Stankiewicz, B.J., Hummel, J.E. & Cooper, E.E. (1998). The role of attention in priming for left-right reflections of object images: Evidence for a dual representation of object shape. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 732-744.

Running Head: ATTENTION AND INVARIANCE WITH LEFT-RIGHT REFLECTION

The Role of Attention in Priming for Left-Right Reflections of Object Images: Evidence for a Dual Representation of Object Shape

Brian J. Stankiewicz, John E. Hummel
University of California, Los Angeles
Eric E. Cooper
Iowa State University

Address correspondence to:
Brian J. Stankiewicz
UCLA
Department of Psychology
Att: Grad Mail Room
Los Angeles, CA 90095-1563
BSTANKIE@PSYCH.UCLA.EDU

Abstract

Three experiments investigated the role of visual attention in priming for object images and their left-right reflections. Objects to which subjects attended were visually primed in both the same view and in the left-right reflected view; ignored objects were primed only in the same view. The effects of attention (attended vs. ignored) and view (same vs. reflected) were strictly additive. These results suggest that two separate representations mediate human object recognition (Hummel & Stankiewicz, 1996): One requires attention but is invariant with left-right reflection, while the other can be activated automatically but is sensitive to left-right reflection. Both representations appear to be invariant with translation across the visual field.

The Role of Attention in Priming for Left-Right Reflections of Object Images: Evidence for a Dual Representation of Object Shape

The human visual system recognizes objects with remarkable speed and accuracy. A fraction of a second after an image falls on the retina, we know what the object is. This is true even if the object is presented in a novel viewpoint or is a new member of a familiar category. Although we are extremely proficient at recognizing objects, we are limited in the number of objects we can recognize simultaneously (see Biederman, Blickle, Teitelbaum, & Klatsky, 1988; Ruthruff & Miller, 1995; Treisman & Gelade, 1980). This capacity limitation suggests that attention plays a critical role in object recognition. What is the role of attention in object recognition? Visual attention is known to influence feature selection (LaBerge & Brown, 1986,1989), processing speed (Hikosaka, Miyauchi, & Shimojo,1993a, 1993b), and visual binding (e.g., feature binding [Treisman & Gelade, 1980; Treisman & Schmidt, 1982; Wolfe, Cave, & Franzel, 1989] and part-relation binding [Enns, 1992; Logan, 1994]), and each of these is likely to affect object recognition. However, the role of attention in binding is especially suggestive. As elaborated shortly, some theories of object recognition predict that different kinds of binding will result in *qualitatively* different representations of object shape (Hummel & Stankiewicz, 1996). We present a series of experiments exploring the properties of the visual representations generated in response to attended and ignored object images.

Object constancy and visual attention

Our capacity to derive unchanging interpretations from highly variable images is typically referred to as *object constancy*. Human object constancy takes two forms. The best known is robustness to variations in viewpoint: Any given object can be viewed in an infinity of perspectives, each projecting a different image to the retina, but our ability to recognize an object is largely unaffected by these variations. Using a priming paradigm, Biederman and his colleagues have shown that this capacity to ignore image variations due to viewpoint at least partially reflects the perceptual representation of object shape (Biederman & Cooper, 1991a; Biederman & Cooper, 1992; Biederman & Gerhardstein, 1993; 1995; but see Tarr & Bülthoff, 1995). The second form of object constancy is our ability to classify novel instances of familiar object classes: Entering a furniture store, we immediately recognize the chairs as chairs and the tables as tables, even if we have never seen exactly those chairs or tables before. This capacity underscores the importance of representation in human object constancy: Our visual system represents object shape in a form that is robust—not only to variations in viewpoint—but also to variations in an object's actual three-dimensional shape.

Theories of object recognition fall into two primary classes: *normalization-based* theories and *structural description* theories. Both classes of theories predict that attention will play an important role in object recognition. According to normalization-based theories, objects are represented in a viewpoint-specific format (such as two-dimensional views). Objects are recognized in novel viewpoints by means of normalization procedures—such as alignment, rotation, or view interpolation—that bring viewed images into correspondence with stored views (e.g., Edelman & Weinshall, 1992; Poggio & Edelman, 1990; Olshausen, Anderson & Van Essen, 1993; Seibert & Waxman, 1993; Tarr & Pinker, 1989, 1990; Ullman & Basri, 1991; Vetter, Poggio, & Bülthoff, 1994; see Tarr, 1995, for a review). In these models, attention would control the processes that perform the normalization. For example, according to the normalization-based model of Olshausen, Anderson and Van Essen (1993), attention serves to map viewpoint-specific images into a view-like coordinate space that is invariant with size and translation. Once mapped into this space, the represented view is invariant with respect to the location and size of the original retinal image.[1]

Structural description theories differ from normalization-based theories in that they emphasize the role of representation (rather than normalization processes) in object constancy. These theories assume that objects are represented in a format that is invariant with some (but not necessarily all) variations in viewpoint (e.g., Marr & Nishihara, 1978; Biederman, 1987; Dickenson, Pentland & Rosenfeld, 1993; Hummel & Biederman, 1990; 1992; Hummel & Stankiewicz, 1996; Sutherland, 1979; Winston, 1975; see Quinlan, 1991 for a review). In these models, recognition is invariant with viewpoint to the extent that the representation on which it is based is invariant with viewpoint. Attention enables the generation of a view-invariant structural description from an object's image. For example, the models of Hummel and Biederman (1992) and Hummel and Stankiewicz (1996) represent objects as structural descriptions specifying the interrelations among their volumetric parts (geons). Importantly, the parts and relations are represented independently, making it necessary to actively—or *dynamically*-—bind parts to their relations (see Hummel & Biederman, 1992). There is strong evidence that binding independent visual attributes requires visual attention (Enns, 1992; Logan, 1994; see also Treisman & Gelade 1980; Wolfe et. al., 1989), suggesting that attention may be necessary for generating a structural descriptions (Hummel & Biederman, 1992). Accordingly, these models predict that the visual system will generate a structural description in response to an attended image, but not in response to an ignored image; as a consequence, the representation of an attended image will be more robust to variations in viewpoint than the representation of an ignored image. Thus, regardless of one's perspective on how the human visual system achieves object constancy (e.g., by normalization or structural description), there are reasons to hypothesize that object constancy will require visual attention.

Importantly, this is not to say that attention is necessary for object recognition (see Tipper, 1985; Tipper & Driver 1988; Treisman & DeShepper, 1995). Rather, the hypothesis is that the representations mediating recognition will *differ* as a function of whether an object is attended or ignored: The visual representations activated in response to an attended object will be more robust to variations in viewpoint than that activated in response to an ignored object. We tested this hypothesis by exploring the role of attention in priming for left-right reflections of object images.

Visual priming and invariance with left-right reflection

Using a priming paradigm, Biederman and Cooper (1991a) showed that some component of the visual representation mediating object recognition is invariant with left-right reflection. Participants viewed and named line drawings of common objects in two blocks of trials (*prime* block and *probe* block). Images displayed during the probe block belonged to one of three different conditions: An image was either identical to an image the participant had named during the prime block (*identical*), a left-right reflection of an image named in the prime block (*reflected*), or an image of an object the participant had not named during the first block (*unprimed*).[2] The dependent measure of primary interest was priming for images and left-right reflections of images participants had named during the prime block (operationalized as the difference between response times to name identical or reflected images and response times to name unprimed images). Biederman and Cooper found that object images visually primed their left-right reflections just as much as they visually primed themselves: That is, visual priming was completely invariant with left-right reflection. This result suggests that the visual representation mediating object recognition is invariant with left-right reflection (Biederman & Cooper, 1991a). (Or, as elaborated shortly, a *component of* this representation is invariant with reflection.) The models of Hummel and his colleagues account for this invariance in terms of the parts and relations that comprise the structural description of an object image (see Hummel & Biederman, 1992). On the assumption that these representations are activated only in response to attended images (due to the need for dynamic part-relation binding), these models predict that attended images will prime their left-right reflections, but ignored images will not (Hummel & Stankiewicz, 1996)[3].

The basic experimental paradigm was as follows. Trials were grouped into prime/probe pairs (Figure 1). On a *prime* trial, two line drawings of common objects were displayed on a computer screen. One object was precued. The participant's task was to name the cued image (the *attended prime*); they were not required to respond to the other (the *ignored prime*). Each prime display was followed by a *probe* display, during which a single image (a *probe*) appeared on the screen. The participant's task was again to name the image. The probe image depicted either the same object as the attended prime, the same object as the ignored prime, or an object the subject had not previously seen in the experiment (an *unprimed* probe). In Experiments 1 and 3, probe images (except for unprimed probes) were either identical to the corresponding primes, or were left-right reflections of them. Priming is operationalized as the difference between naming response times to primed and unprimed probe images. In Experiments 1 and 3 attended prime images reliably primed both themselves and their left-right reflections. However, ignored prime images only primed themselves. Experiment 3 also replicated Biederman and Cooper's (1991a) finding by measuring long-term (rather than short-term) priming. Experiment 2 served as a control for name and concept priming, and showed that a component of the priming observed in Experiments 1 and 3 was specifically visual. In this experiment, identical probes were replaced with same name, different exemplar (*SNDE*) probes--probes that had the same names as the corresponding primes, but were different exemplars with different 3D shapes (e.g., if the prime was a jet [name: "airplane"], the SNDE probe would be a Cessna [name: "airplane"]).

Insert Figure 1 about here

Experiment 1

Experiment 1 served as a basic test of whether attention plays a role in visual priming for left-right reflections of object images. Participants fixated the center of a computer screen while two line drawings of common objects were displayed simultaneously on the left and right sides of the screen. Participants were instructed to name, as quickly and accurately as possible, the object that appeared within the cueing square. The cueing square appeared on the screen just before the images and remained on the screen until the images were replaced with a pattern mask. Approximately 3 seconds after the prime display, the participant viewed a probe display consisting of a single image in the center of the screen. Again, the task was to name the object as quickly and as accurately as possible. The probe was either the same as the prime they had just named (an *attended* probe), the same as the prime they had just ignored (an *ignored* probe), or an image of an object the subject had not previously seen in the experiment. Probes were presented either in the same view as the corresponding primes (*identical* probes), or were left-right reflections of those primes (*reflected* probes). The dependent measure of interest is the effect of attention on priming for identical and left-right reflected probes.

Method

Participants. 28 native English speakers with normal or corrected-to-normal vision participated for credit in introductory psychology courses at the University of California, Los Angeles.

Materials. Line drawings of 56 asymmetrical objects were displayed on a Macintosh color monitor connected to a Macintosh IIVX computer. A program written in MacProbe (Aristometrics) controlled the experiment. Response times were collected using a dynamic trigger microphone attached to a MacProbe Computer Interface Box (Aristometrics). Subjects sat approximately 100 cm from the display.

Procedure: Each object was placed into one of seven conditions (Attended-Identical, Attended-Reflected, Attended-Not-probed[4], Ignored-Identical, Ignored-Reflected, Ignored-Not-probed, and Unprimed). All objects appeared in all seven conditions equally often. The ordering of the trials was randomized for each participant, as was the pairing of attended and ignored objects on prime trials. Prime orientation (left facing vs. right facing) was counterbalanced across participants.

The experimenter read instructions to the participant, after which the participant paraphrased the instructions back to the experimenter. The participant then received 24 practice trials. After the practice phase, the computer displayed “End of Practice”, and the participant was asked if he/she had any questions.

Prime displays began with an unfilled circle (subtending 0.032° of visual angle) in the center of the screen. This circle remained on the screen until the participant initiated the trial with a keypress. After the keypress, the circle was replaced with a filled fixation dot, which remained on the screen for 495 ms. A white screen was then displayed for 30 ms, followed by an attentional cueing square (4.57° on a side) presented either left or right of the fixation dot. The center of the cueing box was 4.0° from the center of the screen. After 75 ms, two line drawings were displayed on the computer screen for 120 ms. One appeared inside the square (the attended image) and the other appeared on the other side of the screen (the ignored image). Both images were centered 4.0° from the center of the screen (fixation). Following a 30 ms blank screen, a random-line pattern mask was displayed on the screen for 495 ms. The mask covered the entire screen (subtending 15.6° of visual angle). To prevent participants from moving their eyes to the fixation square or either object, the entire prime display lasted less than 200 ms, a duration too brief to permit a saccade. The participant’s task was to say the name of the cued (attended) object as quickly and as accurately as possible into the microphone. Response times were recorded by the computer through a voice key attached to a microphone.

Following the prime display, the screen was left blank for 1995 ms. This was followed by a fixation dot (0.032°) displayed for 495 ms. Following a 30 ms blank screen, the probe image was displayed in the center of the screen for 150 ms. The probe was either the attended image, the left-right reflection of the attended image, the ignored image, the left-right reflection of the ignored image, or an unprimed image.[5] A total of 3015 ms elapsed between the end of the prime display and the beginning of the probe display (495 ms for the prime mask, 1995 ms for the blank screen, 495 ms for the probe fixation dot, and 30 ms blank). Following the probe, a single pattern mask (4.57°) appeared in the center of the screen for 495 ms. Participants were instructed to name the probe as quickly and as accurately as possible. After the participant named the probe, the computer displayed the names of the attended prime and the probe along with the probe response time. Participants were allowed to view this feedback for as long as they wished, and indicated they were finished with a keypress. At the end of each trial, the experimenter used the keyboard to record the participant’s accuracy on the prime and probe displays. The experimenter also recorded all trigger errors (i.e., when the voice key did not trigger, or triggered before the subject actually responded).

Results: Figure 2 summarizes the priming results of Experiment 1. Trials on which either the prime or probe responses were incorrect and trials on which probe response times exceeded 3 seconds were omitted from these calculations (4.15%). In all conditions, priming was calculated as the difference between each participant's mean response time in the unprimed (baseline) condition and the participant’s mean response times in each of the other probe conditions.

Insert Figure 2 about here

A 2 (attended vs. ignored) x 2 (identical vs. reflected) within participants analysis of variance (ANOVA) revealed a reliable main effect of attention ($\underline{F}(1,27)=50.33$, $\underline{p}<.001$) and view ($F(1,27)=7.93$, $p<.01$), but the interaction between attention and view did not approach reliability ($F(1,27)=0.23$)

Matched pairs t-tests were conducted on each priming condition to determine which conditions demonstrated priming (i.e., faster naming responses relative to unprimed probes). Priming was reliably greater than zero in the attended-reflected ($t(27)=7.35$, $p<.001$), attended-identical ($t(27)=9.69$, $p<.001$), and ignored-identical conditions ($t(27)=2.14$, $p<.05$), but not in the ignored-reflected condition ($t(27)=0.13$, $p>.05$). Attended images were primed in both the same and reflected views, but ignored images were primed only in the same view. The difference between the attended-identical and attended-reflected conditions is statistically reliable ($t(27)=2.96$, $p<.01$), as is the difference between the ignored-identical and ignored-reflected conditions ($t(27)=2.11$, $p<.05$). Identical probes enjoyed a similar advantage in priming over reflected probes in the attended and ignored conditions (~50ms).

Discussion

Experiment 1 revealed more than 180 ms of priming for left-right reflections of attended images, but no priming for left-right reflections of ignored images. However, both attended and ignored images were primed in the same view. Like negative priming for ignored images (Tipper, 1985; Tipper & Driver, 1988; Treisman & DeShepper, 1996), this result suggests that object recognition may take place in the absence of attention. However, ignored objects were primed only in the identical view, suggesting that the representation mediating recognition without attention is sensitive to left-right reflection. Perhaps most interestingly, the data revealed approximately 50 ms more priming for identical probes than for reflected probes in both the attended and ignored conditions. That is, the effects of attention (attended vs. ignored prime) and view (identical vs. reflected probe) were strictly additive. This pattern suggests that two separaterepresentations mediate object recognition: one is invariant with left-right reflection but requires attention, and the second is sensitive to left-right reflection but does not require attention.

Another interesting aspect of these findings is that the ignored images were primed in the same view even though the prime and probe images appeared in different parts of the visual field. (Recall that prime images were presented either left or right of fixation whereas probe images were presented at fixation, and that the prime displays were too brief to permit eye movements.) Even for ignored images, priming obtains despite translation across the visual field. This suggests that the same-view priming for ignored images does not simply reflect priming in an early representation of local (i.e., retinotopic) image features. It also suggests that the visual system may achieve invariance with translation even without attention[6]. (If attention were required to correct for translation, then we would have observed no priming for the ignored images in either the same or reflected views.) This result is inconsistent with Olshausen et. al's (1993) hypothesis that attention serves to correct for the location of an image in the visual field. Other implications of this finding, including its relation to previous findings (e.g., Ellis, Allport, Humphreys, & Collis, 1989), are elaborated in the General Discussion.

An important limitation of Experiment 1 is that attention was perfectly confounded with naming: Not only did subjects attend to the cued images and ignore the uncued images, but they also named the cued images and did not name the uncued images. It is possible, therefore, that all the priming we observed for left-right reflections in the attended condition is simply name and/or concept priming. Perhaps the visual representations that mediate object recognition are completely viewpoint-specific, and only the name and concept are invariant with left-right reflection. Experiment 2 was designed test this possibility. In particular, it was designed to provide an index of how much of the priming for attended objects is name or concept priming, and how much is specifically visual priming.

Experiment 1 showed that the attended primed representation is at least partially sensitive to left-right reflection. This finding is inconsistent with Biederman and Cooper’s

(1991a) finding that visual priming is invariant with left-right reflection. An important difference between Experiment 1 and Biederman and Cooper’s studies is in the duration between the prime and probe displays. Experiment 1 used a short-term priming paradigm in which probe displays followed each prime display with an inter-display time of about three seconds. Biederman and Cooper used a long-term priming paradigm in which prime and probe displays were separated on the order of five to ten minutes. Experiment 3 investigated the effect of prime-probe inter-display time on viewpoint-sensitive versus viewpoint-invariant priming.

Experiment 2

To measure non-visual priming (and, by subtraction, estimate visual priming) this experiment replaced the identical-image condition of Experiment 1 with a same-name different exemplar (SNDE) condition (following Biederman & Cooper, 1991a). Here, probes were either unprimed images (*unprimed* probes, as in Experiment 1), left-right reflections of prime images (*reflected* probes, as in Experiment 1), or objects that shared the name and concept—but not the same 3D shape—with a primed image (*SNDE* probes, instead of the *identical* probes in Experiment 1). For example, if a left-facing grand piano (name: "piano") served as a prime image, then a right-facing grand piano would serve as the reflected probe, an upright piano (name: "piano") would serve as the SNDE probe, and the image of some other object (e.g., a jet, name: "airplane") would serve as the unprimed (baseline) probe. If some of the priming for reflections of attended objects observed in Experiment 1 was specifically visual, then this experiment will reveal more priming for reflected probes than for SNDE probes. But if all the priming for left-right reflections observed in Experiment 1 was simply name or concept priming, then the reflected and SNDE probes in this experiment will show equal amounts of priming (see Biederman & Cooper, 1991a).

Methods

Participants. 40 native English speakers with normal or corrected-to-normal vision participated for credit in introductory psychology courses at the University of California, Los Angeles.

Materials: The experiment used two sets of objects: a set of SNDE pairs and a set of filler objects (objects with no SNDE counterpart). The conditions in Experiment 2 were identical to those of Experiment 1 except that the identical image condition was replaced by the SNDE condition and the “unprobed” conditions were replaced with filler objects. In the SNDE condition, each prime image was paired with an image of a different object with the same name (rather than being paired with itself, as in the identical image condition of Experiment 1). There were 35 pairs of SNDE objects and 42 filler objects. Each SNDE object appeared equally often in each of the five probe conditions (i.e., Attended-Reflected, Attended-SNDE, Ignored-Reflected, Ignored-SNDE, and Unprimed). The filler objects were used as unprobed objects (e.g., if the attended prime was probed, then the ignored image was a filler).

Procedure: The procedure was identical to that of Experiment 1 except that the identical condition was replaced with the SNDE condition.

Results: Priming data are shown in Figure 3. Trials in which the prime or probe responses were incorrect or response times exceeded 2 seconds were omitted from the analysis (11.28 %)[7]. The remaining data were subjected to a 2 (attended vs. ignored) X 2 (SNDE vs. reflected) within participants ANOVA. There was a reliable main effect of attention (F(1,39)=78.04, p<.001), and a main effect of object type, (F(1,39)=4.12, p<0.05). The interaction between attention and object type was marginally reliable (F(1,39)=3.59. p=.065). Post-hoc matched-pairs t-tests revealed that there was reliably more priming for reflected probes than SNDE probes in the attended condition (t(39)=4.84, p<.01) but not in the ignored condition (t(39)=0.20). On average, attended images primed their left-right reflections 65 ms more than they primed their SNDE counterparts. Ignored images primed neither their reflections nor their SNDE counterparts.

Each priming condition was subjected to a matched-pairs t-test to determine whether it revealed non-zero priming (i.e., decreased naming response times relative to unprimed probes). Priming was reliably greater than zero in both the attended-SNDE ($\underline{t}(39)=4.98$, $\underline{p}<.001$) and attended-reflected conditions ($\underline{t}(39)=7.56$, $\underline{p}<.001$), but it was not reliably greater than zero in either the ignored-SNDE ($\underline{t}(39)=0.43$) or ignored-reflected conditions ($\underline{t}(39)=0.81$).

Insert Figure 3 about here

Discussion

The purpose of this research is to explore the role of attention in *visual* priming for left-right reflections of object images. Experiment 1 showed that attended images prime their left-right reflections, and Experiment 2 showed that at least a component of this priming (approximately 85 ms) is specifically visual. This estimate of visual priming is arguably conservative, as it is premised on the assumption that an object and its SNDE counterpart share no visual attributes (specifically, they do not share the same object model; see Cooper, Biederman, & Hummel, 1992). To the extent that an object and its SNDE counterpart share visual attributes, then priming for SNDE probes likely overestimates the magnitude of non-visual priming; and to the extent that SNDE priming overestimates non-visual priming, the difference between SNDE and reflected priming underestimates visual priming. These results thus provide further evidence that attention plays an important role in building visual representations that are invariant with left-right reflection.

This experiment revealed no SNDE priming for ignored images, suggesting that priming for ignored images is specific, not only to viewpoint (left- vs. right-facing), but also to an object's particular 3D shape. In turn, this result suggest that attention may play a role in object constancy broadly defined, rather than simply in generalization over viewpoint. In contrast to the findings of Tipper (1985), this result revealed no evidence for higher-level (e.g., conceptual or semantic) priming of ignored objects.

One limitation of Experiments 1 and 2 is that the confounding of attention and naming makes it difficult to determine whether it was attention per se or simply naming that caused the priming of the viewpoint invariant shape representation: the results of Experiments 1 and 2 are consistent with both accounts. However, previous studies suggest that naming is not necessary to prime a left-right invariant representation of shape. For example, Srinivas (1996) used a "facing" judgment (i.e., "Is the object facing to the left or right?") during the prime phase and during the probe phase subjects either named the objects or made the facing judgment. In both experiments she found left-right reflection-invariant priming. Because Srinivas' subjects did not name the objects during the prime phase, these studies suggest that naming is not necessary for priming a left-right invariant representation of object shape.

Experiment 3

The purpose of Experiment 3 was twofold. First, it was designed to test whether the effects observed in Experiments 1 and 2 would replicate under a different attention manipulation: Is attending vs. ignoring, per se, responsible for the observed differences in priming, or is it the case that the primes' differing positions in the visual field were also instrumental in the observed effects? This question is especially important given that the effects of attention selection depend, in part, on the ease of selection (Ruthruff & Miller, 1995). Selecting objects on the basis of location in the visual field is arguably easier than selecting overlapping objects on the basis of, say, color (e.g., Tipper, 1985). Perhaps this difference explains our failure to replicate Tipper's finding of high-level priming for ignored images. The second purpose of Experiment 3 was to help us understand why our findings differ from those of Biederman and Cooper (1991a). Biederman and Cooper found complete invariance with left-right reflection in visual priming—that is, images primed their reflections just as much as they primed themselves. By contrast, our results showed an advantage for same-view priming over reflected-view priming, even in the attended condition. One plausible explanation for this difference is that the time between our prime and probe displays was on the order of three seconds, whereas several minutes elapsed between prime and probe displays in Biederman and Cooper's experiment. It is possible that the same-view priming effect we observed is short lived, and that, under longer prime-probe delays, we (like Biederman and Cooper) will observe complete invariance in priming for left-right reflections. This seems especially plausible given that, Ellis and Allport (1985; Ellis et. al., 1989) found evidence for short-lived viewpoint-sensitive representations of object shape.

The prime displays in Experiment 3 consisted of two superimposed objects, one red and one green. Participants selected objects for naming, not based on location (attended and ignored images were in the same location), but based on color. (Half the participants named red objects, and half named green ones.) To test the role of prime-probe delay, we also added a long-term probe condition to this experiment. Objects that were not immediately probed (i.e., on the following trial) were probed in a series of trials at the end of the experiment. If our failure to observe complete invariance with left-right reflection in Experiments 1 and 2 reflects the short prime-probe delays in those experiments, then we should observe complete invariance in priming with the delayed probes. This explanation of the distinction between out findings and those of Biederman and Cooper thus predicts a three-way interaction between attention, left-right reflection, and prime-probe delay.

Method

Participants. 36 native English speakers with normal or corrected-to-normal vision participated for credit in introductory psychology courses at the University of California, Los Angeles.

Materials: The materials were the same as those used in Experiment 1 with the following exceptions: The prime displays consisted of a red object superimposed over a green object on a black background. Probe objects were drawn in white on a black background. Attended and ignored primes were paired as in Experiment 1.

Procedure: The procedure was identical to that of Experiment 1 except that participants selected objects for naming according to color. 18 participants named red objects and 18 named green objects.

Another difference between Experiments 1 and 3 was the use of long-term and immediate probes in Experiment 3. The long-term probe displays occurred in a single block after the prime and immediate probe trials. The long-term probes consisted of the primed and unprimed images that were not probed in the immediate probe condition. For example, if the prime display consisted of an attended jet airplane and an ignored piano, and the immediate probe was a jet airplane, then the piano served as one of the long-term probe images. The display sequence for the long-term probes was identical to the display sequence for the immediate-probe displays. The time between primes and long-term probes averaged about seven minutes.

Results: Figure 4 shows the pattern of priming observed in the immediate probe condition. Trials with incorrect prime or probe responses and trials with probe response times over 3 seconds were omitted from the analysis (12.32%).

Insert Figure 4 about here

A 2 (attended vs. ignored) X 2 (identical vs. reflected) within subjects ANOVA revealed a reliable main effect of attention, ($\underline{F}(1,35)=88.92$, $\underline{p}<.001$) and view, ($\underline{F}(1,35)=9.40$, p<.01). As in Experiment 1, the interaction between attention and view did not approach statistical reliability (F(1,35)=.10).

Each priming condition was subjected to a matched-pairs t-test to determine whether it revealed statistically reliable priming. As before, all conditions revealed reliable priming (attended-LR, t(35)=7.83, p<.001; attended-ID, t(35)=10.02, p<.001; ignored-ID, t(35)=1.98, p=0.056) except the ignored-reflected condition (t(35)=0.58, p>0.05). Priming was again reliably greater in the attended-identical condition than in the attended-reflected condition (t(35)=2.38, p<.05), and reliably greater in the ignored-identical condition than in the ignored-reflected condition (t(35)=2.11, p<.05). This pattern of results is identical to that observed in Experiment 1, even though attended and ignored primes were designated by color rather than location.

Figure 5 summarizes the long-term priming results. Trials on which prime or probe responses were incorrect and trials on which probe response times exceeded 3 seconds were omitted from the analysis (7.32%).

Insert Figure 5 about here

A 2 (attended vs. ignored) X 2 (identical vs. reflected) within subjects ANOVA revealed a reliable main effect of attention (F(1,35)=43.37, p<.001) but not of view, (F(1,35)=0.04). The interaction between attention and view was not statistically reliable (F(1,35)=1.00).

Each priming condition was subjected to a matched-pairs t-test to determine whether it revealed statistically reliable priming. Priming was observed in the attended-reflected (t(35)=4.84, p<.001) and attended-identical (t(35)=5.28, p<.001) conditions only (ignored-identical: t(35)=1.70, p>.05; the ignored-reflected: t(35)=1.04, p>.05). Matched pairs t-tests revealed no reliable differences between priming for identical and reflected images, in either the attended condition (t(35)=1.04, p>0.05) or the ignored condition (t(35)=0.39, p>0.05). These data--in particular, the lack of difference between identical and reflected probes under delayed priming--replicate the complete invariance with left-right reflection observed by Biederman and Cooper.

A 2 (attended vs. ignored) X 2 (identical vs. reflected) X 2 (probe delay: short vs long) within sugjects ANOVA revealed a main effect of attention (F(1,35)=111.22,p<.01), a main effect of view (F(1,35)=6.63, p<.05) and a main effect of probe delay (F(1,35)=13.68,p<.01). As predicted, the view by delay interaction was also statistically reliable (F(1,35)=8.65,p<.05). Furthermore, the attention by delay interaction was also statistically reliable (F(1,35)=8.65,p<.01). However, the attention by view interaction was not statistically reliable (F(1,35)=.05). The three-way interaction was also not statistically reliable (F(1,35)=.60).

Discussion

The immediate probe condition of Experiment 3 replicated Experiment 1 with objects selected by color rather than location. Like the results of Experiments 1 and 2, these results provide further support for the notion that object recognition is mediated by two separate representations, one that requires attention, but is invariant with left-right reflection, and one that is sensitive to left-right reflections but does not require attention. The fact Experiment that 3 replicated Experiment 1 so cleanly suggests that the effect speaks to the role of attention, per se, rather than the role of specific cues (such as location or color) whereby attentional selection may operate.

The long-term probe condition replicated Biederman and Cooper's (1991a) finding of complete invariance with left-right reflection under long prime-probe delays (several minutes). This three-way interaction (between attention, left-right reflection, and prime-probe delay) suggests that the viewpoint-specific priming observed in both the attended and ignored conditions of Experiments 1 and the immediate probe condition of this experiment is only short-lived. By contrast, the component of the priming that is invariant with left-right reflection is longer-lasting.

General Discussion

Attention plays an important role in visual perception. Among other things, it serves to select elements in a scene (LaBerge & Brown, 1986, 1989; Wolfe et. al., 1989), speed the processing of selected elements (Hikosaka, Miyauchi, & Shimojo, 1993a; 1993b), and bind visual attributes into groups (Enns, 1992; Logan, 1994; Treisman & Gelade, 1980). Theoretical work in object recognition suggests that attention should also play an important role in generating object representations for recognition, either by normalizing view-like representations (Olshausen et. al., 1993), or by enabling the generation of viewpoint-invariant structural descriptions (Hummel & Biederman, 1991, 1992; Hummel & Stankiewicz, 1996). Consistent with these predictions, the findings reported here suggest that attention affects the *qualitative* properties of the visual representation of object shape.

In three experiments, participants viewed pairs of object images, one of which they attended (and named), and the other they ignored. Both attended and ignored images were visually primed in the same view (Experiments 1 and 3). This effect obtained despite the fact that prime images were viewed off fixation, whereas probe imaged were viewed at fixation. That is, priming, even for ignored images, is robust to translation across the visual field (see also Biederman & Cooper, 1991a). This finding is important because it shows that the observed priming reflects representations and/or processes other than simple local image features. It also suggests that attention may not be needed to activate a representation that is invariant with translation across the visual field. Although attention appears to play little role in discounting location in the visual field, it does enable the visual system to represent object shape in a manner that is invariant with left-right reflection. (Attended objects were primed in both the same view and the left-right reflected view, whereas ignored objects were primed only in the same view.). Experiment 2 showed that a substantial component of the priming for left-right reflections is specifically visual, rather than just name or concept priming.

The visual priming we observed has two separate components. One is short lived (lasting at least three seconds, but less than five minutes; Experiment 3), relatively viewpoint-sensitive (robust to translation [Experiments 1 and 2], but affected by reflection [Experiments 1-3]), and independent of attention (Experiments 1-3). The other is longer lasting (lasting at least five minutes; Experiment 3), invariant with both translation and reflection, and dependent on attention (it obtains for attended but not ignored images; Experiments 1, 2, and 3). The effects of viewpoint and attention were additive in the short prime-probe delay conditions of Experiments 1 and 3: Attended objects enjoyed the same amount of short lived same-view priming as ignored objects (about 50 ms).

Additive Effects of Attention and Viewpoint

Attention (attended vs. ignored) and viewpoint (same vs. reflected) had additive effects on priming in the short prime-probe delay conditions of Experiments 1 and 3. This additivity is important in the context of the theoretical interpretation of the current findings. Most generally, it suggests that the automatic, view-sensitive representation of object shape is relatively independent of the attention-consuming reflection-invariant representation (Hummel & Stankiewicz, 1996). However, there are at least two specific versions of this interpretation of the additivity.

The account that follows most naturally from the usual interpretation of additive effects is that the data reflect processes operating on independent parts of the system (Sternberg, 1969). For example, in the context of the current findings, perhaps the representation that gets primed without attention (the automatic representation) resides in an early part of the processing stream, while the representation that gets primed only with attention (the controlled representation) resides at a later part of the stream. Object memory resides at the end of the stream. If we assume that priming in the early, automatic representation is short-lived and view-sensitive (i.e., because the representation, itself, is view-sensitive), whereas priming in the later, controlled representation is longer-lived and reflection-invariant (i.e., because the representation is invariant with reflection), then this model can account for all the findings reported here. When an image is attended, both representations will be primed. If that image is then immediately presented again (probed) in the same view, then the system will profit from the priming in both representations: It will be faster (by, say 50 ms) to activate the automatic representation, and it will also be faster (by, say, an additional 100 ms) to activate the controlled representation. The total priming effect will be 150 ms. If the image were probed in the left-right reflected view, then the system would profit only from the priming in the controlled representation (because the "wrong" automatic representation has been primed), and the total priming effect will be 100 ms. Next consider what will happen to an ignored image. Here, only the automatic representation will be primed (again by 50 ms). Probing immediately with the same image will result in 50 ms of total priming, but probing with a left-right reflection will result in no priming. The effects of attention and viewpoint would be additive: no priming for the reflected image without attention; 50 ms of priming for the same view without attention, 100 for attention without the same view, and 150 ms for same view plus attention. Moreover, at long prime-probe delays, priming will be completely invariant with left-right reflection because all the (short-lived) automatic viewpoint-sensitive priming will have disappeared over the delay.

Although this serial model accounts for the findings reported here, it is limited in that it cannot account for the findings that motivated the experiments. Specifically, on the serial model, the automatic representation does not make any direct contact with object memory. (Rather, it only makes contact with the controlled representation, which requires attention to make contact with memory.) The serial model therefore predicts no recognition in the absence of visual attention, a prediction that is inconsistent with the findings of Tipper (1985), and Tipper and Driver (1988).

An alternative account of the additivity is suggested by a model recently proposed by Hummel and Stankiewicz (1996). This model differs from the serial model in that it assumes that both the automatic and controlled representations make direct contact with object memory. This "parallel" model accounts for the finding that recognition does not require attention, but to a first approximation, it also seems to predict a sub-additive interaction between attention and viewpoint (rather than the observed additivity). However, the parallel model only predicts sub-additivity if we assume that the automatic representation is primed just as much without attention as it is with attention. (If we do not make this assumption, then the "automatic" representation is perhaps better termed the "semi-automatic" representation.) In fact, according to the Hummel and Stankiewicz model, even the semi-automatic representation profits from visual attention. According to this model, the interaction between attention and viewpoint could be sub-additive, additive, or even super-additive, depending on the proportion of the semi-automatic representation that gets primed without attention. The smaller the fraction that gets primed without attention, the more the interaction should tend toward super-additivity; the greater the fraction, the more the interaction should tend toward sub-additivity; additivity lies somewhere in the middle. In its current state, the model predicts an interaction in the additive to super-additive range. In all other respects, this model accounts for the current findings in the same way as the serial model. For example, the parallel model -- like the serial model -- predicts complete invariance at long prime-probe delays, assuming that the semi-automatic priming is short-lived.

Attention and Viewpoint in the Hummel and Stankiewicz Model

The reason for the super-additivity prediction of the Hummel and Stankiewicz model is subtle, but it is worth considering because it relates closely to the reasoning behind the prediction that attention should be necessary to prime the left-right reflection of an object's image. Readers who are uninterested in these details can proceed to the next section. The Hummel and Stankiewicz model is motivated by two seemingly contradictory sets of findings in the object recognition literature. One set of findings suggests that human object recognition is mediated, at least in part, by the activation of a structural description specifying an object’s parts and their interrelations (for reviews, see Cooper et. al., 1992; Hummel & Stankiewicz, 1996; Quinlan, 1991); the other suggests that we recognize objects both automatically and rapidly—arguably faster than it is theoretically possible to generate a structural description from an object’s image (due to the need for dynamic binding; see Hummel & Stankiewicz, 1996). The Hummel and Stankiewicz model resolves this dilemma by recognizing objects on the basis of two independent (but integrated) representations of object shape.

One representation, the Independent Geon Array (IGA), is a collection of units that represent the shape attributes of an object's parts (geons) and their interrelations. In the IGA, every attribute is represented independently of every other (i.e., each attribute has its own unit). This independence gives the IGA considerable flexibility, and permits it to discount some variations in viewpoint (such as left-right reflection) as a natural consequence (Hummel & Biederman, 1992; Hummel & Stankiewicz, 1996). The cost of the independence is that it requires a mechanism for actively (dynamically) binding attributes into collections corresponding parts (Hummel & Biederman, 1992). If dynamic binding requires attention (as discussed in the Introduction), then visual attention should be required to generate a useful representation on the IGA. Taken in isolation, the IGA (like the Hummel & Biederman, 1992, model) predicts that object recognition should require visual attention.

The second representation, the Substructure Matrix (SSM), is a collection of units that represent geon attributes, not independently, but at each of several locations in a coarse coordinate system. (The coordinate system is semi-object-centered. It is invariant with translation and scale, but it sensitive to rotation and left-right reflection.) The SSM (in contrast to the IGA) does not rely on dynamic binding to preserve the grouping of shape attributes into geons. This function is performed by the geons' differing positions in the coordinate space. As such, the SSM (but not the IGA) can represent an object's shape in a manner suitable for recognition, even when the resources necessary for dynamic binding (i.e., attention and processing time) are not available. However, the SSM is sensitive to left-right reflection: The representation of a given object facing to the left may be completely different from the representation of the same object facing to the right.

If recognition without attention is mediated by the SSM alone, then priming without attention should be sensitive to left-right reflection (consistent with the findings reported here). If recognition with attention is mediated by both the SSM and the IGA, then priming with attention should be partly invariant with left-right reflection (also consistent by the findings reported here). And if priming in the SSM is short-lived, then priming over long prime-probe delays should be completely invariant with left-right reflection (as observed by Biederman and Cooper, 1991, and in Experiment 3). This model predicts a super-additive relationship between attention and viewpoint (which the findings reported here do not support): Priming for attended images in the same view should be (slightly) greater than the sum of priming for ignored objects in the same view and priming for attended images in the reflected view. The reason for this prediction has to do with the relationship between attention and the SSM.

When an image is first presented to the model, it is initially represented (as a whole) as a pattern of activation on both the IGA and the SSM. If the image is attended, then it will eventually be segmented into its parts (following Hummel & Biederman, 1992), and each part will be represented by its own a pattern on both the IGA and the SSM. Thus, an attended image with N parts will be represented as a collection of 2N+2 patterns: N+1 on the IGA (one pattern for each part plus one for the image as a whole) and N+1 on the SSM. Ignored images are not segmented, so an ignored image will be represented by only one pattern on the IGA and one on the SSM.

With these details in hand, it is now posssible to explain why the model predicts super-additivity between attention and viewpoint. Attending to an image activates and primes N+1 patterns on the IGA and N+1 on the SSM (a total of 2N+2 patterns). If that image is subsequently presented again in the same viewpoint (after a short enough interval), then all 2N+2 primed patterns will facilitate recognition. If the attended image is subsequently presented in the left-right reflected view, then the N+1 primed patterns on the IGA will facilitate recognition, but the N+1 primed patterns on the SSM will not (recall that the IGA, but not the SSM, is invariant with left-right reflection). Ignoring an image will prime one pattern on the IGA an one on the SSM. If that ignored image is presented again in the same view, then these two primed patterns will facilitate recognition. Adding the priming for an ignored image in the same view (corresponding two patterns) to the priming for an attended image in the left-right reflected view (N+1 patterns) results in the sum 2+N+1, or N+3. This sum is less than the priming obtained for attended images in the same view (2N+2): Attention and viewpoint are super-additive. This analysis is completely linear in that the sources of priming were compared simply by adding them together. If we assume that these sources of priming are subjected to a non-linearity in the processes of activating a representation in memory (the very reasonable assumption that leads to the original prediction of sub-additivity), then the magnitude of the predicted super-additivity may be reduced substantially. But as this analysis shows, the parallel model of the automatic (or semi-automatic) and controlled representations does not necessarily predict a sub-additive relationship between attention and viewpoint.

Early vs. late selection

The findings reported here have important implications in the context of the question of early vs. late selection: Consistent with the late selection view, our findings (like those of Tipper and others) suggest that we can recognize objects without attending to them. But consistent with the early selection view, our findings suggest that attention nonetheless plays an important role in object representation: Although attention does not determine *whether* an object will be recognized, it determines *how it will be represented* for recognition.

Properties of the two shape representations

The present experiments revealed some of the properties of the visual representations activated with attention (the controlled representation) and without attention (the automatic representation). We can draw further insights about the nature of the controlled representation from other experiments that have used long-term priming to investigate the visual representation of shape. (These experiments can inform our understanding of the controlled representation because we have shown that priming for the automatic representation is short-lived whereas the priming for the controlled representation is longer lasting.) Together, these experiments and the experiments reported here are beginning to paint a picture of the differences between the shape representations generated with and without attention. Based on such experiments, we can infer that the controlled representation is based on volumetric parts (Biederman & Cooper, 1991b), is invariant with scale (Biederman & Cooper, 1992), and possibly invariant with rotation in depth (Biederman & Gerhardstein, 1993; but see Bülthoff & Edelman, 1992; Tarr, 1995). Some of these properties are summarized in Table 1.

Insert Table 1 about here

The paradigm used in the experiments reported here suggests itself as a tool to help fill the empty cells of Table 1. An important question concerns the nature of the visual primitives used by the automatic representation. Although the controlled representation appears to use simple volumetric parts (e.g., geons), it remains to be seen whether the automatic representations also uses geon-like primitives. Another open question about automatic representation concerns its sensitivity to changes in scale, translation and rotations in depth. It is important to recognize that the current experiments speak to the role of attention in generating a representation of shape that is invariant with left-right reflection and not to the role of attention in compensating for all changes in viewpoint. Some models, such as the Hummel and Stankiewicz model, generate scale- and translation- invariant shape representations without attention, while other models, such as the Olshausen et al. model, explicitly implicate attention in compensating for these viewpoint changes. Future experiments will investigating the role of attention in compensating for other viewpoint changes.

Some findings suggest that the visual system stores one or more canonical (or preferred) views of each known object. Presumably, which view is canonical is a function of experience (e.g., Jolicoeur, 1985; Palmer, Rosch & Chase, 1981; Tarr & Pinker, 1989). Another interesting question concerns whether the automatic and controlled representations show different sensitivities to experience. The Hummel and Stankiewicz model predicts that the controlled representation should be less sensitive to experience (i.e., showing less “preference” for canonical views) than the automatic representation.

Viewpoint invariance vs. viewpoint sensitivity

There is a long-standing debate in the object recognition community regarding the extent to which the representations mediating human object recognition are invariant with vs. sensitive to viewpoint (see Biederman & Gerhardstein, 1995; Tarr & Bülthoff, 1995). Proponents of the normalization-based approach argue that objects are represented in a view-sensitive format, while proponents of the structural description approach argue that objects are represented in a more view-invariant format. There is evidence for both views (for reviews see Biederman & Gerhardstein, 1993, 1995; Cooper, et. al., 1992; Tarr, 1995). For example, Biederman and Cooper (1992) reported long-term priming results suggesting that object recognition is scale-invariant. But using a same different paradigm, Bundesen and Larsen (1975) found that people responded faster when the to-be-compared objects were the same size than when they were different sizes, suggesting that object recognition is scale sensitive. Similarly, using a same-name/different name task, Ellis and Allport (1986) found evidence for view-sensitivity at short stimulus onset asynchronies (SOAs) (<2 seconds), but no effects of viewpoint at longer SOAs. In response to these and other seemingly inconsistent results, it is not uncommon to suggest that object recognition is based on more than one representation, one viewpoint sensitive, and the other more viewpoint invariant (Corballis, 1988; Ellis & Allport, 1986; Farah, 1992; Jolicoeur, 1990). However, with the exception of a few studies (Ellis & Allport, 1986; Ellis, et. al., 1989; Farah, 1992; Tarr & Pinker, 1990) most of the evidence for such a dual representation has been largely indirect because it comes from different experiments using different paradigms.

The experiments presented here provide additional direct evidence for multiple representations of object shape, one more robust to variations in viewpoint than the other. In addition, these experiments demonstrate another important difference between these two forms of visual representation: one requires attention and the other does not. It is tempting to speculate that some of the invariances observed by others reflect the properties of the controlled representation while some of the view sensitivities reflect the automatic representation. If we entertain this speculation, then we may begin to better understand some of the properties of these representations. For example, many demonstrations of viewpoint sensitivity are based on same/different judgment tasks, in which the time between the target and the probe is relatively short (on the order of seconds) (for a review see Cooper, et. al., 1992). Inter-stimulus-intervals of this duration are within the time frame for the priming we found for the viewpoint-sensitive automatic representation. It is therefore plausible that both the automatic and controlled representations will manifest themselves in subject’s performance on such tasks. If either (or both) of these representations is sensitive to viewpoint, then such tasks will demonstrate viewpoint sensitivity. By contrast, experiments using a long-term priming paradigm (with prime-probe delays on the order of minutes) more often show viewpoint invariance (Biederman & Cooper, 1991a; 1992; Biederman & Gerhardstein, 1993). In the context of our findings, the lag between the prime and probe displays in these experiments is longer than the implicit memory trace we observed for the automatic representation. If the memory trace for the automatic representation has dissipated by the time the probe is displayed, then one may only observe the effects of the more view-invariant representation.

Conclusions

The findings reported here suggest that the visual representation generated in response to an attended object is qualitatively different from that generated in response to an ignored object. Importantly, recognition takes place in both cases: Object recognition does not require attention. But the properties of the representations that mediate recognition depend substantially on whether the object is attended or ignored.

References

Biederman, I. (1987). Recognition-By-Components: A Theory of Human Image Understanding. *Psychological Review*, *94*, 115-147.

Biederman, I., Blickle, T.W., Teitelbaum, R. C.; & Klatsky, G.J. (1988) Object search in nonscene displays. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 456-467.

Biederman, I., & Cooper, E. E. (1991a). Evidence for complete translational and reflectional invariance in visual object priming. *Perception*, *20*, 585-593.

Biederman, I., Cooper, E.E. (1991b). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology*, *23*, 393-419.

Biederman, I., & Cooper, E.E. (1992) Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception & Performance*, *18*, 121-133.

Biederman, I., & Gerhardstein, P.C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception & Performance*, *19*, 1162-1182.

Biederman, I. & Gerhardstein, P. C. (1995). Viewpoint-dependent mechanisms in visual object recognition: A critical analysis. *Journal of Experimental Psychology*: *Human Perception and Performance*. in press.

Bundesen, C., & Larsen, A. (1975). Visual transformation of size. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 214-220.

Cooper, E.E.; Biederman, I.; & Hummel, J.E. (1992) Metric invariance in object recognition: A review and further evidence. *Canadian Journal of Psychology*, *46*,191-214.

Corballis, M.C. (1988). Recognition of disoriented shapes. *Psychological Review*, *95*, 115-123.

Dickinson, S.J., Pentland, A.P., & Rosenfeld, A. (1992). 3-D shape recovery using distributed aspect matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *14*, 174-198.

Edelman, S., & Weinshall, D. (1991). A self-organizing multiple-view representation of 3-D objects. *Biological Cybernetics*, *64*, 209-219.

Ellis, R., & Allport, D. A. (1986). Multiple levels of representation for visual objects: A behavioural study (pp. 245-257). In A.G. Cohen & J. R. Thomas (Eds.) *Artificial Intelligence and its Applications*. New York: Wiley.

Ellis, R., Allport, D.A., Humphreys, G.W., & Collis, J. (1989) Varieties of object constancy. *Quarterly Journal of Experimental Psychology*, *41*, 775-796.

Enns, J.T (1992) Sensitivity of early human vision to 3-D orientation in line drawings. *Canadian Journal of Psychology*, *46*,143-169.

Farah, M. J. (1992). Is an object an object an object? Cognitive and neuropsychological investigations of domain specificity in visual object recognition. *Current Directions in Psychological Science*, *1*, 164-169.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993a). Visual attention revealed by an illusion motion. *Neuroscience Research*, *18*,11-18

Hikosaka, O., Miyauchi, S., & Shimojo, S.(1993b). Voluntary and stimulus-induced attention detected as motion sensation. *Perception*, *22*, 517-526.

Hummel, J. E., Biederman, I. (1990). Dynamic binding: A basis for the representation of shape by neural networks. *Proceedings of the 12th Annual Conference of the Cognitive Science Society*, (614-621).

Hummel, J. E., & Biederman, I. (1991). Binding by phase locked neural activity: Implications for a theory of visual attention. Paper presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, Fl. May.

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

Hummel, J. E., & Stankiewicz, B. J. (1995). An architecture for rapid, hierarchical structural description. T. Inui and J. McClelland (Eds.) In *Attention and Performance XVI*. In Press.

Jolicoeur, P. (1990). Identification of disoriented objects: A dual- systems theory. *Mind & Language*, *5*(4).

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory & Cognition, 13*, 289-303.

LaBerge, D. & Brown, V. (1986). Variations in size of the visual field in which targets are presented: An attentional range effect. *Perception and Psychophysics, 40*, 188-200.

LaBerge, D. & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, *96*, 101-124.

Logan, G.D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception & Performance*, *20*,1015-1036.

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

Marr, D. & Nishihara, H.K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London B* 200:269-294.

Olshausen, B. A.; Anderson, C. H.; & Van Essen, D. C. (1993) A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience*, *13*, 4700-4719.

Palmer, S.E., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp.135-151). Hillsdale, NJ: Erlbaum.

Poggio T.; Edelman S. (1990). A network that learns to recognize 3-Dimensional objects. *Nature*, *343*, 263-266.

Quinlan, Philip T. (1991). Differing approaches to two-dimensional shape recognition. *Psychological Bulletin*, *109*, 224-241.

Ruthruff, E. & Miller, J. (1995). Negative priming depends on ease of selection. *Perception & Psychophysics*, *57*, 715-723.

Seibert, M., & Waxman, A.M. (1992). Adaptive 3-D object recognition from multiple views. *IEEE Transactions on pattern analysis and machine intelligence*, *14*,107-124.

Srinivas, K. (1996). Size and reflection effects in priming: A test of transfer-appropriate processing. *Memory & Cognition*, *24*, 441-452.

Sternberg, S. (1969) Memory scanning: Mental processes revealed by reaction time experiments. *American Scientist*, 57, 421-457.

Sutherland, N.S. (1979). The representation of three-dimensional objects. *Nature, 278* ,395-398.

Tarr, M.J. (1995). Rotating objects to recognize them - a case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review, 2*, 55-82.

Tarr, M. J., & Bülthoff, H. (1995). Conditions for viewpoint-dependence and viewpoint-invariance: What mechanisms are used to recognize an object? *Journal of Experimental Psychology*: *Human Perception and Performance*. in press.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation dependence in shape recognition. *Cognitive Psychology, 21*, 233-283.

Tarr, M.J. & Pinker, S. (1990) When does human object recognition use a viewer-centered reference frame? *Psychological Science, 1* ,253-256.

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

Tipper, S. P. & Driver, J. (1988). Negative priming between pictures and words in a selective attention task: Evidence for semantic processing of ignored stimuli. *Memory & Cognition*, v16: 64-70.

Treisman, A. & DeShepper (1995). Object Tokens, Attention, and Visual Memory. In *Attention and Performance XVI*. In Press.

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

Treisman, A., & Schmidt, H. (1982) Illusory conjunctions in the perception of objects. *Cognitive Psychology* 14:107-141.

Ullman, S.; Basri R. (1991). Recognition by linear combinations of models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *13*, 992-1006.

Winston, P.H. (1975). Learning structural descriptions from examples. In P.H. Winston (Ed.), *The psychology of computer vision* (pp. 157-209). New York: Academic Press.

Wolfe, J.M., Cave, K.R., & Franzel, S.L. (1989). Guided search: An alternative to the feature integration model for search. *Journal of Experimental Psychology: Human Perception & Performance*, *15*,419-433.

Author Notes

This work was supported by a Grant-in-Aid of Reseach to BJS from the National Academy of Sciences, through Sigma Xi, and a research grant from the UCLA Academic Senate to JEH. The authors would like to thank Mike Corballis and two anonymous reviewers for their comments on earlier drafts of the manuscript. The authors would also like to thank Nancy Kanwisher, Jun Saiki for discussions about this work. Please direct all correspondence to Brian J. Stankiewicz, Department of Psychology, Attn: Grad Mail Room, 405 Hilgard Ave, University of California, Los Angeles, Los Angeles, CA 90095-1563, Email: bstankie@psych.ucla.edu.

Footnotes

[1] It is important to realize that the predictions made by the Olshausen et al. model do not necessarily speak to all models that use normalization procedures.

[2] There were other conditions as well, but these three are of primary interest in the context of our current discussion.

[3] With respect to the Hummel and Stankiewicz model the prediction about left-right reflection is particularly clear: an attended image should prime itself and its left-right reflection whereas an ignored image should prime only itself and not its left-right reflection.

[4] "Not-probed" objects are objects that appeared on a prime trial, but that did not appear on the subsequent probe trial. For example, when subjects were probed with the ignored prime, the attended object in that condition would be an Attended-Not-probed object.

[5] Although the unprimed image does not have a direction vis. either prime, each unprimed object was presented facing one direction for half of the subjects and the other direction for the other half of the subjects.

[6] Recently we found direct evidence that the representation generated without attention is invariant with translation (Stankiewicz & Hummel, 1996).

[7] The greater number of ommitted trials in the current experiments relative to those reported by Biederman and Cooper is due primarily to the nature of the prime display. Often, participants either could not name the attended prime, or they named the "ignored" image by mistake. These trials were ommited from analysis.

Table 1

Properties of the Controlled and Automatic Representations

Controlled Representation	**Automatic Representation**
• Requires Attention	• Activated without attention
• Invariant with left-right reflection (Experiments 1-3; Biederman & Cooper, 1991a)	• Sensitive to left-right reflection
• Long-lived priming (> ~5 minutes)	• Short-lived priming (> 3 seconds < ~5 minutes)
• Invariant with location (Experiment 1; Biederman & Cooper, 1991a)	• At least partially invariant with location
• Invariant with scale (Biederman & Cooper ,1992)	• **Unknown**
• Invariant with Depth Rotation (up to part occlusion) (Biederman & Gerhardstein ,1993)	• **Unknown**
• Based on volumetric primitives (Biederman & Cooper ,1991b)	• **Unknown**

Figure Captions

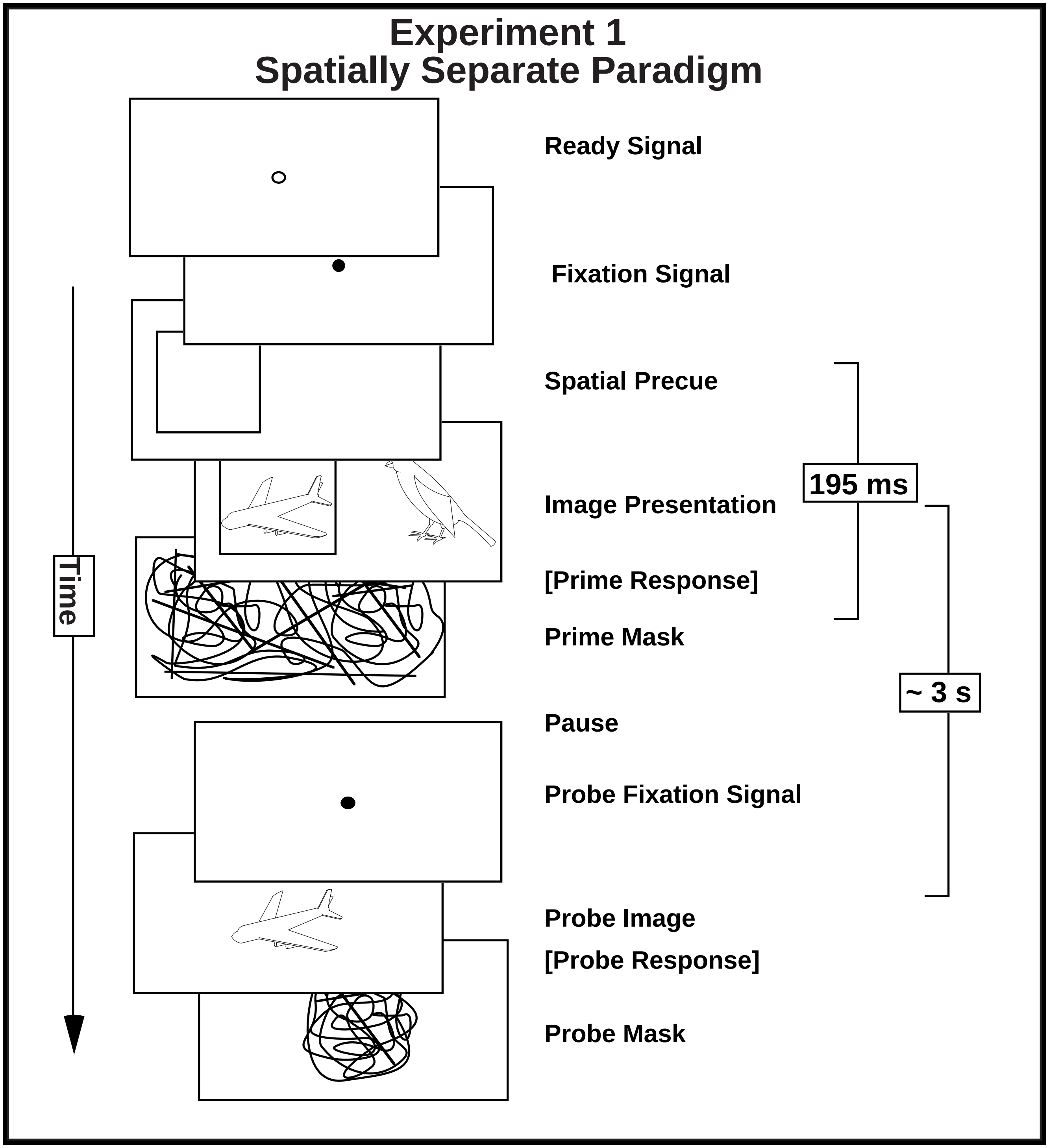
Figure 1: Illustration of display sequence for Experiment 1.

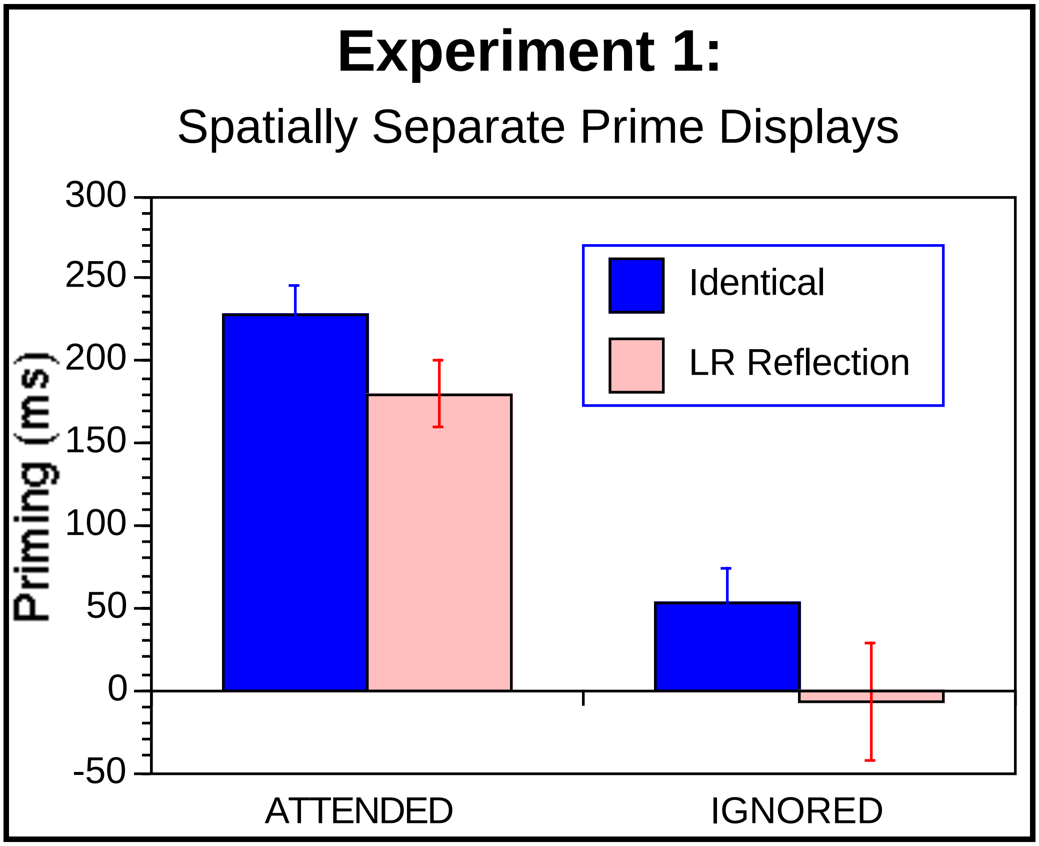
Figure 2: Priming means and standard errors for Experiment 1 as a function of whether the object was attended or ignored and whether the object was identical to the prime image or a left-right (LR) reflection of the prime.

Figure 3: Priming means and standard errors for Experiment 2 as a function of whether the object was attended or ignored and whether the object was identical to the prime image or an object having the same name but being a different exemplar (SNDE).

Figure 4: Short-term priming means and standard errors for Experiment 3 as a function of whether the object was attendedor ignored and whether the object was identical to the prime image or a left-right (LR) reflection of the prime.

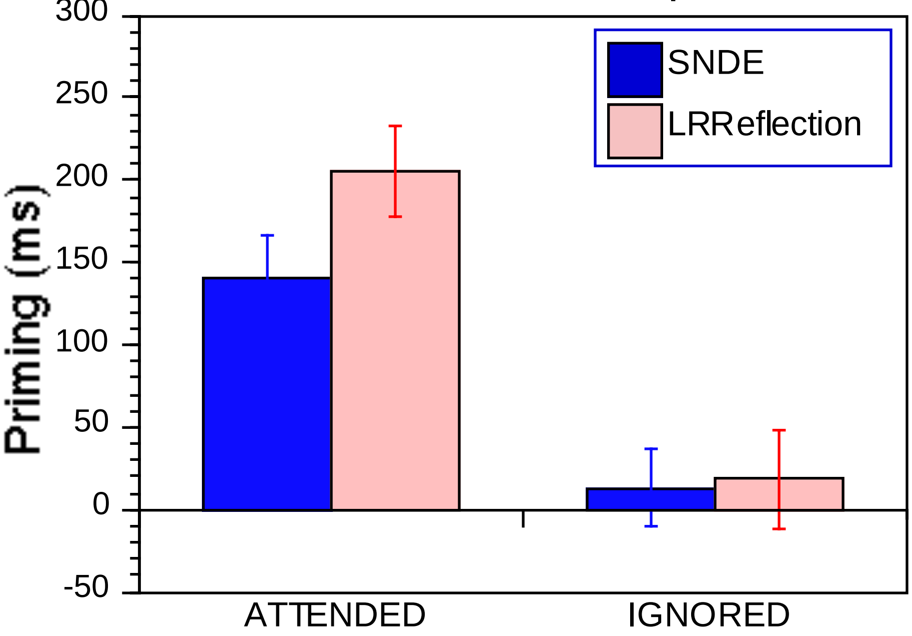
Figure 5: Long-term priming means and standard errors for Experiment 3 as a function of whether the object was attended or ignored and whether the object was identical to the prime image or a left-right (LR) reflection of the prime.

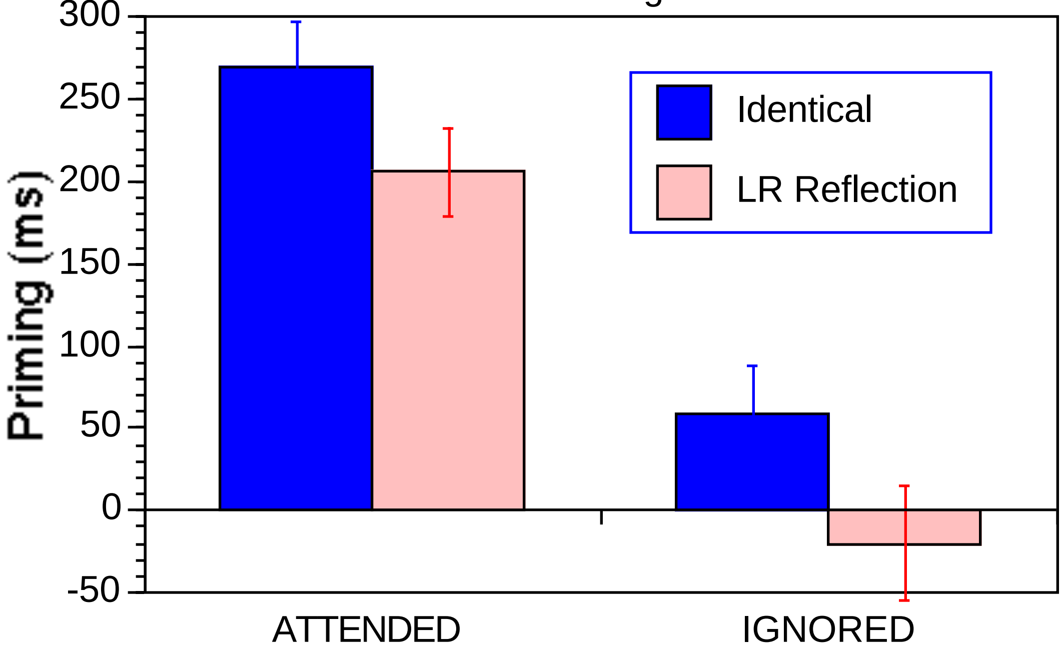

Experiment 1
Spatially Separate Paradigm
Ready Signal
Fixation Signal
Spatial Precue
Image Presentation
[Prime Response]
Prime Mask
Pause
Probe Fixation Signal
Probe Image
[Probe Response]
Probe Mask
195 ms
~ 3 s
Time


Experiment 1:
Spatially Separate Prime Displays
Priming (ms)
300
250
200
150
100
50
0
-50
Identical
LR Reflection
ATTENDED
IGNORED

Experiment 2:

Same Name Different Shape Control

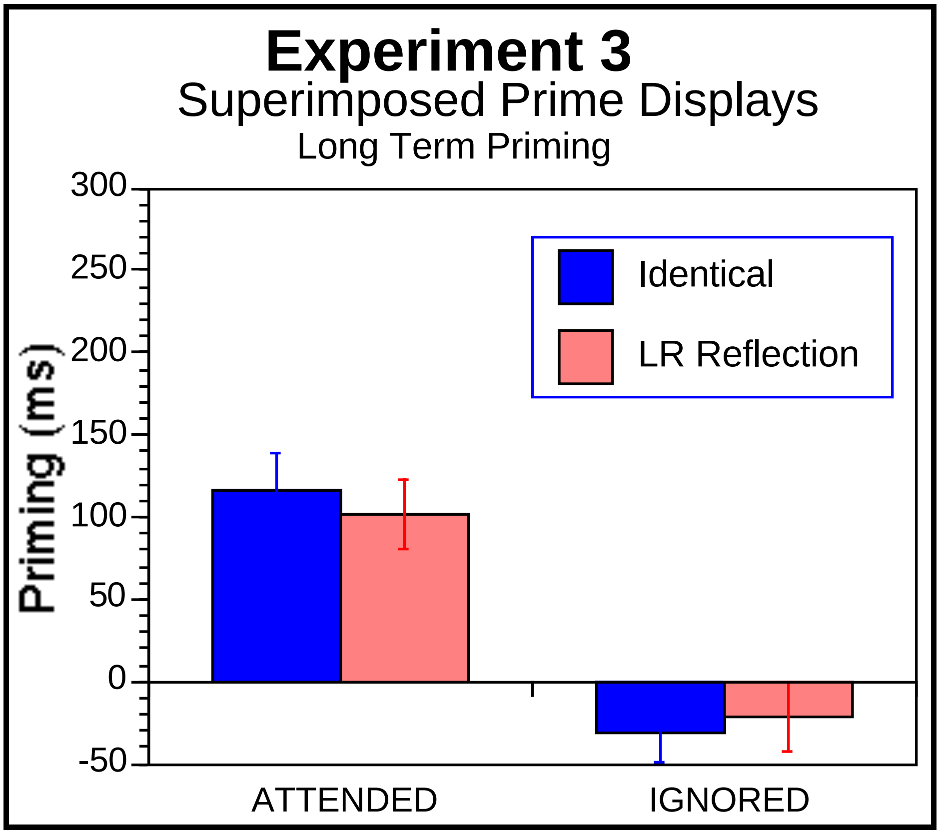




Experiment 3

Superimposed Prime Displays

Short-Term Priming


Experiment 3
Superimposed Prime Displays
Long Term Priming
300
250
200
150
100
50
0
-50
Priming (ms)
Identical
LR Reflection
ATTENDED
IGNORED