) but not at the longest lead time of 500 ms (17 ms difference, $p > .20$). In short, there were modest benefits from the scene primes that were reliable overall, with individually reliable effects occurring at lead times as small as 50 ms.

Scene complexity. Total reaction time increased with scene complexity by 82 ms, $F(3, 75) = 63.27$, $p < .001$, $MSE = 8,763$, $\eta_p^2 = .717$. The scene means, from simplest to most complex, were 733, 810, 840, and 816 ms, respectively. However, complexity did not interact with prime or lead time ($Fs < 1$). The additivity of complexity and these factors is consistent with Experiment 1 and the idea that priming influences the early phase of processing, whereas complexity influences the later phases of processing. Lead time functions for the simple and complex scene primes are shown in Figure 8. As in Experiment 1, there is a sizable difference in overall reaction time but a similarly rapid priming process.

Target difficulty. We calculated overall reaction times for each target number and then divided target numbers into quartiles of varying difficulty. For the four positive lead times, the joint effects of target difficulty with prime type and scene complexity are shown in Figures 9A and 9B. The effects of prime type and difficulty were generally additive; $F(3, 75) = 1.49$, $p > .20$, $\eta_p^2 = .056$, for the interaction. Difficulty and complexity interacted, $F(3, 75) = 10.84$, $p < .001$, $\eta_p^2 = .302$. As can be seen, there was an overall greater increase in reaction times with difficult targets and greater scene complexity; however, the pattern is less consistent than in Experiment 1. The reason for the lack of complexity effect at the third level of difficulty is not clear at this time. Note that there tended to be less of a complexity effect in Experiment 1 at the third level of difficulty as well.
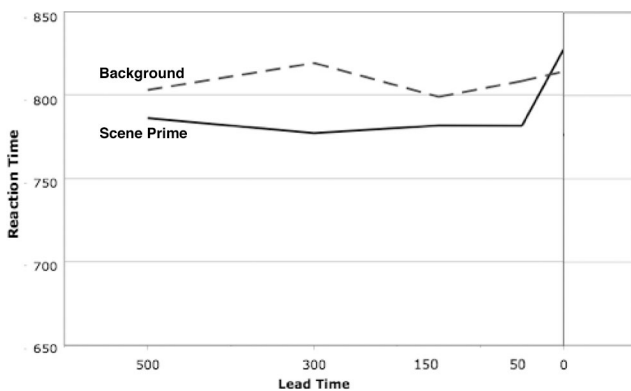

Figure 7. Lead time functions for background prime and scene primes in Experiment 2 (reaction time as a function of lead time).
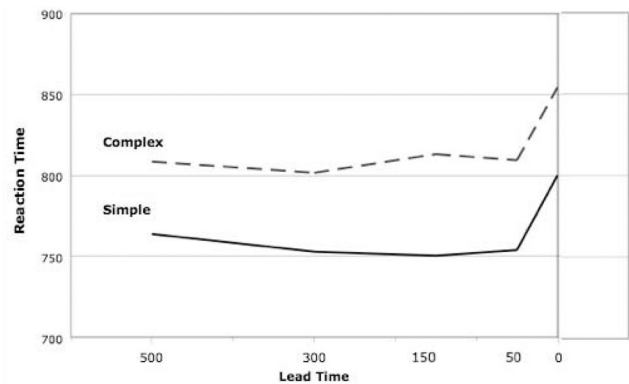

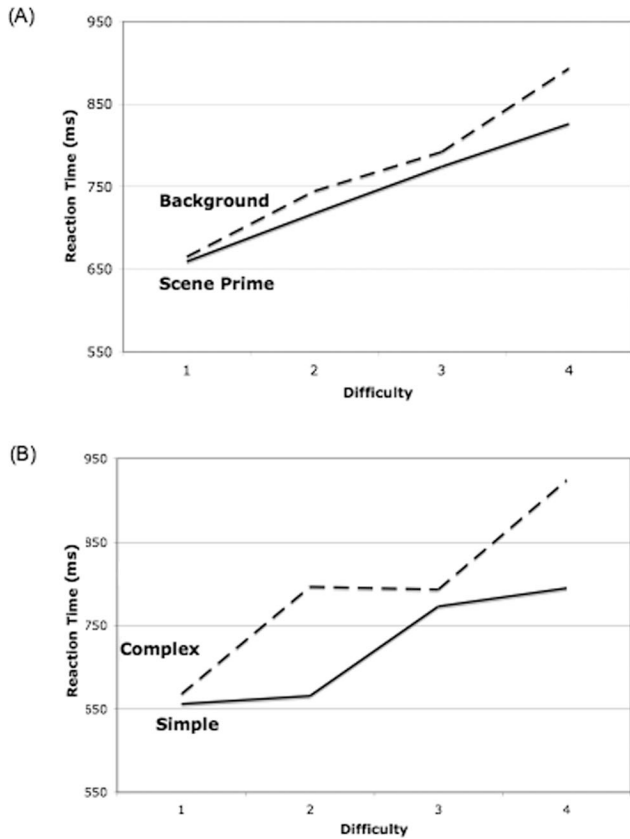*Figure 8.* Lead time functions for simple and complex scene primes in Experiment 2.

Figure 9. Effects of target difficulty with prime type (A) and scene complexity (B) in Experiment 2.

Accuracy. Accuracy was above 96% in all 10 of the main conditions, which are shown in Table 3. There was a main effect of scene complexity, $F(3, 75) = 7.98$, $p < .001$, $MSE = 36.15$, $\eta_p^2 = .591$; the scene means were 98.1%, 95.8%, 96.3%, and 97.6%, respectively.

Discussion

Experiment 2 indicates that scene representations can become functional very rapidly, with reliable benefits occurring at 50 ms. As in Experiment 1, there was no evidence that the benefits increased as lead time become greater.

However, note that all three types of benefits tended to be smaller in Experiment 2 than in Experiment 1 (see Table 1). None of the lead times approached the magnitude of the 200-ms benefits in Experiment 1 (73-ms absolute benefit, 56-ms relative benefit). We suspect that benefits were larger in Experiment 1 because the benefits vary with the utility of the prime information, and primes are more useful when observers have more time to process them. The average prime duration across test trials was twice as long in Experiment 1 (400 ms) as in Experiment 2 (200 ms). Thus, in Experiment 1 observers may have learned that they can see the primes well and used them more than observers in Experiment 2, producing larger benefits for both prime types. Perhaps as a result of greater use of prime information during the lead time interval, the observers processed the targets faster in Experiment 1

than in Experiment 2, $t(52) = 2.09$, $p = .04$, for the difference in overall reaction times.

In both experiments, the reaction time functions were fairly flat across the four positive lead times. (The marginal dip at 200 ms in Experiment 1 did not replicate either in Experiment 2 or in Experiment 3.) It appears that, in general, observers use the primes in the same way across trials varying in lead time. Perhaps observers set a strategy for using prime information based on prime utility, which determines the magnitude of the benefits throughout the experiment.

If this interpretation is correct, then there would be no specific point in time at which prime functionality begins. When most lead times are long enough, a maximum benefit can occur across lead times. When most lead times are short, the benefit is reduced across lead times. Given the results so far, it appears that a maximum benefit can occur with 200-ms lead times. Shorter lead times can also provide a benefit but apparently not as much as when observers have more exposure to the primes.

Further results from Experiment 2 were consistent with the idea that priming influences the early phase of processing, which is distinct from a later phase in which target details are resolved. The time course functions were similar for simple and complex scene primes (see Figure 8), while the effects of prime type were additive with scene complexity. The results involving target difficulty were generally consistent with these conclusions as well; target difficulty was generally additive with priming. There was some evidence of the predicted interaction between target difficulty and scene complexity, although there was increased noise in those results compared to Experiment 1.

Experiment 3

Experiments 1 and 2 converge on the conclusion that scene representations became functional rapidly, with maximal benefits requiring only 200 ms of prime presentation. There appears to be an initial, fast activation of a scene prime's representation that provides a head start on the processing of the target. The early, prime-facilitated phase of target processing appears to be distinct from later phases of target processing that resolve target details.

The purpose of Experiment 3 was to explore these time relations further, by adding another major variable, inversion of the entire scene. On half of the trials, both the prime and target were inverted (e.g., see Figure 10). Inversion increases the difficulty of spatial processing in general (e.g., Schone, 1980/1984) and prolongs processing in the present task (Sanocki et al., 2006). How might inversion influence the phases of processing in the present task?

One would expect that inversion would influence late phases of processing more so than early phases. The initial stage of process-

Table 3

Percentages Correct for Each of the Main Conditions of Experiment 2

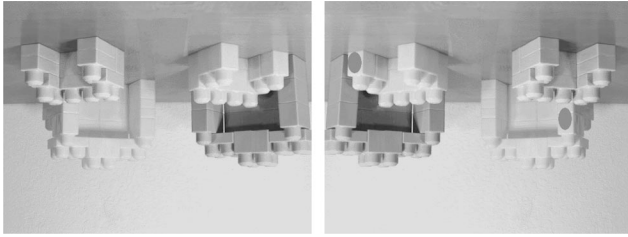| Prime	Lead time				
	500	300	150	50	0
Scene	96.5	97.2	97.3	96.9	96.2
Background	96.8	96.8	97.6	97.2	97.1

Figure 10. Inverted scene prime and target (examples) from Experiment 3.

ing in almost all models, feature extraction, is the efficient, parallel processing of relatively simple local features in primary visual cortex—features no more complex than some basic feature relations (e.g., Enns & Rensink, 1991; Treisman & Gormican, 1988). These features should be similar for upright and inverted scenes; therefore, inversion should have relatively little effect on this stage.

The lack of an inversion effect may extend to later stages of the early phase of processing, in which an intermediate representation develops. If so, then scene prime representations should become functional as quickly with inverted scenes as with upright scenes.

As processing continues, one can assume an increase in the processing of complex spatial relations. The relations may include attachments between objects and surfaces and a global reference frame based on the scene. Initially, coarse general relations may be processed in parallel, whereas eventually precise and complex relations may result from more capacity-limited processes. Inversion would seem more likely to prolong the later, capacity-limited processes than the early coarse processes. It is also possible than inversion would influence the strength of the detailed final representation (e.g., inversion may lower asymptotic activation levels) or reduce the certainty of the response, due to unfamiliarity. De Caro and Reeves have provided evidence that inversion causes an additional late checking process because of unfamiliarity (De Caro & Reeves, 2000). All of these effects would occur relatively late in processing.

Method

We used the wide range of lead times from Experiment 1 (800–0 ms) because of the possibility that inversion might increase the time needed for maximum prime-based preparation. We simplified the complexity manipulation to two scenes (S0 and S4O) and crossed inversion with the other factors. On half of the trials, both the prime and target were inverted. There were 64 initial practice trials with 1-s lead times, followed by 480 test trials. Twenty-eight participants participated (3 men), along with an additional 6 error-prone participants whose data were not analyzed.

Results

Lead time functions. The lead time functions for the two prime types and for inverted and upright scenes are shown in Figure 11. Overall reaction times were substantially longer in this experiment; the scale for Figure 11 begins where the scale for lead time figures ended in Experiments 1 and 2. In Experiment 3, inversion had substantial effects, increasing reaction times by an average of 146 ms over upright stimuli, $F(1, 27) = 55.73$, $MSE = 107,160.1$, $p < .001$, $\eta_p^2 = .674$. However, the overall patterns of effects were

similar at each orientation and qualitatively similar to the previous results. The interaction of prime type and lead time was reliable, $F(4, 108) = 3.75$, $MSE = 13,392$, $p < .01$, $\eta_p^2 = .122$, and this interaction did not vary with inversion ($F < 1$, $\eta_p^2 = .004$).

The absolute background benefit averaged 39 ms, $t(27) = 2.33$, $p = .03$, $SE_{difference} = 17$ ms. The benefits were 40 ms for upright scenes and 37 ms for inverted scenes. These benefits were larger in magnitude than in the previous experiments, as can be seen in Table 1. In the present experiment, the background primes provided valid information about the target's orientation.

The absolute scene prime benefit was 105 ms, $t(27) = 7.88$, $p < .001$, $SE_{difference} = 13$ ms (105 ms for upright scenes and 104 ms for inverted scenes). Thus, the scene primes produced substantial benefits in this experiment, as did the background primes.

The benefit of scenic objects relative to background primes was 55 ms across the four positive lead times $F(1, 27) = 44.31$, $p < .001$, $MSE = 15,325$, $\eta_p^2 = .621$. This effect was constant across lead time; $F(3, 81) < 1$ for the interaction of prime and lead time, $\eta_p^2 = .024$. As can be seen in Figure 11, the benefits were maximal at 200 ms, for both upright and inverted scenes. Thus, the scene prime information was used rapidly in this experiment. Given that background primes provided orientation information, the 55-ms relative benefit is additional, reflecting the effects of object configuration information. The benefit for scenic objects did not vary reliably with inversion, $F(1, 27) = 1.79$, $p > .10$, $\eta_p^2 = .062$, although the magnitude was slightly larger with inverted scenes (64 ms) than with upright scenes (46 ms). Given the substantial 146-ms effect of inversion, however, the 18-ms difference in benefits for upright and inverted scenes is relatively small. The finding that prime information was used as rapidly with inverted scenes as with upright scenes is strong evidence that the prime-
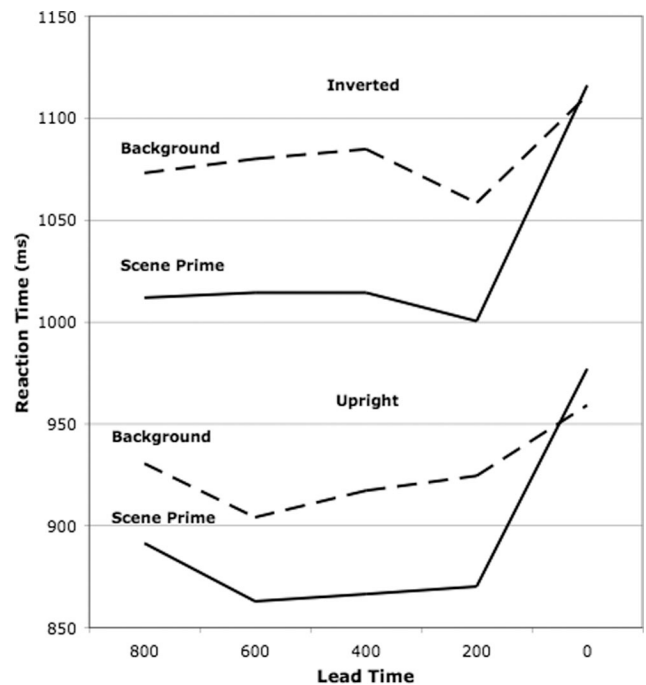

Figure 11. Lead time functions for each orientation and prime type in Experiment 3.

based facilitation process is fast and efficient, consistent with an early priming process.

Scene complexity. Experiment 3 involved two scenes, and the spatial relations took 77 ms longer overall to resolve with the more complex scene, $F(1, 27) = 45.82$, $p < .001$, $MSE = 1,674,395$, $\eta_p^2 = .630$. How did complexity combine with priming and inversion effects? There was no interaction of prime type and complexity, $F(1, 27) < 1$, $\eta_p^2 = .004$; the pattern of benefits was generally similar across levels of complexity. This can be seen in the lead time functions for the simple and complex scene primes, collapsed across orientation and shown in Figure 12. In contrast to the additive priming effects, there was an interaction of inversion and complexity, $F(1, 27) = 6.95$, $p = .01$, $\eta_p^2 = .205$. The more complex scene required 43 ms additional processing time when inverted relative to when upright (see Table 4). This later result is consistent with the idea of a late phase of spatial processing that is sensitive to complexity and compounded by inversion.

Target difficulty. We examined joint effects of difficulty with prime type, scene complexity, and also inversion. Overall reaction time was calculated for each target number, and then target numbers were grouped into quartiles by difficulty. The results for the four positive levels of lead time are shown in Figure 13. In this experiment, prime type interacted with difficulty (see Figure 13A), $F(3, 81) = 2.91$, $p = .04$, $\eta_p^2 = .097$. However, the interaction was fairly small in magnitude, and priming effects were smallest at the third level of difficulty (see Figure 13A). Difficulty also interacted with scene complexity (see Figure 13B), $F(3, 81) = 29.88$, $p < .001$, $\eta_p^2 = .525$. The complexity effect was somewhat greater overall with difficult targets, but the pattern was idiosyncratic and different from Experiments 1 and 2 (where the complexity effect was smallest for the third level of difficulty). Target difficulty and inversion had rather striking additive effects, as can be seen in Figure 13C; $F(3, 81) < 1$, $\eta_p^2 = .016$, for the interaction. We discuss these effects below.

Accuracy. Accuracy was above 92% in all 20 of the main conditions, which are shown in Table 5. There was a main effect of orientation, $F(1, 27) = 16.50$, $p < .001$, $MSE = 104$, $\eta_p^2 = .380$; accuracy was higher with upright scenes ($M = 96.4\%$) than with inverted scenes ($M = 93.9\%$).
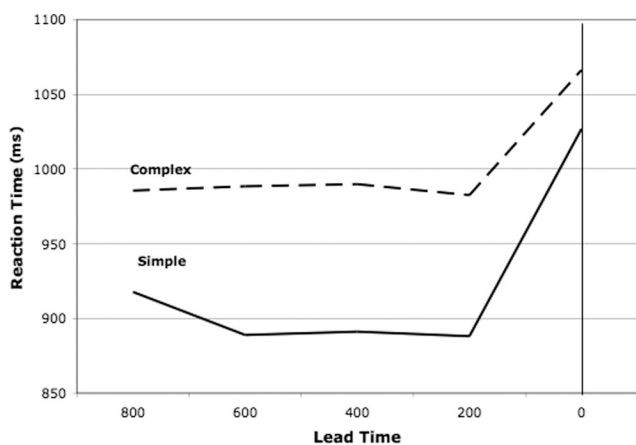
Table 4

Reaction Times (in Milliseconds) for the Simple and Complex Scene in Experiment 3, as a Function of Scene Orientation (Inversion)

| | Scene orientation | | | |
| | Upright | | Inverted | |
Measurement	Simple	Complex	Simple	Complex
Reaction time	883	938	1,007	1,106
Complexity effect (difference)		56		99

Discussion

Experiment 3 replicated and extended the main findings of Experiments 1 and 2. First, prime information was used rapidly; a maximum scene priming benefit occurred with 200 ms of prime duration.

Several further results support the conclusion that scene priming occurs primarily in early phases of processing. First, the time courses of scene priming were similar for upright and inverted scene primes (see Figure 12), as would be expected if priming is an early process that is unaffected by inversion. Second, the magnitudes of priming effects were similar for upright and inverted scenes, as would be expected if priming is an early process. In addition, the effects of priming were independent of the late factors of scene complexity, implying that they influence separate phases of processing.

There was a reliable interaction of priming and the late variable of target difficulty in this experiment, which could be inconsistent with our conclusions. However, the interaction was small in magnitude, and it was not consistent with what would seem to be the most likely interaction between prime and target difficulty—that scene priming would benefit difficult targets the most. The interaction is also inconsistent with the first two experiments.

In the present experiment, all three types of prime benefits were larger in magnitude than in the previous experiments, including the benefits for the background prime. This is consistent with our proposal that prime benefits increase with the utility of the prime information. In the present experiment, both prime types specified the orientation of the scene, which would be salient and useful information in the spatial relation task.[2]

For two reasons, we conclude that the variables of inversion and scene complexity influenced later stages of processing. First, effects on later stages seem more likely for both variables given prior research discussed earlier. Second, the interaction of inversion and complexity, in which reaction times were greatest for scenes that were inverted and complex, is consistent with the idea that these variables influence a common stage. However, the interaction of scene complexity and target difficulty produced some unpredicted variation. Unexpected variation also occurred in Experiment 2 but

Figure 12. Lead time functions for simple and complex scene primes in Experiment 3.

---

[2] A small complication should be noted—there were two scenes in Experiment 3 but four in Experiments 1 and 2. The means for those two scenes in Experiments 1 and 2 are also listed in Table 1 (Scenes S0 and S4Ob). When both Experiments 1 and 2 are considered together, the increase in benefits in Experiment 3, regardless of scene, is still clear.
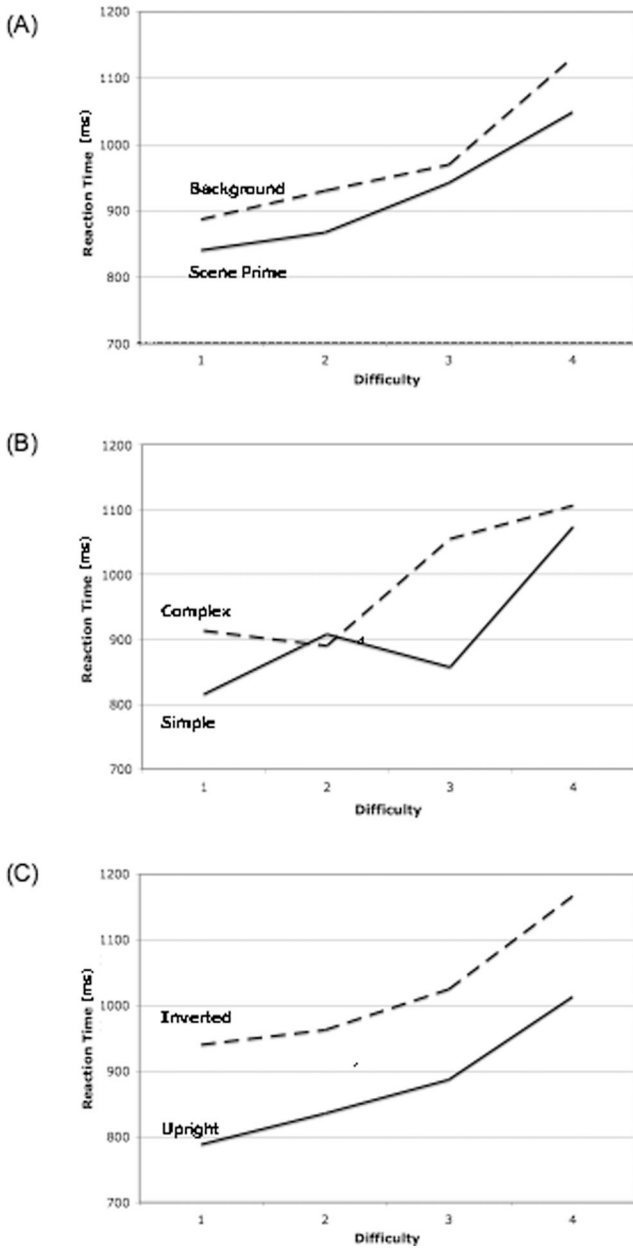
Figure 13. Effects of target difficulty with prime type (A), scene complexity (B), and scene orientation (C) in Experiment 3.

in a different pattern. This variation may come from interactions between probe positions and the structure of the scenes that were not fully controlled.

Even though target difficulty and inversion should both affect stage(s) in later processing, there was a striking additivity of target difficulty and inversion (see Figure 13B), suggesting that they influence separate stages. There may be various interpretations of this. However, one fairly straightforward explanation is that inversion causes a late checking process in reaction time studies (De Caro & Reeves, 2000). The unusual nature of inverted stimuli may decrease participants' confidence in their initial resolution of the probe relation, causing them to engage in an additional checking

process. In the present experiment, participants may initially process inverted targets in a manner similar to upright conditions but then engage in an additional, separate checking process. This process would add to reaction times. For this explanation to be complete, however, one would also have to assume that inversion has a second but earlier effect, in which inversion compounded the effects of complexity. We think this is a reasonable assumption. However, further research would be needed to test this explanation.

General Discussion

The present experiments help to establish scene priming as a potentially important aspect of everyday spatial perception. The experiments provide the first evidence that the layout priming process occurs rapidly for familiar scenes, with a maximum benefit requiring only 200 ms of exposure to a scene. Is this fast enough to influence everyday scene perception with familiar scenes? Given that scenic eye fixations typically last 250 ms or longer (e.g., Henderson & Hollingworth, 1997), there would be enough time during a fixation for an intermediate representation to be extracted. The representation should be functional during the next fixation or two. The speed of scene priming complements other evidence for the generality of scene prime benefits, including invariance across retinal position and lighting (Sanocki, 2003; Sanocki & Epstein, 1997; Sanocki et al., 2006). Further supporting the claim of generality is the fact that the scene primes facilitated processing across most of a typical view, as defined by the distribution of probe pairs across the scene (see Figure 3B).

The experiments indicate that the scene priming effect is likely to involve intermediate-level representations. The main evidence was the separation of priming effects from late influences of other variables (scene complexity and orientation). Scene primes appeared to provide a head start on processing by activating an intermediate-level representation, which allowed late processing to begin sooner. The priming process occurred before, and was additive with, late processing. Figures 5, 8, and 12 show priming effects for scene primes (only) and illustrate the early and late effects. In each experiment, there was similar early priming for simple and complex scene primes (i.e., flat reaction times across the positive lead times, below the 0-ms baseline). These effects were additive with large effects of complexity on late processing— the overall difference between the simple and complex scene functions in the figures. Later processing but not early priming was also affected by scene orientation in Experiment 3.

Table 5
Percentages Correct for Each of the Main Conditions of Experiment 3

Prime and orientation	Lead time				
	800	600	400	200	0
Upright scenes					
Scene	95.8	95.4	96.7	97.3	97.2
Background	96.4	97.2	96.6	95.5	95.5
Inverted scenes					
Scene	92.8	96	94.6	94.2	92.8
Background	94	93.4	93.7	93	94.2

Functional Intermediate-Level Representations

Previous theorizing about intermediate-level representations has emphasized their stimulus-driven nature and their generality (e.g., Marr, 1982; Rensink, 2000). Previous scene priming findings are consistent with generality: Priming is invariant across retinal position and lighting (Sanocki, 2003; Sanocki & Epstein, 1997). The present experiments provide evidence for efficient, stimulus-driven activation of intermediate representations. We measured the speed with which intermediate-level representations became maximally functional, and we found it to be fast, requiring only 200 ms of stimulus exposure. The similar rapid priming for simple and complex scenes suggests that the activation of the representation, up to the point of functionality, is parallel across the scene.

Another potentially interesting property of the functioning of layout representations is local autonomy (Sanocki et al., 2006). In those experiments, we cut scenes into as many as four pieces and then we switched the pieces between scenes, to create new composite scenes. The composite scenes are difficult to comprehend in any unified manner, although the four pieces were locally coherent within themselves. Scene priming effects were as strong for mixed scenes as for the original, coherent scenes. This suggests that local spatial configurations drive the scene priming effect; higher level structural and semantic coherence was not relevant. Local functioning, in which a region can provide facilitation independent of its relation to other regions, would seem to be a natural characteristic of intermediate-level representations.

Scene Familiarity, Speed of Priming, and the Priming Mechanism

To date in our scene priming research, the scenes have been made familiar to observers during practice trials. We have done so to ensure that representations of the scenes would exist in the observers' memories prior to presentation of the scene prime. This practice is representative of everyday perception throughout most of human history; even during travel, humans normally transitioned to a new environment gradually, with continuous changes of location and previews of upcoming layouts. Only through recent media inventions have humans acquired the opportunity to frequently encounter an entirely new environment all at once.

Nevertheless, the process of learning new layout representations is interesting. Previous research suggests that layout learning takes at most a few trials with normal scenes, and about 20 trials when two separate layouts must be integrated (Sanocki et al., 2006). The learning process is an interesting topic for future research.

The power of repetitions to create primeable representations implies a potential explanation of the speed of scene priming, in terms of retrieval of prior episodes with the scene. The influence of memory on processing is now a well-researched phenomenon (e.g., Bowers & Marsolek, 2003; Logan, 1988), including memory for layouts in visual search (Chun & Jiang, 1998). However, if one accepts the conclusion that priming was an intermediate-level function, then the retrieved memories would have to function at the intermediate level. Intermediate functioning would be more general across minor changes in scenes with time. This possibility is worthy of further evaluation.

An alternate possibility would be that the retrieved memories are a later stage representation. However, late representations

should contain some details, including information about the additional objects and relations that define complex scenes. Such information could be expected to reduce the burden of processing the complex scenes. It did not; the amount of priming and its speed were similar with simple and complex primes.

There is also an explanation of the speed of priming that does not require memory retrieval—rapid bottom-up processing. The scene representation might be activated by stimulus information anew on each trial, and the process might be so quick because only a coarse, intermediate-level representation is created from the prime. Research on the rapid categorization of novel natural scenes suggests that bottom-up processing is highly rapid (e.g., Kirchner & Thorpe, 2006; Rousselet et al., 2002, 2005). This explanation should also be considered in future research.

Prime Utility and Prime Benefits

The magnitudes of the priming benefits varied from experiment to experiment, in a manner that can be explained by prime utility—the ease with which observers see the prime, combined with the informativeness of prime information. We argued that prime benefits are larger when the prime durations are long enough for useful information to be extracted from the prime (Experiments 1 and 3 vs. 2), and larger when the prime provides more spatially relevant information (e.g., the orientation of the scene in Experiment 3). However, if the representation that causes benefits is intermediate level and produced by rapid, efficient processes, could its effectiveness vary from experiment to experiment? Perhaps there is learning during the experimental session, in which observers learn to set a temporal attentional window, depending on the availability and usefulness of the prime information. When primes provide highly useful information and the prime durations are long enough, observers make prime processing a high priority. In contrast, if prime durations are very short, observers may devote less effort to selecting prime information from the other events in the stimulus presentation sequence. This explanation entails that the intermediate representation of the prime is not an automatic copy of stimulus information but an entity that varies with the observer's experience and temporal control settings.

Can Scene Primes Influence Later Levels of Processing?

Although intermediate-level scene priming was supported by the results, it is possible scene primes could influence later processes in certain situations. Note that the design of the present experiments may have encouraged observers to use the scene primes in a general, wide-scope manner—each prime was followed by 16 different targets with probe pairs in a variety of locations (e.g., see Figure 3B). Scene prime processing in the present case may be representative of undirected perception, when the observer must be ready for multiple directions of subsequent processing and action. Contrasting situations exist, such as when perception and action are directed at a specific goal repeatedly. For example, a scene prime could consistently predict a particular spatial relation. In such a situation, the scene prime may also provide information about that particular spatial relation, as well as response-related information or other components of a late representation.

However, there was some evidence of another kind of late influence in the present lead time functions—the tendency for reaction time to rise at the longest lead time (800 ms in Experiment 1 [see Figure 4] and 800 ms for upright scenes in Experiment 3 [Figure 11]). The overall interactions were not reliable, but there is an interesting interpretation suggested by a reviewer: Perhaps observers have enough time at 800 ms (with upright scenes) to develop a selective, idiosyncratic late representation. This representation may not integrate with the wide-scope target as well as a less developed intermediate representation, reducing the priming benefit. The data from Experiment 3 suggest that the late representation developed soon enough with upright scenes to reduce performance but not with inverted scenes, as would be expected for a late process.

Layout: Grist for Scene Gist?

Recent evidence suggests that scene layout may be fundamental in tasks that involve scene meaning as well. Torralba, Oliva, Castelano, and Henderson (2006) outlined how global properties of scenes, which include major aspects of layout, can be used to categorize scenes and help guide visual search within the layout. Michod and Intraub (2007) found that the conceptual masking that occurs when scenes follow in rapid succession can be caused by changes in layout only. And as noted earlier, observers can use low- and high-pass images that obscure objects but preserve layout to categorize scenes (Oliva & Schyns, 1997; Schyns & Oliva, 1994). Intermediate-level representations of layout could serve as a foundation for these semantic processes.

References

Alvarez, G., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15,* 106–111.

Biederman, I. (1972, July 7). Perceiving real-world scenes. *Science, 177,* 77–80.

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.

Bowers, J. S., & Marsolek, C. J. (2003). *Rethinking implicit memory.* Oxford, England: Oxford University Press.

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

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25,* 257–271.

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein & S. Rogers (Eds.), *Handbook of perception and cognition: Vol. 5. Perception of space and motion* (pp. 1–46). San Diego, CA: Academic Press.

De Caro, S., & Reeves, A. (2000). Rotating objects to determine orientation, not identity: Evidence from a backward-masking/dual-task procedure. *Perception & Psychophysics, 62,* 1356–1366.

Domini, F., Caudek, C., & Tassinari, H. (2006). Stereo and motion information are not independently processed by the visual system. *Vision Research, 46,* 1707–1723.

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

Evans, K., & Treisman, A. (2005). Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance, 31,* 1476–1492.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Henderson, J. M., & Anes, M. D. (1994). Roles of object-file preview and type priming in visual identification within and across eye fixations. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 826–839.

Henderson, J. M., & Hollingworth, A. (1997). Eye movements during scene viewing: An overview. In G. Underwood (Ed.), *Eye guidance while reading and while watching dynamic scenes.* New York: Elsevier.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24,* 175–219.

Kirchner, H., & Thorpe, S. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research, 46,* 1762–1776.

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

Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology, 28,* 103–174.

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

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 86,* 287–330.

McConkie, G. W., & Rayner, K. (1976). Identifying the span of the effective stimulus in reading: Literature review and theories of reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 137–162). Newark, NJ: International Reading Association.

Michod, K. O., & Intraub, H. (2007). Conceptual masking: Is it really all about the concept or does layout matter? [Abstract]. *Journal of Vision, 7*(9), 191a.

Miller, J. (1988). Discrete and continuous models of human information processing: Theoretical distinctions and empirical results. *Acta Psychologica, 67,* 191–257.

Neisser, U. (1967). *Cognitive psychology.* East Norwalk, CT: Appleton-Century-Crofts.

Ni, R., Braunstein, M., & Andersen, G. (2005). Distance perception from motion parallax and ground contact. *Visual Cognition, 12,* 1235–1254.

Oliva, A., & Schyns, P. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology, 34,* 72–107.

Potter, M. (1975, March 14). Meaning in visual search. *Science, 187,* 965–966.

Potter, M. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 509–522.

Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition, 7,* 17–42.

Rousselet, G., Fabre-Thorpe, M., & Thorpe, S. (2002). Parallel processing in high-level categorization of natural images. *Nature Neuroscience, 5,* 629–630.

Rousselet, G., Joubert, O., & Fabre-Thorpe, M. (2005). How long to get to the 'gist' of real-world natural scenes?. *Visual Cognition, 12,* 852–877.

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.

Sanocki, T. T., Michelet, K., Sellers, E., & Reynolds, J. (2006). Functional representations of spatial layout can consist of independent pieces. *Perception & Psychophysics, 68,* 415–427.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human

information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1–66.

Schone, H. (1984). *Spatial orientation* (C. Strausfeld, Trans.). Princeton, NJ: Princeton University Press. (Original work publisher 1980).

Schyns, P., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science, 5,* 195–200.

Sedgwick, H. A. (1986). Space perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 1–57). New York: Wiley.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, Amsterdam, 30,* 276–315.

Sternberg, S. (1998). Discovering mental processing stages: The method of additive factors. In D. Scarborough & S. Sternberg (Eds.), *Methods, models, and conceptual issues: An invitation to cognitive science* (Vol. 4, pp. 703–863). Cambridge, MA: MIT Press.

Torralba, A., Oliva, A., Castelhano, M., & Henderson, J. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113,* 766–786.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Treisman, A., Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95,* 15–48.

Wolfe, J., & Bennett, S. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research, 37,* 25–43.