Psychonomic Bulletin & Review
2000, 7 (3), 480-489

Viewpoint-invariant and viewpoint-dependent object recognition in dissociable neural subsystems

E. DARCY BURGUND and CHAD J. MARSOLEK
University of Minnesota, Minneapolis, Minnesota

Participants viewed objects in the central visual field and then named either same or different depth-orientation views of these objects presented briefly in the left or the right visual field. The different-orientation views contained either the same or a different set of parts and relations. Viewpoint-dependent priming was observed when test views were presented directly to the right hemisphere (RH), but not when test views were presented directly to the left hemisphere (LH). Moreover, this pattern of results did not depend on whether the same or a different set of parts and relations could be recovered from the different-orientation views. Results support the theory that a specific subsystem operates more effectively than an abstract subsystem in the RH and stores objects in a manner that produces viewpoint-dependent effects, whereas an abstract subsystem operates more effectively than a specific subsystem in the LH and does not store objects in a viewpoint-dependent manner.

The ability to recognize the same object from many different viewpoints is remarkable, given that differing viewpoints can produce quite dissimilar input images. The explanation for this ability is the subject of much controversy in the object recognition literature (e.g., Biederman & Gerhardstein, 1995; Tarr & Bülthoff, 1995). One group of theorists proposes that object recognition relies on a viewpoint-invariant process that uses independent representations of invariant parts or features of an object (e.g., Biederman, 1987; Biederman & Gerhardstein, 1993; Hummel & Biederman, 1992; Hummel & Stankiewicz, 1996; Koenderink, 1987; Marr, 1982; Wagemans, Gool, & Lamote, 1996). According to most viewpoint-invariant theories, once a particular object has been stored, recognition of that object from any view (including novel views) should be unaffected by the viewpoint, provided that the necessary features can be recovered from that view. Another group of theorists proposes that object recognition relies on a viewpoint-dependent process involving view-specific images or undifferentiated representations of whole shapes (e.g., Bülthoff & Edelman, 1992; Edelman, 1998; Edelman & Bülthoff, 1992; Humphrey & Khan, 1992; Poggio & Edelman, 1990; Tarr, 1995; Ullman, 1989, 1996). According to this perspective, once a particular object has been stored, recognition of that object from novel views should be impaired, relative to recognition of previously stored views. Unfortunately, the current set of behavioral findings does not unequivocally settle the debate between these two theoretical approaches. In this article, we examine a potential resolution to the controversy. In particular, we investigate whether dissociable neural subsystems underlie object recognition in ways that are differentially prone to the production of viewpoint effects.

Support for both viewpoint-invariant and viewpoint-dependent object recognition has been obtained in previous studies of object priming. Viewpoint-invariant recognition is evidenced when test objects are processed equally well after having been viewed previously in either the same or a different orientation in depth during initial encoding. Viewpoint-dependent recognition is evidenced by greater same-orientation than different-orientation priming. Invariant priming has been observed when the same parts and relations could be recovered from the different-orientation views, but not when different parts and relations could be recovered (Biederman & Gerhardstein, 1993; Srinivas, 1995, Experiment 2). In contrast, viewpoint-dependent priming has been observed in other studies, regardless of whether the same or different parts could be recovered from the different-orientation views (Hayward & Tarr, 1997; Lawson & Humphreys, 1998; Srinivas, 1995, Experiment 3; Tarr, Bülthoff, Zabinski, & Blanz, 1997).

A potential explanation for the empirical discrepancy is that dissociable neural subsystems underlie object recog-

---

Parts of this research were reported earlier at the Annual Meeting of the Cognitive Neuroscience Society, San Francisco (1998). Funding for this work came from the Center for Cognitive Sciences in conjunction with the National Science Foundation (GER 9454163), the Office of the Vice President for Research, and Dean of the Graduate School of the University of Minnesota. In addition, we thank Beth Lavin and the Summer Research Experience for Undergraduates program, Center for Cognitive Sciences, University of Minnesota, as well as Jeremy Blessing, Lorien Parson, Kelly Pucel, Anna Shier, and Brett Stephan for their assistance with data collection and analysis. Correspondence concerning this article should be addressed to E. D. Burgund, Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455 (e-mail: burgu001@tc.umn.edu).

Copyright 2000 Psychonomic Society, Inc.

nition and perform qualitatively different processes that are differentially affected by viewpoint manipulations. Recent evidence indicates that relatively independent subsystems, operating in parallel, underlie abstract and specific object recognition (Marsolek, 1999). After viewing line drawings of objects and words that name other objects in the central visual field, participants named laterally presented objects that were the same exemplars as those previously viewed, different exemplars with the same names as previously viewed objects, objects corresponding to the previously viewed words, and unprimed objects. When test objects were presented directly to the left cerebral hemisphere (LH), same- and different-exemplar primed objects were named with equal accuracy, and both were named more accurately than word-primed objects. In contrast, when test objects were presented directly to the right cerebral hemisphere (RH), same-exemplar-primed objects were named more accurately than both different-exemplar-primed objects and word-primed objects (and accuracy in the latter two conditions did not differ). Several important conclusions can be drawn from this study. First, an *abstract-category subsystem* stores object categories and operates more effectively than a specific-exemplar subsystem in the LH, whereas a *specific-exemplar subsystem* stores object exemplars and operates more effectively than an abstract-category subsystem in the RH. Second, even an abstract-category subsystem relies on visual—rather than postvisual—information for object recognition (as is evidenced by the finding that same- and different-exemplar-primed objects were recognized more effectively than word-primed objects). Finally, evidence for abstract priming coupled with no evidence for specific priming when objects were presented directly to the LH, and vice versa when objects were presented directly to the RH, indicates that abstract-category and specific-exemplar subsystems operate in parallel, rather than in sequence.[1]

These subsystems appear to rely on contradictory processing strategies to accomplish their respective goals (for further explication and evidence from neural network models, see Marsolek & Burgund, 1997). A features-based processing strategy, in which only the features that are relatively invariant across different exemplars within the same abstract category are activated in the long-term memory representations, should be useful (although admittedly not *necessary*) for recognizing abstract categories (Marsolek, 1995). This is because the information that is common to even highly dissimilar inputs belonging to the same abstract category (e.g., consider a grand piano and an upright piano) typically is found in only a subset of the features of any one input form (e.g., the features defining the keyboard, two front legs, etc.). The "features" stored by an abstract subsystem may be very abstract structural primitives reflecting relatively invariant information across exemplars within a category and exhibiting a high degree of tolerance in how they are activated by perceptual inputs (e.g., perhaps not unlike the "complex features" that activate single cells in the inferior temporal cortex [K. Tanaka, 1993] or generalizable cylinders of the sort described by Marr, 1982). Such features would allow parts to vary in some parameters, such as curved versus linear edges, but not to vary much in other parameters, such as rough aspect ratio. Features-based processing may take the form of *sparse* (or relatively local) coding (for the benefits of such coding, see Churchland & Sejnowski, 1992) in the internal processing of an abstract subsystem. In this sort of sparse coding, different elements of a representation would be differentially sensitive to different features of an input object. This would help a subsystem to learn representations that are activated by relatively invariant information across exemplars within a category, are not activated or inhibited by information that varies across exemplars, and are inhibited by information not found in exemplars in that category (benefiting category recognition, but not exemplar recognition).

In contrast, a whole-based processing strategy, in which features of an exemplar are stored all of a piece and not independently of one another at any stage of processing, should be useful (although admittedly not *necessary*) for recognizing specific exemplars (Marsolek, Schacter, & Nicholas, 1996). This is because the information that distinguishes even highly similar exemplars in one abstract category (e.g., consider two similar grand pianos) from each other, as well as from all other exemplars, typically is found in information close to the undifferentiated whole of one input form. Whole-based processing may take the form of distributed (or *population*) coding (for the benefits of such coding, see Ballard, 1986; Hinton, McClelland, & Rumelhart, 1986) in the internal processing of a specific subsystem. In this sort of distributed coding, the entire pattern of activation across a large number of different elements of a representation (each sensitive in different ways to, perhaps, all of an object) would represent an input object. This would help a subsystem to learn representations of entire stored exemplars in such a manner that individual features are not represented independently of one another (benefiting exemplar recognition, but not category recognition). Importantly, sparse coding and distributed coding are contradictory in nature, in that they cannot both be used simultaneously in the same representations, which suggests a reason why at least relatively independent abstract and specific subsystems may have developed (Marsolek & Burgund, 1997).

Indeed, results from a recent visual-priming study indicate that abstract and specific subsystems rely on these different processing strategies (Marsolek et al., 1996). In that study, letter case-specific priming (greater same-case-primed than different-case-primed performance) in RH test presentations was observed only when the visual contexts presented above the target items were the same between initial encoding and subsequent test. Lettercase-specific priming was not observed, even in RH test presentations, when the visual contexts for the target items differed between encoding and test. Conversely, lettercase-abstract priming (equivalent same-case-primed and different-case-primed performance) in LH test presentations was not affected by visual context. Lettercase-abstract priming was observed in LH test presentations regardless of whether the visual contexts for test items differed between encoding and test. Taken together, these findings indicate that a specific-exemplar subsystem represents entire, undifferentiated configurations of inputs (i.e., target words and context words are integrated in the same whole representation, and hence, priming of targets requires repetition of their primed contexts), rather than subwhole elements of the input independently of each other. In contrast, an abstract-category subsystem *does* represent subwhole elements of an input independently of each other (i.e., target words are represented independently of their visual contexts, and hence, priming of targets can occur in a manner that is not influenced by their contexts). Thus, a specific subsystem may rely on a relatively whole-based processing strategy, whereas an abstract subsystem may rely on a relatively features-based processing strategy.[2]

We suggest that, with the development of abstract and specific subsystems, two mechanisms may emerge for underlying viewpoint effects in object priming. In particular, an abstract subsystem that uses a features-based processing strategy should store features that are common to different views of the same exemplar, as well as common to different exemplars in the same abstract category. Hence, the priming from such representations should *not* be viewpoint dependent. In contrast, a specific subsystem that uses a whole-based processing strategy should store whole shapes at each stage of processing and should include at least one stage in which the representations are specific to the distinctive novel view of a particular object as it appears during an initial encoding trial. Hence, the priming from such representations should be viewpoint dependent. We tested these hypotheses in the following experiment.

First, the participants initially encoded a series of objects, each from a particular viewpoint. Then, during a subsequent test phase, they named objects presented directly to the LH or directly to the RH. Test objects were presented in either the same or different orientation in depth, as compared with initial encoding, and viewpoint-dependent priming was measured as significantly greater accuracy in naming objects presented in the same orientation, rather than in a different orientation. According to the dissociable subsystems theory, different results should be observed, depending on the hemisphere of direct test presentations and, hence, which subsystem is given an advantage during test. If an abstract subsystem operates more effectively than a specific subsystem in the LH, no viewpoint-dependent priming should be observed when test objects are presented directly to the LH. If a specific subsystem operates more effectively than an abstract subsystem in the RH, viewpoint-dependent priming should be observed when test objects are presented directly to the RH. Alternatively, single-system viewpoint-dependent theories would predict viewpoint-dependent priming in both LH and RH test presentations.

Another question addressed in this experiment was whether priming would be sensitive to the parts that could be recovered from the different-orientation views. One hypothesis is that viewpoint-dependent priming should be observed only when different parts and relations can be recovered (owing to differential occlusion of parts) from the prime and the test views in the different-orientation condition (Biederman & Gerhardstein, 1993, 1995). To examine this question, we included a parts-occlusion variable. Half of the participants were presented with different-orientation prime views in which different parts could be recovered, as compared with their test views, and the other half were presented with different-orientation prime views in which the same parts and relations could be recovered, as compared with their test views. If, as was proposed by Biederman and Gerhardstein (1993, 1995), different discernible parts and relations must be recovered in order to observe viewpoint-dependent priming, viewpoint-dependent priming should be observed in the different-part condition, but not in the same-part condition.

METHOD

Participants

Sixty-four male and 64 female undergraduate students at the University of Minnesota volunteered to participate for payment or for course credit. All were restricted to be fairly strongly right-handed, as assessed through the Edinburgh Handedness Inventory (mean laterality quotient = .88; range = .50–1.0; Oldfield, 1971) and had normal or corrected-to-normal vision.

Design

A 2 × 2 × 2 × 2 mixed factorial design was used. Prime orientation (same orientation as the test view vs. different orientation than with the test view) and hemisphere of direct test presentations (LH vs. RH) were within-subjects independent variables. Parts and relations recoverable from the different-orientation prime and test views (same parts vs. different parts) and type of test task (sketching and then naming vs. naming) were between-subjects independent variables. Figure 1 depicts examples of the priming conditions.[3]

Materials

Images from 24 objects were selected from the Verfaillie and Boutsen (1995) corpus of depth-rotated images of familiar objects. For each object, three views were selected: an arbitrary 0° orientation (this view was always presented as the test view), a view differently oriented from the test view by 45° such that the same set of parts and relations could be recovered (same-part, different-orientation encoding view), and a view differently oriented from the test view by 45° such that a different set of parts (at least one part difference per object) and relations could be recovered (different-part, different-orientation encoding view). A part was defined as a volumetric primitive that could be recognized from the two-dimensional viewpoint-invariant properties in the image (a geon; Biederman, 1987; Hummel & Biederman, 1992). The mean raw number of different parts between the test and different-part encoding views was 12, but note that this number was large, owing to the fact that some objects had multiple homogeneous or hierarchical parts that were affected in common ways by the viewpoint change (e.g., the keys on the telephone; see

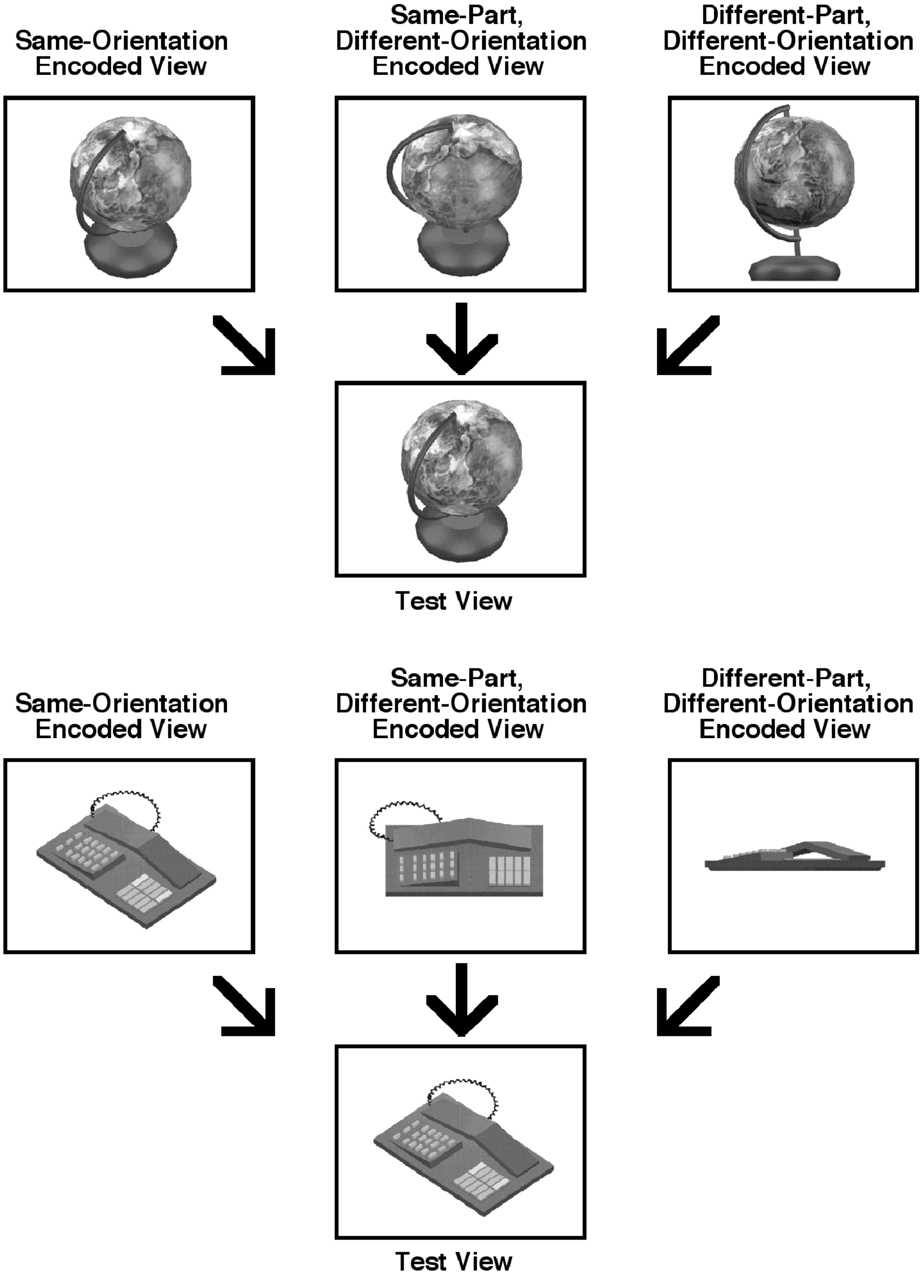

Figure 1. Examples of the types of priming examined in this experiment. Each test view of an object was primed in one of three ways: (1) same orientation, (2) same parts, different orientation, and (3) different parts, different orientation.

Figure 1). Presumably, the raw number of different parts across views should have different effects on priming, depending on the total number of parts evident in both views (e.g., a 1-part difference for a 100-part object should have less of an effect on priming than does a 1-part difference for a 3-part object). Thus, for each object, we calculated a difference proportion (the number of different parts between the test and the different-part encoding views divided by the total number of parts evident across both views) to quantify how much the test view differed from its corresponding different-part, different-orientation encoding view. The mean proportion for the different-part views used in this experiment was .66. We used the difference proportions to counterbalance a measure of perceived difference between test and different-part encoding views across conditions and participants.

The 24 test views were divided into four lists of six objects each. The lists were equated so that they did not differ significantly from each other in terms of their difference proportion (relevant for their different-part encoding views) or in terms of the *goodness* of the test views, as measured and reported by Verfaillie and Boutsen (1995). This goodness measure is essentially a rating of how canonical (Palmer, Rosch, & Chase, 1981) viewers perceive a particular view to be for a given object. Four within-subjects encoding test con-

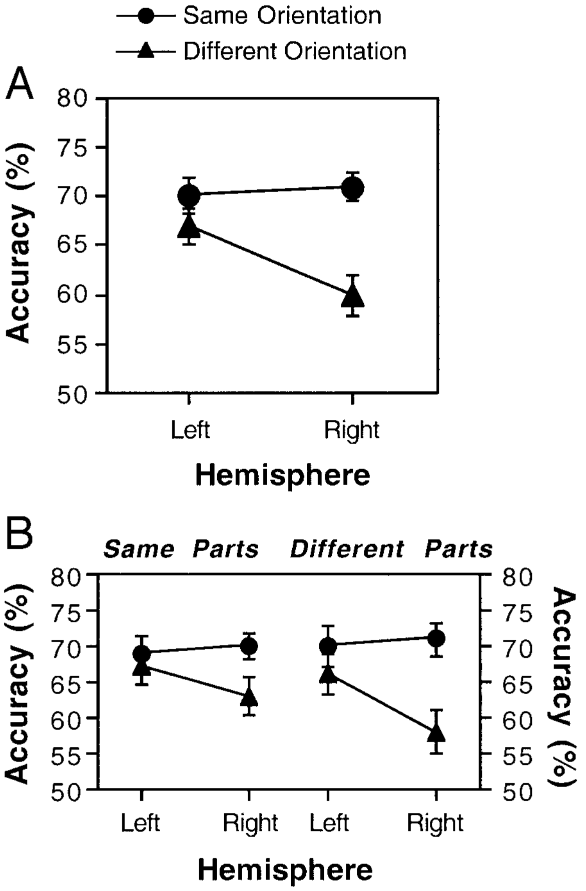

Figure 2. Mean accuracy rates displayed as a function of prime orientation and hemisphere of direct test presentation (A). Also, the same data displayed separately for the group with same-parts recoverable different-orientation primes (B, left) and for the group with different-parts recoverable different-orientation primes (B, right). Error bars indicate standard errors of the mean.

ditions were created by crossing prime orientation (same orientation vs. different orientation) with hemisphere of direct test presentation (RH vs. LH). For counterbalancing purposes, each list was rotated through those four conditions, as well as through the combinations of between-subjects variables. Counterbalancing ensured that each test view (including left/right reflections of each view, which were held constant for an object for each participant) represented each experimental condition equally often across participants.

Objects were presented in 256-level gray scale against a white background and subtended 6.4° of widest (horizontal or vertical) visual angle. During the encoding phase, objects appeared in the center of the monitor. During the test phase, objects were presented in the left or the right visual field so that the center of each object was 7.4° from the center of the monitor and the inner edge of an object never appeared closer than 4.2° from the center. A chinrest was employed to keep participants' eyes approximately 79 cm from the monitor. Stimuli were presented on an AppleVision 1710AV Display connected to a Power Macintosh 7600-132 computer. Stimulus presentation and response time measurement were controlled through the use of the PsyScope software package (J. Cohen, MacWhinney, Flatt, & Provost, 1993) and the PsyScope button box with a connected Optimus Pro 50 MX boom microphone.

Procedure

During the initial encoding phase, the participants were presented with one of the encoding views from each of the 24 objects (plus five filler images, three at the beginning and two at the end, to attenuate primacy and recency effects). Half of the objects were presented in the same orientation as they would be at test, and half were presented in a different orientation. For half of the participants, the different-orientation encoding view contained the same set of parts as those that could be recovered from the test view, and for the other half of the participants, the different-orientation encoding view contained at least one discernible difference in parts, as compared with the test view.

Each trial began with the presentation of a warning signal in the form of a fixation point, which appeared at the center of the display for 500 msec. Then, a recording of the name of the object was presented, in order to ensure that the participant would recognize the object correctly during its initial presentation and, potentially, to help standardize the names produced for multiname objects during the subsequent test phase. A second fixation point followed the auditory recording for 500 msec. The participants were instructed to focus on this fixation point. Immediately after the point disappeared, an object image was presented centrally for 3 sec, and after the object disappeared, the participants pressed a number key from 1 to 5 on the computer keyboard to indicate how much they liked or disliked that object. For these judgments, the participants were asked to consider the meanings associated with the object, as opposed to how they sounded, what they looked like, and so forth. We used this encoding task because both conceptual and structural information about the object may be used to make these decisions. However, on the basis of a similar study in which the conceptual and/or structural nature of the encoding task did not affect the object priming observed (Marsolek, 1999), we did not suspect that the type of encoding task used would be very important for the present results. Trials were presented in orders that were random, but with the constraints that no more than three objects appeared consecutively that would be presented in the right or left visual field or in the same or a different orientation at test.

During the subsequent test phase, the participants were presented with the test views for each of the 24 objects. Half of the objects were presented directly to the RH (in the left visual field), and half were presented directly to the LH (in the right visual field). A test trial began with the presentation of the fixation point at the center of the display for 500 msec. The participants were encouraged to focus their attention on the fixation point when it appeared and not to anticipate on which side of the fixation point the next object would appear. Immediately after the fixation point disappeared, an object appeared in the left or the right visual field for 14 msec. A blank screen followed and remained until the participant spoke the name of the object into the microphone connected to the button box. Half of the participants were instructed to speak aloud the name of the object as quickly and accurately as possible. The other half of the participants were instructed to sketch the object (as it appeared on the display) before naming it as accurately as possible. The sketch-then-name task was included in order to examine whether drawing an object in the particular viewpoint in which it appeared on the display would especially encourage viewpoint-dependent processing. When they were ready, the participants pressed the space bar to begin the next trial. Seven additional trials appeared at the beginning of the test phase for practice and warm-up. The trials were presented in orders that were random, but with the constraints that no more than three objects appeared consecutively in the right or left visual field or in the same or a different orientation, as compared with initial encoding.

RESULTS

A response was scored as correct when the participant named the object, using either a correct basic or a correct subordinate level term. For example, responses of "shoe," "tennis shoe," and "sneaker" were scored as correct responses for the sneaker, and responses of "airplane" and "jet" were scored as correct for the jet. Gender of participant, left/right reflection of the object, and test task did not enter into any significant effects in preliminary analyses; thus, these variables were excluded from the following analyses. Percentage of correct identifications of objects was analyzed in a three-way repeated measures analysis of variance (ANOVA). One ANOVA was performed with participants as the random variable (denoted F_1), and another was performed with items as the random variable (denoted F_2). Prime orientation (same orientation vs. different orientation) and hemisphere of direct test presentation (LH vs. RH) were within-subjects independent variables. Parts recoverable from different-orientation views (same parts vs. different parts) was a between-subjects variable.

Figure 2 illustrates the results from this experiment. The most important result was a significant interaction between prime orientation (same orientation vs. different orientation) and hemisphere of direct test presentation [LH vs. RH; see Figure 2A; $F_1(1,126) = 4.09$, $MS_e = 412.2$, $p < .05$; $F_2(1,94) = 5.15$, $MS_e = 247.6$, $p < .05$]. In particular, when objects were presented directly to the LH, same-orientation-primed objects (70%) and different-orientation-primed objects (67%) were not named with significantly different accuracy [$F_1(1,252) = 1.46$, $p > .22$; $F_2(1,188) = 2.21$, $p > .13$] for the simple effects contrasts. In contrast, when test objects were presented directly to the RH, same-orientation primed-objects (71%) were named more accurately than different-orientation-primed objects [60%; $F_1(1,252) = 19.6$, $MS_e = 395.6$, $p < .0001$; $F_2(1,188) = 22.3$, $MS_e = 260.4$, $p < .0001$] for the simple effects contrasts.

It is important to note that the two-way interaction between prime orientation and parts recoverable from different-orientation encoding views did not approach significance [$F_1(1,126) = 1.10$, $p > .29$; $F_2(1,94) = 1.17$, $p > .28$]; nor did the three-way interaction between those variables and hemisphere of direct test presentation (see Figure 2B; $F_1 < 1$, $F_2 < 1$). However, to directly test whether viewpoint-dependent priming would not be observed in direct LH test presentations but would be observed in direct RH test presentations for both same- and different-parts-recoverable conditions, we conducted the following simple effect contrasts. In the same-parts condition, when objects were presented directly to the LH, same-orientation-primed objects (69%) and different-orientation-primed objects (67%) were not named with significantly different accuracy ($F_1 < 1$, $F_2 < 1$), but when objects were presented directly to the RH, same-orientation primed objects (70%) were named more accurately than different-orientation primed objects [63%; $F_1(1,378) = 4.87$, $MS_e = 401.1$, $p < .05$; $F_2(1,282) = 5.61$, $MS_e = 256.1$, $p < .05$]. Similarly, in the different-parts condition, when objects were presented directly to the LH, same-orientation-primed objects (70%) and different-orientation-primed objects (66%) were not named with significantly different accuracy [$F_1(1,378) = 1.41$, $MS_e = 401.1$, $p > .23$; $F_2(1,282) = 1.63$, $MS_e = 256.1$, $p > .20$], but when objects were presented directly to the RH, same-orientation-primed objects (71%) were named more accurately than different-orientation-primed objects [58%; $F_1(1,378) = 13.31$, $MS_e = 401.1$, $p < .0005$; $F_2(1,282) = 15.36$, $MS_e = 256.1$, $p < .0005$].

The only other significant effect was the main effect of prime orientation [$F_1(1,126) = 15.4$, $MS_e = 379.0$, $p < .001$; $F_2(1,94) = 16.1$, $MS_e = 273.1$, $p < .001$; all other ps $> .08$]. Generally, same-orientation-primed objects (70%) were named more accurately than different-orientation-primed objects (63%).

Finally, for direct RH presentations only, a measure of viewpoint-dependent priming was computed by subtracting the naming accuracy for the different-orientation-primed views from the naming accuracy for the same-orientation-primed views for each test item. We then examined whether viewpoint-dependent priming covaried with the goodness of the test views (Verfaillie & Boutsen, 1995) in either the same- or the different-parts conditions (see Lawson & Humphreys, 1996, 1998, and Srinivas, 1993, 1995, for possible relations between canonicality and viewpoint-dependent priming) or with the difference proportions in the different-parts conditions (as would be predicted by Biederman & Gerhardstein, 1993, 1995). Since the distribution range for prime scores were very different from the ranges for goodness of views and difference scores, we normalized these distributions so that each had equivalent means and standard deviations before performing the correlation analyses. None of the Pearson r correlations between viewpoint-dependent priming and goodness of the test view or between viewpoint-dependent priming and difference proportion for that object was significant (all ps $> .38$).

Note that analyses of response times for correct responses were prohibited in this experiment by two factors. The times were highly variable and lengthy in the group of participants who sketched objects before naming them, and a large number of cells in the design were empty in the group of participants who simply named objects.

DISCUSSION

Viewpoint-dependent priming was observed when test objects were presented directly to the RH, but not when test objects were presented directly to the LH. Thus, contrary to single-system theories of object recognition, a specific subsystem operates more effectively than an abstract subsystem in the RH and stores objects in a manner that can produce viewpoint-dependent priming, whereas an abstract subsystem operates more effectively than a specific subsystem in the LH and stores objects in a manner that does not produce viewpoint-dependent priming.

In addition, neither the viewpoint-dependent priming observed in direct RH presentations nor the lack of viewpoint-dependent priming observed in direct LH presentations depended on whether the same or different parts and relations could be recovered from the different-orientation encoding and test views. Thus, contrary to another theory (Biederman & Gerhardstein, 1993, 1995), viewpoint-dependent priming may not be determined by whether the same or different parts and relations are recoverable across encoding and test views. The fact that the parts manipulation did not affect the lack of viewpoint-dependent priming in LH test presentations may seem at first to contradict our hypothesis that an abstract subsystem relies on a features-based processing strategy to store object categories. However, it is important to note that the parts manipulated in the present experiment were defined as geons (Biederman, 1987; Hummel & Biederman, 1992), in order to directly test predictions from Biederman and Gerhardstein (1993, 1995). As was mentioned in the introduction, it is likely that the features putatively stored in an abstract subsystem are more abstract than geons; thus, they may not have been systematically manipulated across our same- and different-parts conditions. Alternatively, the results may indicate that all the features for one abstract category are activated (and hence contribute to priming) when the viewer recognizes the abstract category of a features-occluded object (owing to a kind of vector completion or mental "filling in" process). If so, it would be difficult to observe a significant difference between same- and different-parts priming in such a subsystem in this kind of experiment. Finally, it should be noted that we did observe a trend (albeit, nonsignificant) for greater viewpoint-dependent priming in the different-parts condition than in the same-parts condition (see Figure 2B). However, this trend was observed in both direct LH and direct RH test presentations and thus, apparently, in both abstract and specific subsystems.

We conclude that it is likely that a resolution to the debate over viewpoint dependency in object recognition requires consideration of dissociable neural subsystems. Both viewpoint-invariant and viewpoint-dependent effects may be observed, depending on which recognition subsystem is given an advantage at test. In the present experiment, presentations of test objects directly to one hemisphere or the other influenced which subsystem was more likely to win the race to guide post visual processing. However, with more typical central-visual-field test presentations, other factors may be critical in determining the relative contributions of abstract and specific subsystems.

One such factor is novelty. Initial memory for an unfamiliar object (a novel exemplar belonging to an unfamiliar category) entails representing at least some new (preexperimentally unfamiliar) visual information. Novel information typically is found in the holistic structures of visual objects, because, at some level, even highly unfamiliar objects can be decomposed into familiar features (e.g., volumes, line segments, pixels). Supporting this idea, we have found that a specific subsystem stores such novel information more effectively than does an abstract subsystem (Burgund & Marsolek, 1999a; Marsolek et al., 1996). Conversely, priming for a familiar object does not necessitate the storage of a new representation but, rather, can be supported by priming of a (now well-learned) preexisting representation. Therefore, using unfamiliar stimuli may create situations in which a specific subsystem contributes to processing to a greater degree than does an abstract subsystem and/or is more likely to win any race that ensues between the two subsystems to guide postvisual processing. In line with this reasoning, viewpoint-dependent performance typically is observed when unfamiliar objects are used as stimuli (e.g., Bülthoff & Edelman, 1992; Edelman & Bülthoff, 1992; Hayward & Tarr, 1997; Srinivas, 1995, Experiment 3; Tarr, 1995; Tarr et al., 1997). Indeed, even when familiar objects are used, performance may vary, depending on the extent of the novelty of the input shape. For example, when noncanonical views of familiar objects are tested, viewpoint-dependent performance is often observed (e.g., Lawson & Humphreys, 1996, 1998; Srinivas, 1993, 1995).[4] Similarly, less familiar objects (e.g., pliers or a corkscrew) may have a greater likelihood of producing viewpoint-dependent effects than do more familiar objects (e.g., a chair or a book; Bartram, 1976, Experiment 2). Finally, studies that intermix stimuli with varying degrees of familiarity (e.g., Lawson & Humphreys, 1996, 1998; Srinivas, 1993, 1995) may observe viewpoint-dependent performance in the means across all of the objects tested (including the highly familiar ones), because the inclusion of the less familiar objects may create a situation or a mental set in which a specific subsystem dominates processing generally.[5]

Several other factors also may be important for determining the relative contributions of abstract and specific subsystems. Previous research has demonstrated that the use of visually degraded inputs creates a situation in which a specific subsystem dominates processing; specific priming is less detrimentally affected than abstract priming when the initial-encoding stimuli are presented in a relatively degraded manner (Marsolek, 1999; Marsolek & Hudson, 1999). The distributed-coding, whole-based processing purported for a specific subsystem should give that subsystem the well-known benefit of distributed representations that they are useful for processing *distorted* or *noisy* versions of previously learned input patterns (see, e.g., Hinton et al., 1986), as long as the distortions are not so large as to change the best-matching exemplar (Marsolek & Burgund, 1997). This finding may help to explain why viewpoint-dependent performance is often observed in experiments in which stimuli are masked or otherwise degraded (as is typical in same/different matching paradigms; e.g., Hayward, 1998; Srinivas, 1995), whereas viewpoint-invariant performance is more likely to be observed when stimuli are not masked or otherwise degraded (e.g., Bartram, 1976, Experiment 2).[6] In addition, task demands per se may influence the relative contributions of these subsystems; abstract and specific priming can be independently manipulated through changes in the test tasks used (Marsolek, 1999; cf. Burgund & Marsolek, 1997). Several such factors may influence processing within a single experiment, and thus, careful consideration of the demands that a particular method places on the underlying processing mechanisms may be essential for predicting which subsystems will contribute to performance and, hence, whether viewpoint-invariant or viewpoint-dependent processing will be observed.

Finally, it is important to note that when both abstract and specific subsystems contribute nearly equally to processing, the net result will be measured as viewpoint-dependent performance. Viewpoint-invariant performance should be found only when an abstract subsystem dominates processing. This may explain why reports of viewpoint-dependent effects are more prevalent in the literature than reports of viewpoint-invariant effects.[7]

The present theory also leads to predictions about recognition of objects rotated in the picture plane. A specific subsystem using a whole-based processing strategy to store objects should exhibit viewpoint dependence for objects rotated in the plane, as well as for those rotated in depth, because, in both cases, (viewer-centered) whole-based representations are affected by the manipulation. This prediction for a specific subsystem is in line with those made by many viewpoint-dependent theorists (Bülthoff & Edelman, 1992; Edelman, 1998; Edelman & Bülthoff, 1992; Humphrey & Khan, 1992; Poggio & Edelman, 1990; Tarr, 1995; Ullman, 1989, 1996). However, the predictions for an abstract subsystem are not as clear. To the extent that rotations in the plane disturb the relations between features or the positions of features in viewer-centered space (e.g., *top of* may become *next to* or *bottom of*) to a greater degree than do rotations in depth, there may be a greater likelihood of observing viewpoint-dependent priming effects for planar-rotated objects than for depth-rotated objects (for analogous reasoning, see Hummel & Biederman, 1992). Indeed, recent evidence from our laboratory supports this contention (Burgund & Marsolek, 1999b); when test objects are presented directly to the LH, priming is sensitive to planar-reorientations. Thus, object representations in an abstract subsystem may include spatial relations between stored features.

In conclusion, two relatively independent subsystems appear to underlie object recognition. A specific subsystem operates more effectively than an abstract subsystem in the RH and supports viewpoint-dependent priming. An abstract subsystem operates more effectively than a specific subsystem in the LH and does not support viewpoint-dependent priming. Such a conclusion suggests not only a possible resolution to the viewpoint dependency debate, but also a fundamental property of the underlying neural architecture of vision and memory.

REFERENCES

Ballard, D. H. (1986). Cortical connections and parallel processing: Structure and function. *Behavioral & Brain Sciences*, **9**, 67-120.

Bartram, D. J. (1976). Levels of coding in picture–picture comparison tasks. *Memory & Cognition*, **4**, 593-602.

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

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

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: Reply to Tarr and Bülthoff (1995). *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 1506-1514.

Bülthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences*, **89**, 60-64.

Burgund, E. D., & Marsolek, C. J. (1997). Letter-case-specific priming in the right cerebral hemisphere with a form-specific perceptual identification task. *Brain & Cognition*, **35**, 239-258.

Burgund, E. D., & Marsolek, C. J. (1999a). *Separable mechanisms for initial storage of unfamiliar three-dimensional objects*. Manuscript submitted for publication.

Burgund, E. D., & Marsolek, C. J. (1999b). *When planar reorientations do and do not affect priming for unfamiliar objects*. Manuscript submitted for publication.

Churchland, P. S., & Sejnowski, T. J. (1992). *The computational brain*. Cambridge, MA: MIT Press.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh Computers. *Behavioral Research Methods, Instruments, & Computers*, **25**, 257-271.

Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system*. Cambridge, MA: MIT Press.

Cooper, L. A., Schacter, D. L., Ballesteros, S., & Moore, C. (1992). Priming and recognition of transformed three-dimensional objects: Effects of size and reflection. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 43-57.

Edelman, S. (1998). Representation is representation of similarities. *Behavioral & Brain Sciences*, **21**, 449-498.

Edelman, S., & Bülthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research*, **32**, 2385-2400.

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

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

Farah, M. J. (1990). *Visual agnosia: Disorders of object recognition and what they tell us about normal vision*. Cambridge, MA: MIT Press.

Farah, M. J. (1991). Patterns of co-occurrence among the associative agnosias: Implications for visual object representation. *Cognitive Neuropsychology*, **8**, 1-19.

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.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "greeble" expert: Exploring mechanisms for face recognition. *Vision Research*, **37**, 1673-1682.

Hayward, W. G. (1998). Effects of outline in object recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 427-440.

Hayward, W. G., & Tarr, M. J. (1997). Testing conditions for viewpoint invariance in object recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 1511-1521.

Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. (1986). Distributed representations. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1: Foundations* (pp. 77-109). Cambridge, MA: MIT Press.

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. (1996). Categorical relations in shape perception. *Spatial Vision*, **10**, 201-236.

Hummel, J. E., & Stankiewicz, B. J. (1998). Two roles for attention in shape perception: A structural description model for visual scrutiny. *Visual Cognition*, **5**, 49-79.

Humphrey, G. K., & Jolicœur, P. (1993). An examination of the effects of axis foreshortening, monocular depth cues, and visual field on object identification. *Quarterly Journal of Experimental Psychology*, **46A**, 137-159.

Humphrey, G. K., & Khan, S. C. (1992). Recognizing novel views of three-dimensional objects. *Canadian Journal of Psychology*, **46**, 170-190.

Humphreys, G. W., & Riddoch, M. J. (1984). Routes to object constancy: Implications from neurological impairments of object constancy. *Quarterly Journal of Experimental Psychology*, **36A**, 385-415.

Jolicœur, P. (1990). Identification of disoriented objects: A dual-systems theory. *Mind & Language*, **5**, 387-410.

Koenderink, J. J. (1987). An internal representation for solid shape based on the topological properties of the apparent contour. In W. Richards & S. Ullman (Eds.), *Image understanding 1985–86* (pp. 257-285). Norwood, NJ: Ablex.

Lawson, R., & Humphreys, G. W. (1996). View specificity in object processing: Evidence from picture matching. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 395-416.

Lawson, R., & Humphreys, G. W. (1998). View-specific effects of depth rotation and foreshortening on the initial recognition and priming of familiar objects. *Perception & Psychophysics*, **60**, 1052-1066.

Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience*, **19**, 577-621.

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

Marsolek, C. J. (1995). Abstract-visual-form representations in the left cerebral hemisphere. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 375-386.

Marsolek, C. J. (1999). Dissociable neural subsystems underlie abstract and specific object recognition. *Psychological Science*, **10**, 111-118.

Marsolek, C. J., & Burgund, E. D. (1997). Computational analyses and hemispheric asymmetries in visual-form recognition. In S. Christman (Ed.), *Cerebral asymmetries in sensory and perceptual processing* (pp. 125-158). Amsterdam: Elsevier.

Marsolek, C. J., & Hudson, T. E. (1999). Task and stimulus demands influence letter-case specific priming in the right cerebral hemisphere. *Laterality*, **4**, 127-147.

Marsolek, C. J., Schacter, D. L., & Nicholas, C. D. (1996). Form-specific visual priming for new associations in the right cerebral hemisphere. *Memory & Cognition*, **24**, 539-556.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, **102**, 419-457.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, **9**, 97-113.

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

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, **343**, 263-266.

Poggio, T. A., & Hurlbert, A. (1994). Observations on cortical mechanisms for object recognition and learning. In C. Koch & J. L. Davis (Eds.), *Large-scale neuronal theories of the brain* (pp. 153-182). Cambridge, MA: MIT Press.

Rock, I., & DiVita, J. (1987). A case of viewer-centered object perception. *Cognitive Psychology*, **19**, 280-293.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, **99**, 195-231.

Srinivas, K. (1993). Perceptual specificity in nonverbal priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 582-602.

Srinivas, K. (1995). Representation of rotated objects in explicit and implicit memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1019-1036.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, **25**, 583-592.

Tanaka, K. (1993). Neuronal mechanisms of object recognition. *Science*, **262**, 685-688.

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. H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 1494-1505.

Tarr, M. J., Bülthoff, H. H., Zabinski, M., & Blanz, V. (1997). To what extent do unique parts influence recognition across changes in viewpoint? *Psychological Science*, **8**, 282-289.

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition*, **32**, 193-254.

Ullman, S. (1996). *High-level vision: Object recognition and visual cognition*. Cambridge, MA: MIT Press.

Verfaillie, K., & Boutsen, L. (1995). A corpus of 714 full-color images of depth-rotated objects. *Perception & Psychophysics*, **57**, 925-961.

Wagemans, J., Gool, L., & Lamote, C. (1996). The visual system's measurement of invariants need not itself be invariant. *Psychological Science*, **7**, 232-236.

Warrington, E. K., & Taylor, A. M. (1978). Two categorical stages of object recognition. *Perception*, **7**, 695-705.

NOTES

1. We should note that other theories have posited multiple representations for shape recognition (Ellis & Allport, 1986; Ellis, Allport, Humphreys, & Collis, 1989; Farah, 1990, 1992; Hummel & Stankiewicz, 1996, 1998; Humphreys & Riddoch, 1984; Jolicœur, 1990; Logothetis & Sheinberg, 1996; Poggio & Hurlbert, 1994; Tarr, 1995; Warrington & Taylor, 1978). The present theory may be distinguished from the others by positing that different types of representations have developed for underlying abstract and specific recognition per se, that abstract and specific subsystems operate with different relative efficiencies in LH and RH, respectively, and that viewpoint-invariant and viewpoint-dependent object recognition do not correspond to the goals of the subsystems but, rather, may be natural consequences of the differing processing strategies utilized by them.

2. We theorize that these results from word-processing studies are applicable to object recognition theories for the following reasons. From an evolutionary perspective, it is unlikely that humans developed sub-

systems specialized for processing words (and not objects), given the relatively short time that word recognition has taken place in our species. Rather, it is more likely that word recognition is learned and performed by the same subsystems as those that evolved to subserve object recognition. Moreover, evidence from patients with associative agnosias is consistent with this hypothesis. In an analysis of a large number of cases of associative agnosia, Farah (1990, 1991) found that the observed (and the unobserved) combinations of word, object, and face recognition deficits exhibited by individual patients indicated that two subsystems (and not, e.g., three) are susceptible to damage leading to agnosia, one involved in word and object recognition and the other involved in object and face recognition.

3. Note that, owing to a limitation in the number of images of objects that met our strict criteria for inclusion in this study, we did not include an unprimed condition. Note also, however, that it is likely that this does not raise problems for our purposes. The goal of this experiment was to measure the presence or absence of viewpoint-dependent priming (measured as a difference between performances in the same-orientation primed-condition and the different-orientation-primed condition), which did not require measuring performance in an unprimed condition. We did not directly test whether (more invariant) priming occurs in the different-orientation-primed condition, in part because previous studies already have demonstrated that it does occur generally when different-orientation-primed conditions are compared against an unprimed condition (e.g., Biederman & Gerhardstein, 1993; Srinivas, 1993, 1995).

4. Reports of patients with RH parietal lesions who are severely impaired at processing noncanonical views of familiar objects but who process canonical views with relative ease (Humphreys & Riddoch, 1984; Warrington & Taylor, 1978) lends further support to this idea (but see also Humphrey & Jolicœur, 1993).

5. It is important to note that this hypothesis should not be understood as contradicting findings that greater stimulus familiarity produces greater ability to recognize specific exemplars and/or greater whole-based processing of the stimuli (often referred to as *expertise effects*; e.g., Gauthier & Tarr, 1997; J. W. Tanaka & Sengco, 1997). In cases demonstrating expertise effects, participants distinguish *highly similar* exemplars (evidently, through whole-based processing) and are able to do so only after sufficient training (or, possibly, sufficient evolutionary development; e.g., with faces). We hypothesize that sufficient training in a specific subsystem is necessary to support these effects; otherwise, the distributed representations in such a subsystem would (incorrectly) treat such similar exemplars as slightly noisy versions of the same exemplar. Our hypothesis in the text is that, after one or two exposures of one novel exemplar from an unfamiliar category, a specific subsystem should store the shape information more effectively than does an abstract subsystem (likely owing to the kind of short-lived representation that can support priming for novel information, rather than to the kind of well-established shape representation that supports expertise effects).

6. Biederman and Gerhardstein (1993) observed viewpoint-invariant performance for unfamiliar objects, using a masked matching paradigm. This at first may seem to counter the present theory. However, as was pointed out by Hayward and Tarr (1997), the different viewpoints used in Biederman and Gerhardstein's (1993) experiment were essentially mirror reversals of each other. Since, typically, mirror reversal transformations have little consequence for object recognition performance (Biederman & Cooper, 1991; Cooper, Schacter, Ballesteros, & Moore, 1992), viewpoint-dependent performance should not be expected.

7. In this analysis, we did not attempt to address studies demonstrating viewpoint-dependent explicit memory, for several reasons. First, explicit memory for both familiar and unfamiliar objects typically is found to be viewpoint dependent (e.g., Edelman & Bülthoff, 1992; Humphrey & Khan, 1992; Rock & DiVita, 1987; Srinivas, 1995); we know of no cases of viewpoint-invariant explicit memory. Thus, viewpoint effects in explicit memory are not subject to the same controversy as those found with other (more implicit) memory measures. In addition, it is likely that explicit memory for objects involves interactions of visual-form subsystems with the hippocampus and related areas (see, e.g., N. J. Cohen & Eichenbaum, 1993; McClelland, McNaughton, & O'Reilly, 1995; Squire, 1992), and these interactions may produce different effects from those observed with "pure" priming measures supported by visual-form subsystems alone (see Marsolek et al., 1996).

(Manuscript received May 28, 1998;
revision accepted for publication September 14, 1999.)