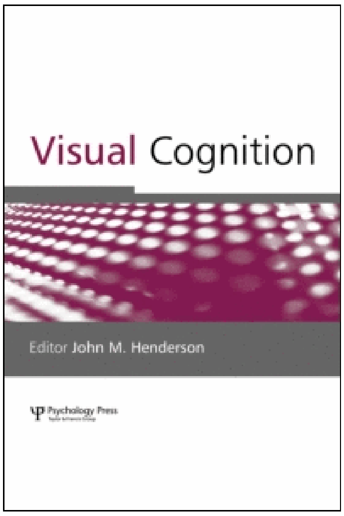

Visual Cognition
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683696

Priming of plane-rotated objects depends on attention and view familiarity

Volker Thoma [a]; Jules Davidoff [b]; John E. Hummel [c]
[a] School of Psychology, University of East London, UK [b] Department of Psychology, Goldsmiths University of London, UK [c] Department of Psychology, University of California, Los Angeles, USA

Online Publication Date: 01 February 2007

To cite this Article Thoma, Volker, Davidoff, Jules and Hummel, John E.(2007)'Priming of plane-rotated objects depends on attention and view familiarity',Visual Cognition,15:2,179 — 210

To link to this Article: DOI: 10.1080/13506280500155627

URL: http://dx.doi.org/10.1080/13506280500155627

VISUAL COGNITION, 2007, 15 (2), 179–210

Psychology Press
Taylor & Francis Group

Priming of plane-rotated objects depends on attention and view familiarity

Volker Thoma

School of Psychology, University of East London, UK

Jules Davidoff

Department of Psychology, Goldsmiths University of London, UK

John E. Hummel

Department of Psychology, University of California, Los Angeles, USA

Three experiments investigated the role of attention in visual priming across rotations in the picture plane. Experiment 1 showed that naming latencies increased with the degree of misorientation for objects commonly seen in an upright view (base objects) but not for objects seen familiarly from many views (no-base objects). In Experiment 2, no-base objects revealed a priming pattern identical to that observed previously for left–right reflections (Stankiewicz, Hummel, & Cooper, 1998): Attended objects primed themselves in the same and rotated views, whereas ignored images primed themselves only in the same view, with additive effects of attention and orientation. In Experiment 3 ignored base objects only primed themselves in a familiar (upright) view, indicating that priming only obtains when that image makes contact with object memory. These data challenge theories of object recognition that rely on any single representation of shape and contribute to evidence suggesting holistic (view-like) representations for ignored and analytic (view-insensitive) representations for attended objects.

The human capacity for visual object recognition is characterized by a complex pattern of strengths and limitations that defy explanation in terms of simple "one size fits all" accounts of the representation of object shape or the processes that generate those representations and match them to object memory. For example, consider our apparent ability to recognize familiar objects in novel views; this led many researchers (Biederman, 1987; Clowes, 1967; Hummel & Biederman, 1992; Marr & Nishihara, 1978; Palmer, 1977;

Please address all correspondence to: Volker Thoma, School of Psychology, University of East London, The Green, London E15 4LZ, UK. Email: v.thoma@uel.ac.uk

http://www.psypress.com/viscog DOI: 10.1080/13506280500155627

Sutherland, 1968) to postulate that we represent objects as *structural descriptions*. Structural representations specify an object's parts, typically volumetric parts, such as geons (Biederman, 1987) or generalized cylinders (Marr & Nishihara, 1978), in terms of their spatial relations to one another. A structural description of a coffee mug might represent its shape (approximately) as a "curved cylinder side-attached to a vertical straight cylinder" (see, e.g., Biederman, 1987). This description does not specify—and so does not vary with—the angle and distance from which the mug is viewed (barring "accidental" views, such as viewing the mug at an angle that hides the curved cylinder and projects the straight cylinder as a simple rectangle). Because of these properties structural descriptions seem to provide a natural account of our ability to recognize objects in a variety of viewpoints. However, structural descriptions have been criticized as a too powerful account of human shape perception in that they predict greater invariance with changes in viewpoint than the human visual system actually exhibits (e.g., Tarr & Bülthoff, 1995).

Alternative accounts to structural descriptions have argued that although we can recognize objects in many novel viewpoints, we are nonetheless faster and more accurate to recognize objects in some views than others. In consequence, a number of researchers concluded that instead of generating and matching parts-based structural descriptions we recognize objects by matching object images to specific holistic *views* in long-term memory (e.g., Edelman & Intrator, 2003; Poggio & Edelman, 1990; Tarr & Bülthoff, 1995). These view-based models account for some of the view-sensitivities of human object recognition (e.g., Bülthoff & Edelman, 1992; Tarr & Pinker, 1989, 1990), but they have difficulty accounting for its invariances (see Biederman & Gerhardstein, 1995). In addition, purely holistic accounts of shape perception, such as view-based models, have difficulty in explaining the role of spatial relations in shape perception (Hummel, 2000), our ability to make judgements about one aspect of an object's shape (e.g., its aspect ratio) independent of other aspects (e.g., axis curvature; Saiki & Hummel, 1998; Stankiewicz, 2002), and accounting for the role of visual attention in shape perception (Stankiewicz, Hummel, & Cooper, 1998). These facts suggest that structural descriptions play a central role in the visual representation of object shape. Yet, at the same time there are further empirical facts that are clearly inconsistent with a purely structural description-based account of human object recognition (Hummel, 2001; Hummel & Stankiewicz, 1996).

According to recent theories (e.g., Hummel & Biederman, 1992; Hummel & Stankiewicz, 1996), generating a structural description from an object's image requires time and visual attention. In consequence, if object recognition were based strictly on structural descriptions of object shape, then it should likewise be time consuming and demanding of visual attention. It is neither, at least for objects depicted in familiar views. Potter (1976) and Intraub (1981) showed that people are capable of recognizing common objects as rapidly as 10 per second; similarly, Thorpe, Fize, and Marlot (1996) showed that people can recognize animals in cluttered displays in exposures lasting less than 100 ms, and Oram and Perrett (1992) observed that face-selective neurons in macaque IT respond to their preferred stimuli within 100–110 ms of that stimulus appearing on the animal's retina. Such rapid processing leaves too little time for the recurrent and feedback flow of information necessary to generate a complex description of an object's parts in terms of their spatial relations (Hummel & Stankiewicz, 1996).

Purely structural description-based accounts of object recognition are also problematic with respect to the attentional demands of object recognition. Generating a structural description requires visual attention in order to bind visual features into parts and to bind parts to their relations (Hummel, 2001; Hummel & Biederman, 1992; Hummel & Stankiewicz, 1996; Logan, 1994; Stankiewicz et al., 1998; Thoma, Hummel, & Davidoff, 2004). However, ignoring an object image on one occasion can prime recognition of that same object on a subsequent occasion (Stankiewicz & Hummel, 2002; Stankiewicz et al., 1998; Thoma et al., 2004; Tipper, 1985; Treisman & DeSchepper, 1996) indicating that object recognition does not necessarily require visual attention in all cases.

Taken together, there is evidence for both structural and view-based representations in human visual object recognition—and there is evidence that neither of them is sufficient to explain all its properties. Recently, theorists are increasingly discussing the possibility of both types of representations processing object shape independently (Foster & Gilson, 2002; Hayward, 2003; Hummel & Stankiewicz, 1996; Tarr & Bülthoff, 1995). So far only one of these approaches is considering the role of attention in relation to the properties of structural and view-based representations. In this paper, we will describe the hybrid model of object recognition by Hummel (2001) because of its detailed predictions for object recognition after view changes such as plane rotations. However, many of the theoretical issues and the experiments described here will also concern traditional models of object recognition as well as the general approach of proposing multiple formats of representation.

A HYBRID MODEL OF OBJECT RECOGNITION

The fact that both structural descriptions and view-based representations of shape can account for some, but not all of the properties of object recognition led Hummel and Stankiewicz (1996; Hummel, 2001) to propose that objects are recognized on the basis of a hybrid representation of shape, consisting of a holistic (i.e., "view"-like) representation working in parallel with an analytic representation (i.e., a structural description). The model (JIM.3) consists of an eight-layer artificial neural network that can be trained to recognize line drawings of objects. In the first three layers, units represent local image features (coding contours, vertices, etc. into surfaces). Gating units in layer 4 project the output of layer 3 to layer 5, which is divided into two components: An analytic component (i.e., a structural description) and the holistic surface map (HSM) representing the shape attributes of an object's surfaces (as coded in layer 3). In layers 6–8 the representation generated by the activation patterns of the units in layer 5 are encoded into long-term memory.

The analytic representation codes an object's shape explicitly in terms of the categorical interrelations among its parts. This representation has the properties of a structural description and is largely robust to many variations in viewpoint (such as translation, changes in scale, left–right reflection and some rotations in depth) but it is sensitive to rotations in the picture plane (see Hummel & Biederman, 1992). Furthermore, it also allows generalization across metric variations in object shape, generalization to novel views and to novel exemplars of known categories, reflecting the desirable properties of a structural description. However, it requires processing time and visual attention to represent parts and spatial relations independently of the parts they relate (Hummel, 2001; Hummel & Biederman, 1992).

The holistic representation, in contrast, does not specify an object's parts or their categorical spatial relations independent of each other. Instead, an object's parts are represented in terms of their topological positions in a 2-D coordinate system (see Hummel, 2001). Since the holistic representation does not require attention for the dynamic binding of parts to their relations, it can be generated rapidly and automatically. However, as the units representing surface attributes are spatially separated, the representation formed on the surface map is sensitive to left–right reflections as well as to rotations in the picture plane and in depth. However, the HSM representation is invariant with translation and scale. Although the surfaces' topological relations are maintained in the mapping from layer 3 to the HSM, their absolute locations in the visual field and their size in the image are not (this is because the holistic representation receives its input from units in layer 4 which are distributed spatially to cover the whole visual field). Thus, the units of the HSM are confined neither to a particular location nor to a receptive field size, which allows them to "shrink-wrap" on a given object, no matter where it is in the visual field.

The holistic representation permits rapid, automatic recognition of familiar objects in familiar views, but it is sensitive to variations in viewpoint (specifically, to rotations and reflections). For example, a holistic representation of a horse would be matched, in its entirety, against an object's image to determine the degree of fit between the image and the holistic representation (i.e., view) in memory: The coordinates of the features in the viewed image would be matched to the coordinates of the features in the stored view and the degree of fit would be computed as a function of the vector similarity of the coordinates of corresponding features. The matching process is analogous to laying a template for one view of the horse over its rotated version and counting the points of overlap. By this holistic measure of similarity, the upright and rotated horse images are very different because few (if any) corresponding features reside in equivalent locations in the two images. By contrast, in an analytic representation of shape the upright and rotated images of the horse are still highly similar because they depict many of the same parts, which in turn should yield analytic priming. However, many of the spatial relations have changed, for example, after a 90° rotation the legs may now appear as "side-attached to" (rather than "below") the torso. Because of the mismatch of spatial relations priming for analytic representations should be reduced compared to view-changes where the analytic representation is not affected, e.g., mirror-reflection (Stankiewicz et al., 1998).

The hybrid holistic/analytic model predicts a complex pattern of relationships between visual attention and visual priming as a function of variations in viewpoint (and other manipulations of an object's image). These predictions derive from the fact that the model represents attended images both analytically and holistically, whereas it represents ignored images only holistically. As such, it predicts that visual priming for attended images should reflect the properties of both representations, whereas priming for ignored images should reflect the properties of the holistic representation alone. Specifically, it predicts that attended images should visually prime themselves, as well as translated, scaled, left–right (mirror) reflected versions of themselves, and even configural distortions of themselves (e.g., in which the image is split down the vertical midline and the left and right halves switch places). Ignored images should prime themselves, translated and scaled versions of themselves but not their mirror reflections or configural distortions of themselves (see Hummel, 2001, and Thoma et al., 2004, for more detailed elaborations of these predictions). These predictions were tested (Stankiewicz & Hummel, 2002; Stankiewicz et al., 1998; Thoma et al., 2004), and the findings were exactly as predicted by the model.

Functional imaging studies (e.g., Vuilleumier, Henson, Driver, & Dolan, 2002) also support the notion that two types of object representations can be distinguished according to view invariance in priming tasks. Vuilleumier et al. (2002) showed that repetition of images of common objects decreased activity (i.e., showed priming) in the left fusiform area independent of viewpoint (and size), whereas a viewpoint-dependent decrease in activation was found in the right fusiform area. Interestingly, the latter area was sensitive to changes in orientation but not in size—properties of the holistic component directly predicted by the hybrid model (Hummel, 2001) and confirmed in behavioural studies (Stankiewicz & Hummel, 2002).

MOTIVATION FOR THE CURRENT EXPERIMENTS

The present experiments return to the question of the role of attention in visual priming across changes in viewpoint, specifically, to the case of rotations about the line of sight. Plane rotated objects were used to further test the general notion of a hybrid model consisting of both a structural (analytic) and a view-based (holistic) representation. Previously, Stankiewicz et al. (1998) tested and confirmed the predictions of the hybrid model using an object naming task with paired prime/probe trials. Attended images reliably primed both themselves and their left–right reflections. However, ignored images only primed themselves in the same view. The priming advantage for same view prime–probe trials was equivalent in both attended and unattended conditions (about 50 ms) and was credited to the contribution of the holistic component.

Left–right reflection, as studied by Stankiewicz et al. (1998), may be an unusual or unrepresentative change in viewpoint. For example, rather than a change in the view of the 3-D object itself, the visual system may instead interpret mirror reflection as a 2-D flip of the 2-D image itself (Davidoff & Warrington, 2001; Murray, 1997). More important, it is not clear whether viewpoint changes in the picture plane can actually be accounted for in the same way as mirror reflections within Hummel's (2001; Hummel & Stankiewicz, 1996) version of the hybrid model.

Object recognition is well-known to be sensitive to orientation in the picture plane (for a review, see Lawson, 1999). People are slower and more error prone to name some objects (specifically, objects with a canonical upright orientation) rotated away from the upright (e.g., presented upside down, or lying on their side) than to name them presented in their canonical upright orientation (e.g., Jolicoeur, 1985; Tarr & Pinker, 1989, 1990). Moreover, naming response times (RTs) get longer, and errors more frequent, as the image is rotated further from the upright (up to about 120°, at which point RTs and errors continue to decrease to 180°, at least for some objects).

Hummel's (2001; Hummel & Stankiewicz, 1996) version of the hybrid model—like its purely structural description-based predecessor (Hummel & Biederman, 1992; see also Hummel, 1994)—accounts for this picture-plane rotation effect primarily in terms of the effects of picture-plane rotations on the mental representation of the relations among an object's parts. For example, if some part of an object (say, the handle of a bucket) is usually on top of (above) some other part (say, the container part of the bucket) then it will be that way in the long-term (LTM) representation of that object. However, if the image of the bucket is rotated 45° clockwise, then the handle will appear both above and beside the container. The structural description of the resulting image will mismatch the representation of the bucket in LTM because it includes the spurious beside relation. In consequence, the mismatch impairs the model's ability to recognize the object as a bucket. If the image is rotated an additional 45–90° off upright, then the description is further distorted to simply *beside* (handle, container), which mismatches the LTM representation on two relations (i.e., it lacks *above*, and it includes *beside*), further impairing recognition. Rotated another 45–135° off upright, the description is distorted further to *below-and-beside* (handle, container), which mismatches the LTM representation on three relations (it lacks *above*, it includes *below*, and it includes *beside*), further impairing recognition. Finally, if the image is rotated all the way to 180° off upright, the description changes to simply *below* (handle, container), which mismatches the LTM representation on only two relations (it lacks *above* and it includes *below*, but it no longer includes *beside*), so performance improves relative to the 135° case (see Hummel, 1994; Hummel & Biederman, 1992).

In this way, the Hummel and Biederman model and the hybrid model that evolved from it (Hummel, 2001; Hummel & Stankiewicz, 1996) account for the effects of picture-plane orientation on object recognition, and even account for the characteristic "dip" in RT and errors between 120° and 180°. In addition, in the case of the hybrid model, picture-plane rotations also impair recognition by distorting the holistic representation because holistic representation of a rotated object will tend to have little or no overlap with a holistic representation of that same object in an upright orientation. Indeed, picture-plane rotations have more catastrophic effects on the holistic representation than on the analytic one: Even a comparatively small rotation (e.g., 45°) can cause the match between a holistic representation of an object image and a holistic representation in LTM to drop to virtually zero.

When considering plane rotation effects, an important empirical fact is that not all objects show the "characteristic" rotation function (Vannucci & Viggiano, 2000; Verfaillie & Boutsen, 1995). When observers rated the goodness of views for common objects, one cluster did not yield a preferred "upright" orientation (Verfaillie & Boutsen, 1995). The resulting distinction between *base* objects (objects with a preferred upright; examples include animals, houses, furniture, etc.) and *no-base* objects (objects with no preferred upright, such as hammers, forks, etc.) has been found to have importance for both behavioural (Vannucci & Viggiano, 2000) and neuropsychological (Davidoff & Warrington, 1999) investigations of object orientation. For example, Vannucci and Viggiano (2000) demonstrated that no-base objects (e.g., hammer) are recognized equally well at all orientations, and Davidoff and Warrington (1999) showed differential performance after brain damage in matching inverted objects according to whether they were base or no-base objects. These two types of objects were used to test view-dependent priming within the hybrid model.

In order to better understand the logic of our predictions, it is important to note at which level visual priming operates and how it affects LTM. Biederman and Cooper (1991) showed that the locus of visual priming, at least for attended objects, is at the level of the representation of an object's parts and their interrelations—or more specifically, as discussed in detail by Cooper, Biederman, and Hummel (1992), at the level of the mapping between the representation of parts and relations and the representation of the complete *object model* (i.e., the presumably relatively localist visual representation of an object's complete shape) in LTM. Integrating these considerations with the hybrid representation the idea is that priming reflects learning the mapping from the analytic *and* holistic representations of object shape to the representation of the object model in LTM. This is instantiated in the hybrid model (Hummel, 2001) as strengthening the connections from layer 5 (the hybrid representation of shape) through layer 8 (the localist representation of the object model).

PREDICTIONS OF THE HYBRID MODEL

In general, the hybrid model predicts that attended objects prime themselves and their plane-rotated versions, whereas ignored objects prime themselves only in the same view, as long as this view is familiar. Both the analytic and the holistic components of the hybrid model are sensitive to rotation in the picture plane, although the holistic component is likely to be much more so. The analytic representation suffers from substantial plane rotations if these change the spatial relations between an object's parts. The holistic representation, in contrast, is view-sensitive because after plane rotations (just as after mirror-reflections) the 2-D coordinates of features (e.g., surfaces) are now different from the original view. Thus, the hybrid model predicts priming for attended objects (which are represented both analytically and holistically) should be sensitive to changes in picture-plane orientation; more specifically, the model predicts that the magnitude of visual priming for rotated prime–probe pairs should systematically diminish with the degree of orientation difference between prime and probe. The predictions for priming for ignored objects are more interesting. Priming should be very sensitive to changes in orientation such as plane rotation, because ignored objects are represented only holistically. But although the holistic representation of both no-base and base objects should be equally affected by changes in rotation, the model predicts differences in which orientations allow priming for ignored objects in the first place.

In Experiment 2, for no-base objects, we predict that an ignored prime image at some orientation, x, should prime recognition of an identical probe image (i.e., a probe at orientation x), but should not prime recognition of the same object at any other orientation, y. In Experiment 3, we use only base objects, and for these we predict that an ignored upright prime image should prime recognition of an identical probe image (as has been observed numerous times in the work of Stankiewicz, Thoma, and colleagues), but an ignored misoriented image should not prime recognition of *anything*, even itself. The reason, according to the hybrid model, is that (a) base objects are represented in LTM only in the upright orientation, and (b) priming resides in the *mapping* from the representation of object shape (which, in the case of an ignored image, is only the holistic representation) to the stored representation in LTM. Since base objects in the hybrid model only have upright representations in LTM, a misoriented ignored prime image, failing to match anything in LTM, simply has nothing to prime. As a result, it should not even prime itself.

In testing these predictions an immediate concern using plane-rotated objects in a priming paradigm is that differences in baseline responding complicate assessing the magnitude of priming. Fortunately, no-base objects are recognized with approximately equal facility at all orientations in the picture plane (Vannucci & Viggiano, 2000; Verfaillie & Boutsen, 1995), so in addition to allowing us to test some of the subtler predictions of the hybrid model, these objects also provide a stable baseline for measuring the magnitude of visual priming. Experiment 1 was designed to verify whether no-base objects do in fact yield equivalent naming latencies in different orientations. Experiments 2 and 3 tested the predictions of the hybrid model with regard to the effects of attention and picture-plane orientation on short-term visual priming of no-base and base objects.

EXPERIMENT 1: NAMING OF BASE AND NONBASE OBJECTS ROTATED IN THE PICTURE PLANE

Experiment 1 was designed to extend Vannucci and Viggiano's (2000) demonstration that no-base objects (e.g., hammer) are recognized equally well at all orientations to the case of visual priming. Specifically, Experiment 1 tested whether the pattern of results observed by Vannucci and Viggiano in an object decision task also hold for the speeded naming tasks to be used in our Experiments 2 and 3. In addition, Vannucci and Viggiano did not equate their sets of base and no-base objects for familiarity and visual complexity, (Snodgrass & Vanderwart, 1980), so either or both of these factors might have affected the pattern of performance. We predict that, even controlling these factors, recognition latencies for base objects (such as animals and houses) will increase with rotation away from the canonical (upright) view, whereas no-base objects (such as hammers and keys) will be identified equally well in all orientations. Finding sets of objects for which recognition performance is independent of orientation will be essential to our tests of the hybrid model in Experiment 2.

Method

Participants. Twenty-nine native English speakers with normal or corrected-to-normal vision participated for credit in introductory psychology courses at Goldsmiths College University of London.

Materials. The experimental program was generated in E-Prime 1.0 (PSN). Participants sat approximately 90 cm from the screen. Three subsets of 24 images (animals, base objects, and no-base objects, see Appendix) were taken from Snodgrass and Vanderwart (1980). The base and no-base object sets were matched for familiarity (means: 3.75 vs. 3.60; max = 5) and visual complexity (means: 2.64 vs. 2.77; max = 5) according to the norms obtained by Snodgrass and Vanderwart. The means for familiarity and visual complexity of the animal set were 2.69 and 3.70; it was not possible to match animals to the other two sets. For each object, the standard view (as obtained from the original set) was assigned as the 0° view. Clockwise rotations in the picture plane resulted in 60° and 120° orientations for each object (see Figure 1). Participants saw a given object only once during the experiment in one of the three orientations. The allocation of objects to the experimental conditions was randomized for each participant. Thus, there were eight different objects in each of the three object sets for each of the three orientation conditions (0°, 60°, and 120°) resulting in a total of 72 trials per participant.

Procedure. The participants first read instructions which they paraphrased back to the experimenter. After four practice trials with objects not chosen from the experimental sets, participants were asked whether they had any questions. Each subsequent test trial was initiated by the participant. A trial began with an unfilled circle (subtending 0.032° of visual angle) in the centre of the screen that was replaced by the participant's key press with a fixation cross for 495 ms. An object (subtending 4.57° of visual angle) was then shown in the centre of the screen for 195 ms followed by a single pattern mask for 495 ms. The participant's task was to name the object as quickly and as accurately as possible. After the response, a feedback display with the name of the object and the response time was shown. At the end of each

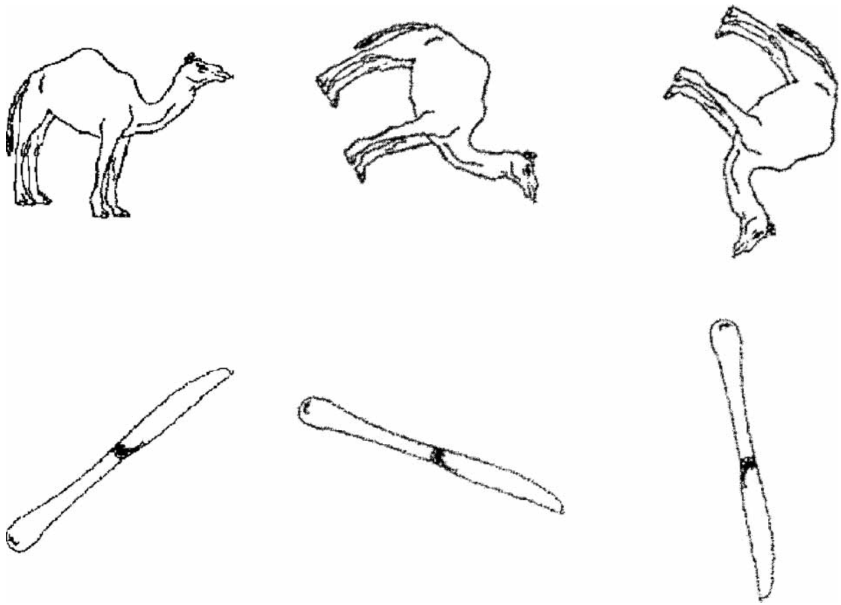

Figure 1. Examples of object images used in Experiment 1.

trial, the experimenter used the keyboard to record the participant's accuracy and any voice key errors.

Results

The overall error rate was 6.32% (of which 1.1% were voice key errors elicited by the subject). Response times for correct trials and error rates were submitted to a 3 (object type: Animals vs. base vs. no-base) $\times$ 3 (rotation: 0° vs. 60° vs. 120°) ANOVA. For latencies, there were significant effects of object type, $F(2, 56) = 33.41$, $MSE = 6020.73$, $p < .001$, rotation, $F(2, 56) = 21.84$, $MSE = 5201.83$, $p < .001$, as well as for the interaction, $F(4, 112) = 4.15$, $MSE = 6650.41$, $p < .01$ (see Figure 2 and Table 1). Overall, no-base objects showed no effects of rotation, whereas naming RTs for base objects and animals increased with greater rotation. A similar pattern was found for errors: Significant effects were again found for object type, $F(2, 46) = 15.75$, $MSE = 0.46$, $p < .001$; rotation, $F(2, 46) = 12.06$, $MSE = 0.39$, $p < .001$, and for the interaction, $F(4, 92) = 9.07$, $MSE = 0.36$, $p < .001$.

For latencies, post hoc comparisons using Tukey's HSD test revealed that, for animals, only the increase in response times from 0° compared to 120° rotation was significant ($p < .01$; all other ps $> .33$). Similar analyses revealed that, for artificial base objects, the RT differences between 0° and

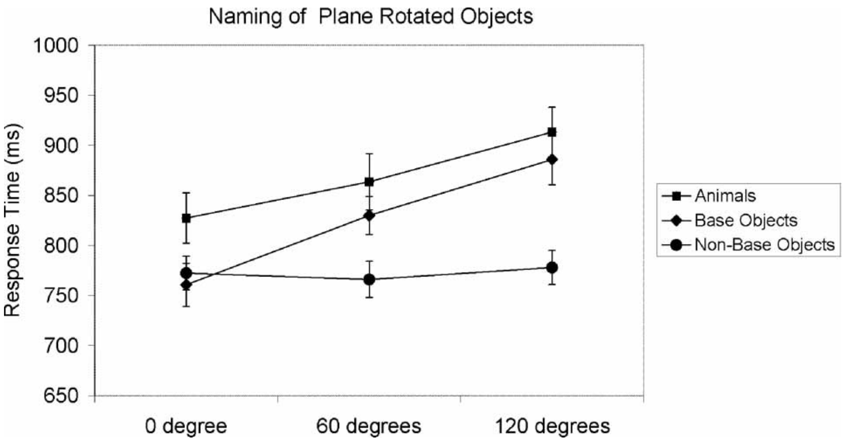

Figure 2. Response time means and standard errors for Experiment 1 as a function of the degree of rotation in the picture plane and the type of objects ($n = 29$).

60° ($p < .05$) and between 0° and 120° ($p < .001$) were significant, but not the difference between 60° and 120° ($p > .19$). For no-base objects, however, there were no conditions in which the RT differences even approached significance (all ps $> .9$). Error rates were small and post hoc comparisons only revealed effects for animals. There were significantly more errors for naming animals shown rotated 120° compared to 0° as well as compared to 60° conditions ($p < .001$). There were no significant differences in error rates over changes in orientation for artificial base objects (all ps $> .1$) and no-base objects (all ps $> .9$).

An additional ANOVA was run for RTs with items as a random factor (across subjects) and it revealed a similar pattern, with significant main effects of object type, $F(2, 69) = 5.4$, $MSE = 38{,}285.92$, $p < .01$, and rotation, $F(2, 138) = 16.4$, $MSE = 7014.79$, $p < .001$, as well as a significant interaction between them, $F(4, 138) = 2.7$, $MSE = 7014.79$, $p < .05$.

TABLE 1
Mean response times (RT in milliseconds), standard errors and errors (frequency and percentage errors) for conditions in experiment 1

	Animals			*Base Objects*			*No-Base Objects*		
	0°	*60°*	*120°*	*0°*	*60°*	*120°*	*0°*	*60°*	*120°*
RT	827	863	913	760	829	885	772	766	778
SE	25	28	24	21	19	25	16	18	17
errors	9	14	40	6	11	19	6	6	1
& errors	4	6	17	3	5	8	3	3	0

Discussion

Experiment 1 clearly demonstrated that orientation in the picture plane had differential effects on base and no-base objects. It was only artificial base objects and animals that incurred increasing recognition costs when they were rotated from their standard view. Even though the subset of animal pictures could not be matched for familiarity and visual complexity with other base objects, and were, in general, harder to identify, they showed the same pattern of increasing response times and similar trend of increasing error rates as artificial base objects. In contrast, no-base objects were equally recognizable in all picture-plane orientations. Thus, Experiment 1 extends the findings of Vannucci and Viggiano (2000), who used an object decision task, to the case of naming RTs and errors.

The aim of Experiment 1 was not to distinguish between structural description and view-based accounts of object recognition. The results would clearly be predicted from the latter account according to which objects are stored in correspondence with the familiarity of different viewpoints. For a similar reason, the results would also be predicted from the structural description account. Both base and no-base objects contain relations that could be coded for the object's relative positions in terms of "on-top-of" or "below-of" and that operation would automatically be carried out in a structural representation, such as the analytic route of Hummel's hybrid model. However, as we usually see no-base objects—but not base objects—in multiple orientations routinely, we also encode them that way in LTM. It would be access to those multiple representations in LTM that would be required for recognition and allow equivalent naming latencies for no-base objects.

The goal of Experiment 1 was to verify that no-base objects would provide images to test the predictions from the hybrid theory of Hummel and Stankiewicz (1996; Hummel, 2001). Priming studies, such as those of Stankiewicz et al. (1998) and Thoma et al. (2004) require comparisons against a baseline. No-base objects will provide a means of accurately matching baseline latencies for different rotations of an object because we have now shown that these are identical for this class of object.

EXPERIMENT 2: THE EFFECTS OF VIEWPOINT AND ATTENTION ON PRIMING

The aim of Experiment 2 was to extend the priming results of Stankiewicz et al. (1998) to picture-plane rotations. The hybrid model postulates two qualitatively different representations (or processing routes) feeding into LTM. One of these (holistic) clearly does not generalize (i.e., shows view-dependent recognition performance or priming) over plane rotation but there are recognition costs also associated with the analytical route (Hummel & Biederman, 1992). Thus, most object recognition accounts, including the hybrid model, would make the same predictions for attended objects: Substantial variation in picture-plane rotation from prime to probe should weaken priming for attended objects. However, predictions for prime–probe variation for ignored objects differ between models. In the ignored condition, structural description accounts (Biederman, 1987) would predict no priming for ignored objects whatever their orientation because attentional processes are necessary to code parts and their relations into structural descriptions (Hummel & Biederman, 1992). Multiple views accounts also do not allow strong predictions for the ignored conditions as, likewise, they do not explicitly incorporate attention but it is unlikely that these accounts would predict access to stored object representations for unattended stimuli (e.g., see Olshausen, Anderson, & van Essen, 1993).

We make the following predictions based on the hybrid model. For ignored no-base objects we expect only same view priming. For attended objects, we expect priming from both identical and rotated views. The priming would be greater for identical views as shown in many studies (e.g., Warren & Morton, 1982). However, unlike in Stankiewicz et al. (1998), the Hummel model should now predict an interaction between attention and rotation. There should still be a predicted 50 ms priming cost for rotated views in the ignored and attended condition but there should also be an additional cost in the attended condition with the analytic priming component reduced after plane rotations (due to the mismatch of spatial relations). We examine these predictions using a priming paradigm similar to Stankiewicz et al. in Experiment 2.

Method

Participants. Thirty native English speakers with normal or corrected-to-normal vision participated for credit in introductory psychology courses at Goldsmiths College University of London.

Materials. We used 56 no-base images from the Snodgrass and Vanderwart (1980) set (see Appendix). They included the ones used in Experiment 1 (except for two items: Gun and ring). Of these, 32 were filler items that were never used as probes (e.g., used as the "ignored" item in the prime display of trials in which attended prime objects were repeated in the probe). The 24 critical (probe) items were counterbalanced across participants by placing each object in one of six clusters of four objects. Each cluster (and thus, each object) was placed into one of six conditions (attended-same, attended-rotated, ignored-same, ignored-rotated, unprimed-same-view, and unprimed-rotated-view). Thus, an object appeared in only one trial for a given subject. The target objects appeared in all six conditions equally often across participants. Prime and probe objects were shown in the standard view (as in the original Snodgrass and Vanderwart set) or rotated 90° clockwise in the picture plane. The two views appeared in all conditions equally often.

Procedure. The basic procedure followed the paradigm of Stankiewicz et al. (1998). Specifically, trials were presented in prime–probe pairs, with the probe trial immediately following the corresponding prime trial. Prime trials presented two images to the left and right of fixation, one of which (the attended prime) was precued, and which it was the subject's task to name; the other (ignored) prime was nominally irrelevant to the subject's task. The probe task presented a single image at fixation. The subject's task was to name the probe image. The critical conditions manipulate the relationship between the prime and probe images, as detailed below.

The ordering of the trials and the pairing of attended and ignored objects on prime trials were randomized for each participant. The participants first read the instructions, which they paraphrased back to the experimenter. The participants then read a list of names of objects that would appear in the experiment. There were six practice trials with a set of objects different from the experimental set. After the practice trials, the computer displayed "End of Practice", and the participants were asked whether they had any questions. Each experimental trial began with an unfilled circle (subtending 0.032° of visual angle) in the centre of the screen that was removed by the participant's key press and was replaced with a fixation cross for 495 ms. Participants then saw a white screen briefly for 30 ms followed by an attentional cueing square (4.57° ×4.57°) either to the left or right of the fixation cross at a distance of 4.0°. After 75 ms, images of two different objects were displayed simultaneously on the computer screen for 120 ms; one object was inside the square (the attended image) and the second (ignored) object on the other side of the fixation cross (see Figure 3). Both images were centred 4.0° from the fixation cross. The entire prime display lasted less than 200 ms, a duration that is too short to allow a saccade to either object. After the images disappeared, a 30 ms blank screen was shown followed by a random-line pattern mask displayed for 495 ms covering the entire screen (15.6° of visual angle). Participants named the cued (attended) object as quickly and as accurately as possible. Latencies were recorded by the computer through a voice key attached to a microphone.

After the mask, a blank screen was displayed for 1995 ms followed by a fixation cross (0.032°) displayed for 495 ms. Following a 30 ms blank screen, the probe image was shown in the centre of the screen for 150 ms. In total, 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,

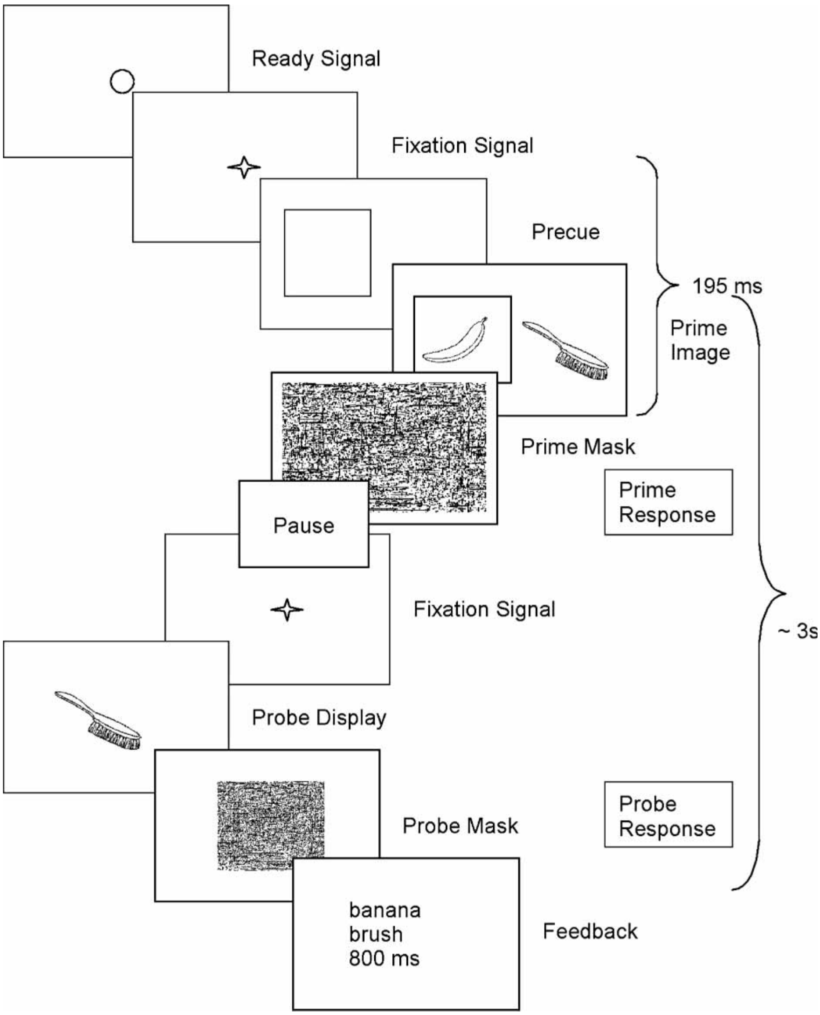

Figure 3. Sequence of displays in Experiment 2.

495 ms for the probe fixation dot, and 30 ms blank). Following the probe display, a single pattern mask (4.57°) was shown in the centre of the screen for 495 ms. The probe object was the attended object (attended conditions), the ignored object (ignored conditions), or a third object not seen previously in the experiment (unprimed baseline condition). In the attended and ignored conditions, half the probes were the same view of the target and half the rotated view. Again the participant's task was to name the probe as quickly and as accurately as possible. Naming was followed by a display of

the names of the attended prime and the probe along with the probe response time. At the end of each trial, the experimenter recorded the participant's accuracy on the prime and probe displays, and all voice key errors. The participant then could initiate the next trial with a key press.

Results

Trials on which either the prime or probe responses were incorrect were excluded from the analysis (13.1%) as were voice key errors and response times above 3000 ms (5.0%). The mean response time for the standard view was 846.3 ms (*SE* 39.5) and 805.5 ms (*SE* 24.95) for the rotated view, a nonsignificant difference, $t(1, 29) = 1.12$, $p > .05$. The mean error rates for the standard view was 13.3% (*SE* 3.3) and 14.1% (*SE* 3.0) for the rotated view. For all conditions, priming was calculated as the difference between each participant's mean latency in the unprimed (baseline) condition and the participant's mean latency in each of the other probe conditions (see Figure 4 and Table 2). A 2 (attention: Attended vs. ignored) $\times$ 2 (rotation: Same view vs. rotated view) within-subjects ANOVA was performed on priming latencies. The analysis revealed a reliable main effect of attention, $F(1, 29) = 10.87$, $MSE = 34522$, $p < .01$, and a main effect of rotation, $F(1, 29) = 14.79$, $MSE = 7692$, $p < .001$. The interaction between attention

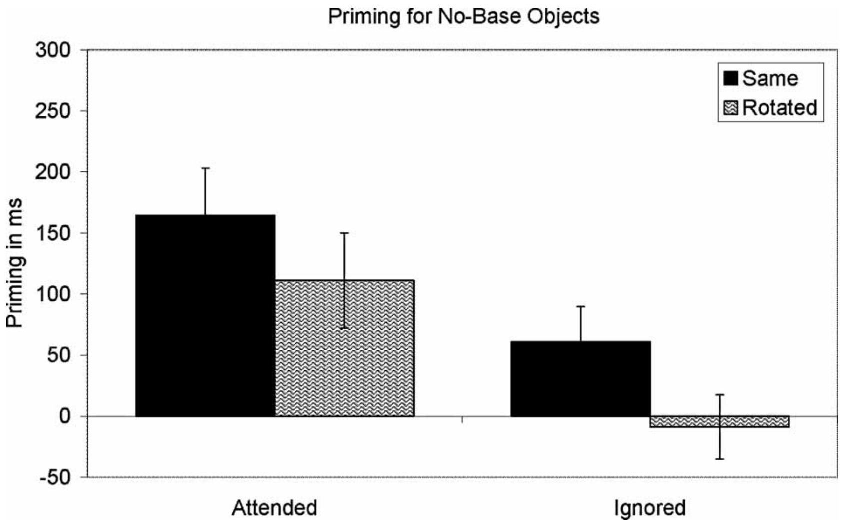

Figure 4. Priming means in response times and standard errors for base objects in Experiment 2 as a function of whether the object was attended or ignored in the prime display prior to the probe and whether the probe objects were presented in the same orientation as the prime image or rotated in the picture plane ($n = 30$).

TABLE 2
Mean response times (RT, in milliseconds), standard errors, and percentage errors for probe objects Experiment 2

Variable	*Attended*		*Ignored*		*Unprimed*	
	Same	*Rotated*	*Same*	*Rotated*	*Same*	*Rotated*
RT	655	708	758	828	814	823
SE	28	29	21	25	36	34
% errors	5	3	4	3	5	3

and rotation was not reliable, $F(1, 29) < 1$. A Friedman ANOVA on all errors revealed no significant differences in the four priming conditions, $\chi^2(3) = 1.99$, $p > .57$. An additional ANOVA was run with target items as random variable to examine item effects. Again, there was a reliable main effect of attention, $F(1, 23) = 18.64$, $MSE = 16{,}964$, $p < .01$, and a main effect of rotation, $F(1, 23) = 5.12$, $MSE = 15{,}372$, $p < .05$, but no reliable interaction, $F(1, 23) < 1$.

Matched pairs t-tests revealed priming reliably greater than zero in the attended-same, $t(29) = 4.27$, $p < .001$; attended-rotated condition, $t(29) = 2.85$, $p < .01$; and ignored-same conditions, $t(29) = 2.12$, $p < .05$; but not in the ignored-rotated condition, $t(29) < 1$, $p > .05$ (see Figure 4). Thus, attended images in the prime display primed the probe image in both the whole and the rotated view but ignored images primed the probe object only when it was presented in the same view.

Discussion

The pattern of priming effects observed in Experiment 2 clearly replicated the findings of Stankiewicz et al. (1998) and Thoma et al. (2004). This outcome was not entirely predicted. As predicted, attended images primed both themselves and their plane-rotated image, whereas ignored images only primed themselves and not a plane-rotated version. Thus, the general notion of a hybrid model consisting of a holistic and analytic representation is supported by the fact that attended objects primed themselves in both the same view and the rotated view, whereas ignored objects only primed themselves in the same view (Hummel, 2001). However, unlike predicted from the model, the priming advantage for same views over rotated views was the same in both attended and ignored conditions. That is, as in the case for left–right reflections (Stankiewicz et al., 1998) and configural distortions (Thoma et al., 2004), the effects of attention and rotation were strictly additive. Moreover, as would be expected from "ideal" data, these previous studies and Experiment 2 obtained the same priming advantage of ~50 ms for ignored and attended conditions when holistic properties remained unchanged between prime and probe display. Thus, in terms of the hybrid model, this additive relationship suggests that the analytic representation is fully invariant with picture-plane rotations because the observed priming difference (~50 ms) between rotation conditions (attended or not) is attributed to the holistic representation. By contrast, the hybrid model, in its current state, predicts nonadditivity; greater priming costs for rotation in the attended condition than for ignored conditions because in addition to affecting the holistic component plane rotations should also change spatial relations in the analytic representation.

In considering priming for "attended" (cued) objects, we note that in the present experiments, recognition and naming were confounded in the attended conditions. Thus, not all the priming in the attended condition is visual but this cannot be an explanation for the effects in our data. The priming observed will contain a semantic or name prime component, as well as a component for visual priming but it is likely that the visual priming component is the larger (Bruce, Carson, Burton, & Ellis, 2000). Previous studies (Stankiewicz et al., 1998; Thoma et al., 2004) estimated visual priming by substituting the image in the identical conditions with a different object that had the same basic-level name (e.g., a grand piano instead of an upright piano). These "same-name-different-exemplars" produced no priming in the ignored condition, and significantly less priming than reflected (Stankiewicz et al., 1998) or split (Thoma et al., 2004) images in the attended-changed conditions. In both studies, subtraction produced a conservative estimate of about 80 ms of purely visual priming in the analytic representation. It is, therefore, reasonable to assume that the priming found in the attended conditions in Experiment 2 (using an almost identical paradigm and similar stimuli) also contained a significant and large visual component (see also Biederman & Cooper, 1991, 1992).

The present experiment was the first to test the model with stimuli that should show nonadditivity because of additional processing costs for attended rotated objects. Yet, the additive effects remain. It seems that models of object recognition like RBC (Biederman, 1987) and JIM (e.g., JIM.3; Hummel, 2001) underestimate the view invariance of the visual representation of shape. Nevertheless, while the additive effects of viewpoint and attention were not predicted by the hybrid model they would not have been predicted by any other current model either. We return later in the General Discussion to how the hybrid model might deal with the results from the attended condition. For now we consider how the results of the ignored conditions of Experiment 2 affect models of object recognition.

In contrast to the hybrid theory, geon theory (Biederman, 1987) would not predict view-dependent priming in the ignored conditions, as binding of parts should require attention (Hummel & Biederman, 1992). Similarly, multiple view accounts do not predict differences in priming between ignored conditions because in these models there is no explicit role for attention in shape representation. However, one multiple view account that does incorporate attention is the model of Olshausen et al. (1993) in which attention serves a gating function in early visual processing.

According to Olshausen et al. (1993), the outputs of retinotopic visual neurons (as found in V1 and V2) are mapped under attention to neurons whose receptive fields are invariant with translation and scale (and possibly other variations in viewpoint) in higher visual areas such as inferotemporal cortex (IT). Ignored information is either not mapped from V1 to IT or, if it is mapped, then it is sensitive to metric variations such as translation, scaling and rotation (see Olshausen et al., 1993). This model either predicts no priming at all for ignored objects (because such objects are not recognized) or, if priming in the ignored condition is assumed to reside in early visual representations (e.g., V1 or V2), then priming for ignored identical images only. Thus, a possible alternative explanation for the results of ignored stimuli in Experiment 2 is that the difference in priming effects between objects in the same view and a rotated view are due to the matching of simple features or global shape properties on a lower level of visual processing rather than the involvement of higher visual representations. The priming pattern in Experiment 2 could emerge simply because same views prime themselves always more than different views. Experiment 3 seeks to rule out such low level activation as the cause of view-dependent priming in the ignored conditions by testing whether images in unfamiliar (rotated) views prime themselves when ignored.

Experiment 3 served as an additional test of the hypothesis that activation in early visual representations is responsible for the observed view-dependent priming in the ignored conditions. It does so by investigating whether images in unfamiliar (rotated) views prime themselves when ignored: To the extent that priming in the ignored condition reflects activation in early visual representations, ignored base objects in unusual orientations should visually prime themselves since they activate the same early visual representations. But to the extent that the observed priming in the ignored condition reflects priming in the mapping from the representation of shape to object models in LTM, as we assume it does, then ignored images depicting base objects in off-upright orientations should not even prime themselves.

EXPERIMENT 3: PRIMING FOR IDENTICAL UPRIGHT OR ROTATED IMAGES AND ATTENTION

The goal of this experiment was to establish whether the pattern of priming observed in the ignored condition of Experiment 2 was due to activation of identical view-dependent object representations or to simple facilitation due to extraction of identical low level features. The pattern of priming in Experiment 2 was predicted by the hybrid account that the visual system generates holistic representations of ignored images and analytic representations of attended images (Hummel, 2001; Hummel & Stankiewicz, 1996). However, an alternative interpretation of these results is that the observed priming resides in early visual representations (i.e., rather than in the representations responsible for object recognition, as assumed by the hybrid model), and that identical images simply prime one another more than nonidentical images, and attended images prime one another more than unattended images. If this alternative explanation is correct, then the advantage for identical images over nonidentical images and the advantage for attended images over unattended images could produce the additive priming effects observed in Experiment 2. This interpretation is challenged by the results of Stankiewicz and Hummel (2002), who showed that priming for ignored images is invariant with translation and scale—and thus cannot be explained by an "early" locus of priming. However, the results from Experiment 1 allow us to test the different accounts directly with plane rotated objects.

On the low level matching account, it should follow that any view of an object primes its identical self but not a changed view (of the same object). Priming would be solely dependent on the view. A different prediction would follow from the hybrid model of object recognition. On that model, it is only *familiar* views of objects that cause priming in the ignored condition as long as there is no view change that alters the holistic properties of an object (Stankiewicz & Hummel, 2002). Priming in the ignored condition reflects the activation of holistic representations in LTM, but these exist only for familiar views (Hummel, 2001). Therefore, identical unfamiliar views would not show priming from ignored images. Plane-rotated views of base objects constitute such unfamiliar views.

Experiment 3 used only objects with a definite base because they presumably have only one familiar view (i.e., the upright) stored in LTM; no-base objects do not have unfamiliar views. In Experiment 3, for prime–probe pairs both the relevant prime objects (attended or ignored) and the corresponding probe images were shown in the same orientation—both appeared in either an upright (familiar) or rotated (unfamiliar) view. The particular interest is in the ignored trials. If ignored images make contact with stored representations in object memory, then objects that have a definite base and are seen almost exclusively in an upright position should exhibit no priming when both prime and probe are shown in a rotated view. At the same time, they should exhibit priming when both prime and probe are presented in the identical upright (familiar) view. If, however, the priming observed for ignored objects is due to simple low level priming, then even unfamiliar (rotated) views of base objects should prime themselves.

Method

Participants. Thirty native English speakers with normal or corrected-to-normal vision participated for credit in introductory psychology courses at Goldsmiths College University of London.

Materials. The materials consisted of objects taken from Snodgrass and Vanderwart (1980). There were 36 objects with a definite base that were used as probes, and 48 filler objects that were never used as probes (see Appendix).

Procedure. The basic procedure was similar to that of Experiment 2 with the following two differences. First, only base objects were used as probes; second, corresponding prime and probe images (ignored-repeated or attended-repeated) were always shown in the same orientation in a single trial—either rotated in the picture plane (90°) or upright. Attended or ignored prime objects that were not probed could appear in either an upright or rotated view.

Results

Trials on which either the prime or probe responses were incorrect (8.7%) were excluded from the analysis of latencies, as were voice key errors (3.3%) elicited by the subject. The mean response time for the standard view was 787.3 ms (*SE* 31.2) and 805.6 ms (*SE* 19.8) for the rotated view. This difference was not statistically reliable, $t(1, 29) < 1$. However, the mean error rate for the upright view was 4.4% (*SE* 1.5) and 13.8% (*SE* 2.5) for the rotated view, a significant difference, $t(29) = 3.32$, $p < .01$. The data were treated as in Experiment 2, however, different baselines were used to calculate priming RTs: Unprimed-rotated conditions formed the baseline for attended-rotated and ignored-rotated conditions, whereas unprimed-upright conditions were used to calculate priming for attended-upright and ignored-upright conditions. The priming RT were analysed in a 2 (attention: Attended vs. ignored) $\times$ 2 (rotation: Same view vs. rotated view) within-subjects ANOVA for latencies (see Figure 5 and Table 3). The analysis revealed a reliable main effect of attention, $F(1, 29) = 123.40$, $MSE = 7152.16$, $p < .001$, but no main effect of rotation, $F(1, 29) = 1.70$, $MSE = 21{,}607.37$, $p > .05$. The interaction between attention and rotation

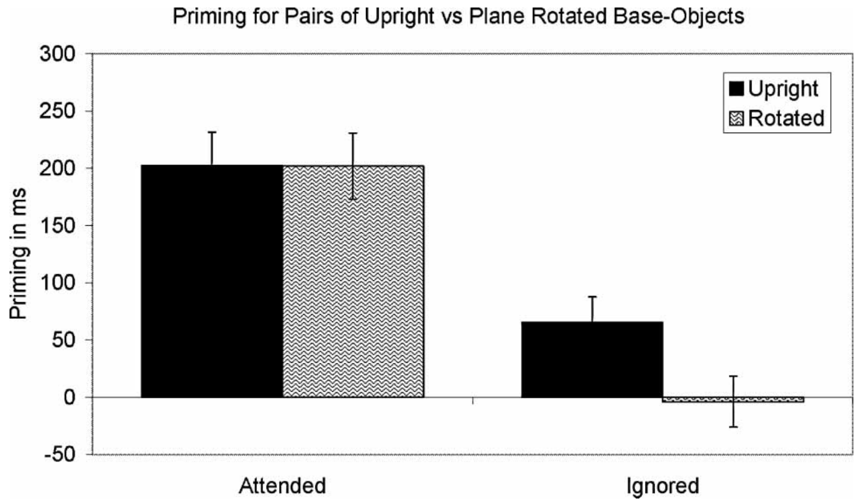

Figure 5. Priming means in response times and standard errors for base objects in Experiment 3 as a function of whether the object was attended or ignored in the prime display prior to the probe and whether the prime and the probe objects were presented in the upright orientation or rotated in the picture plane ($n = 30$).

was reliable, $F(1, 29) = 7.64$, $MSE = 4637.15$, $p < .05$. The difference between the attended-same and attended-rotated conditions was not reliable, $t(29) < 1$, whereas the difference between the ignored-same and ignored-rotated conditions was reliable, $t(29) = 2.27$, $p < .05$. Thus, there was no difference in priming between attended conditions, but a significant difference in priming latencies between ignored-familiar and ignored-unfamiliar conditions. An additional ANOVA was run with target items as random variable to examine item effects. Again, there was a reliable main effect of attention, $F(1, 29) = 100.75$, $MSE = 8641$, $p < .001$, no main effect of rotation, $F(1, 29) = 1.37$, $MSE = 34{,}227$, $p > .05$, and a reliable interaction effect, $F(1, 29) = 11.59$, $MSE = 3958$, $p < .01$.

TABLE 3
Mean response times (RT, in milliseconds), standard errors, and percentage errors forprobe objects Experiment 3

	Attended		*Ignored*		*Unprimed*	
Variable	*Same*	*Rotated*	*Same*	*Rotated*	*Same*	*Rotated*
RT	584	603	722	809	787	805
SE	13	17	19	21	31	20
% errors	2	3	3	3	2	7

Errors were unevenly distributed and occurred mainly in the attended rotation condition. A Friedman analysis of ranks gave $\chi^2(3) = 20.45$, $p < .001$. Wilcoxon tests revealed more errors in the attended rotated condition than all other conditions (all ps $< .01$) but none of the other comparisons showed significant differences (all ps $> .4$).

Matched pairs t-tests were conducted on each priming condition to determine savings in response time when naming the probe (i.e., relative to unprimed probes). Priming was reliably greater than zero in the attended-upright, $t(29) = 6.92$, $p < .001$; attended-rotated, $t(29) = 9.04$, $p < .001$; and ignored-upright conditions, $t(29) = 2.26$, $p < .05$; but not in the ignored-rotated condition, $t(29) < 1$. Attended images in the prime display primed themselves regardless of whether they were upright or rotated, but ignored images primed themselves only when they were upright (see Figure 5).

In Experiment 3, different baselines were required (unprimed-upright vs. unprimed-rotated) to calculate priming because better performance was expected for upright images than for rotated images. In the previous experiment, only a single baseline was required because of equal performance across views of no-base objects. To more clearly compare the priming pattern with Experiment 2, an additional ANOVA was run over priming data obtained by pooling baselines (i.e., priming was established from the mean of unprimed-upright and unprimed-rotated condition). There were reliable main effects of attention, $F(1, 29) = 123.39$, $MSE = 7152.16$, $p < .001$, and rotation, $F(1, 29) = 28.45$, $MSE = 2992.58$, $p < .001$, and the interaction between attention and rotation was reliable, $F(1, 29) = 7.64$, $MSE = 4637.15$, $p < .01$. There was no difference in mean priming RT between attended images in the same and the rotated condition, $t(29) < 1$, but the difference between ignored images was reliable, $t(29) = 4.45$, $p < .001$. Further matched-pairs t-tests showed that priming was reliably greater than zero in the attended-same, $t(29) = 9.91$, $p < .001$, attended-rotated, $t(29) = 8.03$, $p < .001$, and ignored-same conditions, $t(29) = 3.59$, $p < .05$, but not in the ignored-rotated condition, $t(29) < 1$. Thus, the important effects found in Experiment 3 were not due to the way the baseline was established. Pooling the baseline to calculate priming yielded the same priming pattern as the previous analysis with separate baselines. The only difference was that absolute priming differences between upright and rotated conditions were accentuated, hence the now significant main effect of rotation.

Discussion

The critical result of Experiment 3 is the replication of the pattern of performance found with ignored images in Experiment 2. Once more, we find a significant amount of priming in one ignored condition (familiar views) and no priming in the other (unfamiliar views). The lack of priming in Experiment 2 was for familiar images that were rotated between prime and probe displays. Importantly, the lack of priming in Experiment 3 was for unfamiliar views that were *identical* in prime and probe trials. Thus, the priming differences for ignored objects in Experiment 2—in which the orientation between prime and probe object was either the same or rotated—cannot be attributed to a simple low level priming advantage for the unchanged condition nor can they be trivially attributed to the amount of similarity between prime and target views. The priming pattern for ignored objects observed here is perhaps the most direct evidence that ignored images prime subsequent recognition by making contact with object models in LTM. Additional evidence that only familiar (stored) holistic representations get primed in the ignored route was obtained by Thoma et al. (2004). They showed that ignored images primed themselves when they were intact, whereas ignored images that were split (with two halves moved to the contralateral side) did not prime the very same split image in the probe trial. Furthermore, in fMRI studies, Henson, Shallice, and Dolan (2000) found that repetition priming resulted in different patterns of attenuation in the right fusiform area depending on whether the repeated stimuli were familiar or unfamiliar. Together with the data from Vuilleumier et al. (2002) these imaging results would imply that the right fusiform area exhibits properties directly predicted from the holistic component of the model.

For the attended conditions in Experiment 3, we note differences to the priming effects observed by Stankiewicz et al. (1998) for same view versus mirror reflection. They reported an additive effect of attention and viewpoint in the attended condition. Here there was equivalent priming (i.e., no additivity) for attended objects in an upright view and rotated view. There are three possible explanations for these data from the attended condition, all of which derive from the fact the rotated objects are harder to recognize than upright objects. First, the parts of the rotated prime would activate the analytic components of the object representation. If the rotated images take longer to be recognized, this could allow extra analytic priming. A related second explanation is that performance for rotated images is farther from ceiling so there is more room for learning during a rotated prime trial and hence more room for improvement on a rotated probe trial. A third possibility is that by attending to and recognizing the rotated image, a holistic representation is thereby encoded. Hence, on subsequent presentation of the target, both upright and rotated images benefit equally from analytic and holistic representations. The third interpretation fits with experimental data reported elsewhere showing that previously attended rotated objects are recognized more quickly during subsequent presentation whereas formerly ignored objects do not show such an advantage of prior exposure (Murray, 1995).

Whatever the best explanation of the equal priming in the attended conditions, it in no way detracts from the more important findings for the ignored conditions. In the ignored conditions, we found no priming for unfamiliar views of objects. The lack of priming is more notable because it might be thought likely that increased baselines for rotated (unfamiliar) images ought to increase the amount of priming for rotated objects over upright images. Yet, it was only for the ignored upright images with the lower baseline for recognition, that we observed priming.

GENERAL DISCUSSION

The present experiments showed that object recognition depends on attention and view familiarity—properties which are proposed in the hybrid model by Hummel and colleagues (Hummel, 2001; Hummel & Stankiewicz, 1996; Stankiewicz et al., 1998; Thoma et al., 2004). At the same time, some aspects of the current instantiation of the theory were falsified. In Experiment 2, ignored no-base objects only primed themselves when the prime and probe images were presented in the same orientation. This result suggests, as predicted by the hybrid model, that no-base objects are represented in LTM by multiple object models, each responsible for recognizing the object at a particular orientation in the picture plane, but all making contact with the same object concept. By contrast, in Experiment 3, ignored base objects in unfamiliar views did not prime themselves at all, even in the identical view. This result suggests that the object model for a base object is specific to a particular (upright) picture-plane orientation, and that priming in the ignored condition reflects a holistic representation of the object's shape making contact with that object model. More generally, the lack of transfer of priming between different orientations in the ignored conditions of Experiments 2 and the absence of priming between identical noncanonical views in Experiment 3 are consistent with the hypothesis that ignored images are recognized on the basis of view-specific holistic representations of shape.

By contrast, the attended condition of Experiment 2 revealed robust transfer of priming across changes in picture-plane orientation—indeed, it revealed greater transfer than predicted by any current models of object recognition, including the hybrid model that motivated the experiment. Although the invariance with picture-plane rotation observed in the attended condition of Experiment 2 is greater than predicted by the hybrid model in its current form, the pattern of attended priming effects is still best explained by a hybrid representation. First, many previous studies have shown the involvement of part-based (structural) representations when attending to objects (e.g., Biederman & Bar, 1999; Biederman & Cooper, 1991; Foster & Gilson, 2002; Thoma et al., 2004). Second, the stark dissociation between the patterns of priming observed in the attended and ignored conditions is highly consistent with the hybrid model. Like the dissociation observed by Stankiewicz et al. (1998) using attention and mirror reflection, and the dissociation observed by Thoma et al. (2004) using attention and configural distortions, the dissociation observed here—that priming for attended images is more view invariant than priming for ignored images—demands explanation in terms of a hybrid representation of shape, in which one part of the representation (the "analytic" component) is robust to various image distortions but requires visual attention to generate, and another part (the "holistic" component) that can be generated automatically but is intolerant of image distortions (except for translation and scale changes (Stankiewicz & Hummel, 2002).

In no previous test of the hybrid model did attention interact with the other factor being manipulated; there are additive effects on visual priming for reflection or configural distortion (Stankiewicz et al., 1998; Thoma et al., 2004) or no effects for changes in translation and scale (Stankiewicz & Hummel, 2002). The present Experiment 2 was the first to test the model with stimuli that it predicts should produce an interaction with attention: The model predicts that attended images should prime more than ignored images, that identical picture-plane orientations should prime more than nonidentical orientations, and that the effects of identical vs. nonidentical orientation should be greater for attended images than ignored ones. Yet, the additive effects remain, suggesting that, counter to the predictions of the hybrid model, priming for the analytic representation may be invariant with picture-plane rotation.

An explanation for the invariance of priming with rotation may come from Stankiewicz (2002). Using a noise masking paradigm with simple one-part shapes, Stankiewicz (2002, Exp. 3) showed that the representation of object shape is independent of (i.e., invariant with) viewpoint. Stankiewicz noted that, although his result demonstrating independence of shape and viewpoint appears at first blush to be at odds with the (apparent) view sensitivity of recognition (e.g., as measured by naming RT and errors), it really is not: Even if viewpoint is represented independently of shape (i.e., that the representation of shape is itself view *invariant*), for many objects, viewpoint information is nonetheless diagnostic, in the sense that many objects appear in some views (e.g., upright) more than others; hence, a rational visual system would use this view information (at least when it is diagnostic). As a result, when it is diagnostic, if it is also unusual, viewpoint information is expected to impede object recognition, even though the representation of shape is, itself, view invariant. Consistent with this conjecture, objects that do not appear in prototypical views (such as knives, forks, hammers, etc., our "no-based" objects) do not show the typical "rotation function" in object identification (Experiment 1) shown by objects that do have a canonical upright view (such as cars, houses and animals).

It seems that the object recognition system may use viewpoint when it is diagnostic (e.g., for cars) but ignore it when its not (e.g., for forks). However, this would seem to suggest a paradox: How does the visual system "know", *before* recognizing an object, whether the object is one for which viewpoint should be used or ignored? The circularity can be avoided simply by assuming, as proposed by Stankiewicz (2002), that the visual system represents view-specific information (e.g., viewer-centred spatial relations) independently of view-invariant information (e.g., various aspects of geons and perhaps various object-centred relations) and that it simply learns stronger connections from view-sensitive aspects of the representation to object models of objects that frequently appear in a canonical view (such as cars) than for object models of objects that do not appear in a canonical view (such as forks). A system that represents information independently is free to attend to (i.e., use; have strong connections from) or ignore (i.e., do not use; have weak connections from) various aspects of that information as demanded by the statistics of experience. That is, perhaps the visual system learns weaker view-sensitive connections for no-base objects precisely because the neurons representing that information are not consistently activated by no-base objects: Sometimes the tines of a fork are "above" the handle, sometimes "beside" and sometimes "below", so the visual system does not learn to prefer one over another. Thus, when you see a fork in some particular orientation (e.g., with the tines above the handle), that relation is activated as part of the representation of the (as yet to be recognized) object's shape, but because of its weak connections to the object model it neither helps nor hurts object identification.

Viewpoint should matter to the extent that it is regular for any given object. Each object may have its own range(s) of views over which recognition is unaffected (i.e., because all views within the range(s) are about equally likely) but have other range(s) over which it is sensitive to viewpoint). For presumably the same reason, colour has little effect on object recognition for most objects (Biederman & Ju, 1988; Ostergaard & Davidoff, 1985). Colour gets activated, but since it is not part of the object model (or is only a weak part), it has little or no effect on recognition. By contrast, for base objects, a single set of view-specific relations *will* tend to be systematically activated by the object and so *will* have strong connections to the object model. Hence, for these objects, unfamiliar view information impairs recognition. The modification to Hummel's (2001; Hummel & Stankiewicz, 1996) hybrid model necessary to account for the findings of Experiment 2 may be, to a first approximation, simple: Add to its vocabulary of relations a population of more view-invariant relations (such as connectedness and other object-centred relations). The more challenging part will be adding the routines, in the lower layers of the model, that compute these relations and to explain why these effects show up in first-order measures of object identification (e.g., naming RTs and errors; exposure duration necessary for recognition, etc.) but not in second-order measures (i.e., visual priming).

In conclusion, the present study is the first that tested the hybrid model's predictions on priming for plane-rotated objects. Ignored objects prime themselves, but only when presented in a canonical view and in a strictly view-dependent manner, as predicted by the model. Attended objects, however, prime themselves in rotated views, but in a relatively view-independent manner. The current results are in agreement with the general notion of a hybrid representation of object shape, consisting of an analytic (structural) and holistic (view-like) component. They also indicate, however, that view dependency in the analytical route is less pronounced (at least for some types of objects such as no-base objects) than previously assumed.

REFERENCES

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115–147.

Biederman, I., & Bar, M. (1999). One-shot viewpoint invariance in matching novel objects. *Vision Research*, *39*, 2885–2899.

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., & Cooper, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 121–133.

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 and Performance*, *21*(6), 1506–1514.

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, *20*, 38–64.

Bruce, V., Carson, D., Burton, M. A., & Ellis, A. W. (2000). Perceptual priming is not a necessary consequence of semantic classification of pictures. *Quarterly Journal of Experimental Psychology*, *53A*(2), 289–323.

Bülthoff, H. H., & Edelman, S. (1992). Psychophysical support for a 2-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences of the USA*, *89*(1), 60–64.

Clowes, M. B. (1967). Perception, picture processing and computers. In N. L. Collins & D. Michie (Eds.), *Machine intelligence* (Vol. 1, pp. 181–197). Edinburgh, UK: Oliver & Boyd.

Cooper, E. E., Biederman, I., & Hummel, J. E. (1992). Metric invariance in object recognition: A review and further evidence. *Canadian Journal of Psychology–Revue Canadienne De Psychologie*, *46*(2), 191–214.

Davidoff, J., & Warrington, E. K. (1999). The bare bones of object recognition: Implications from a case of object recognition impairment. *Neuropsychologia*, *37*(3), 279–292.

Davidoff, J., & Warrington, E. K. (2001). A particular difficulty in discriminating between mirror images. *Neuropsychologia*, *39*(10), 1022–1036.

Edelman, S., & Intrator, N. (2003). Towards structural systematicity in distributed, statically bound visual representations. *Cognitive Science*, *27*, 73–110.
Foster, D. H., & Gilson, S. J. (2002). Recognizing novel three-dimensional objects by summing signals from parts and views. *Proceedings of the Royal Society of London Series B: Biological Sciences*, *269*, 1939–1947.
Hayward, W. G. (2003). After the viewpoint debate: Where next in object recognition. *Trends in Cognitive Sciences*, *7*, 425–427.
Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science*, *287*, 1269–1272.
Hummel, J. E. (1994). Reference frames and relations in computational models of object recognition. *Current Directions in Psychological Science*, *3*, 111–116.
Hummel, J. E. (2000). Where view-based theories break down: The role of structure in human shape perception. In E. Dietrich & A. B. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines* (pp. 157–185). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Hummel, J. E. (2001). Complementary solutions to the binding problem in vision: Implications for shape perception and object recognition. *Visual Cognition*, *8*, 489–517.
Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape-recognition. *Psychological Review*, *99*(3), 480–517.
Hummel, J. E., & Stankiewicz, B. J. (1996). An architecture for rapid, hierarchical structural description. In T. Inui & J. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 93–121). Cambridge, MA: MIT Press.
Intraub, H. (1981). Identification and processing of briefly glimpsed visual scenes. In D. Fisher (Ed.), *Eye movements: Cognition and visual perception* (pp. 181–190). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory and Cognition*, *13*(4), 289–303.
Lawson, R. (1999). Achieving visual object constancy across plane rotation and depth rotation. *Ada Psychologica*, *102*(2–3), 221–245.
Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1015–1036.
Marr, D., & Nishihara, H. K. (1978). *Representation and recognition of three dimensional shapes*. Paper presented at the Proceedings of the Royal Society of London.
Murray, J. E. (1995). The role of attention in the shift from orientation-dependent to orientation-invariant identification of disoriented objects. *Memory and Cognition*, *23*(1), 49–58.
Murray, J. E. (1997). Flipping and spinning: Spatial transformation procedures in the identification of rotated natural objects. *Memory and Cognition*, *25*(1), 96–105.
Olshausen, B., Anderson, C, & van Essen, D. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience*, *13*(11), 4700–4719.
Oram, M. W., & Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *Journal of Neurophysiology*, *68*(1), 70–84.
Ostergaard, A. L., & Davidoff, J. B. (1985). Some effects of color on naming and recognition of objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 579–587.
Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, *9*, 441–474.
Poggio, T., & Edelman, S. (1990). A network that learns to recognize 3-dimensional objects. *Nature*, *343*(6255), 263–266.

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

Saiki, J., & Hummel, J. E. (1998). Connectedness and the integration of parts with relations in shape perception. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 227–251.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174–215.

Stankiewicz, B. J. (2002). Empirical evidence for independent dimensions in the visual representation of three-dimensional shape. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(4), 913–932.

Stankiewicz, B. J., & Hummel, J. E. (2002). Automatic priming for translation- and scale-invariant representations of object shape. *Visual Cognition*, *9*(6), 719–739.

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*(3), 732–744.

Sutherland, N. S. (1968). Outlines of a theory of visual pattern recognition in animals and man. *Proceedings of the Royal Society of London, Series B*, *171*, 95–103.

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 and Performance*, *21*(6), 1494–1505.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape-recognition. *Cognitive Psychology*, *21*(2), 233–282.

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

Thoma, V., Hummel, J. E., & Davidoff, J. (2004). Evidence for holistic representation of ignored images and analytic representation of attended images. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(2), 257–267.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.

Tipper, S. (1985). The negative priming effect: Inhibitory effects of ignored primes. *Quarterly Journal of Experimental Psychology*, *37A*, 571–590.

Treisman, A., & DeSchepper, B. (1996). Object tokens, attention, and visual memory. In T. Inui & J. McClelland (Eds.), *Attention and performance: XVI. Information integration in perception and communication* (pp. 15–46). Cambridge, MA: MIT Press.

Vannucci, M., & Viggiano, M. P. (2000). Category effects on the processing of plane-rotated objects. *Perception*, *29*(3), 287–302.

Verfaillie, K., & Boutsen, L. (1995). A corpus of 714 full-color images of depth-rotated objects. *Perception and Psychophysics*, *57*(1), 925–961.

Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J. (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience*, *5*(5), 491–499.

Warren, C. E. J., & Morton, J. (1982). The effects of priming on picture recognition. *British Journal of Psychology*, *73*, 117–130.

Manuscript received July 2004
Manuscript accepted March 2005
First published online March 2006

APPENDIX
Stimuli

Experiment 1

No-base objects: banana, brush, carrot, comb, football, fork, glove, guitar, gun, hammer, key, knife, leaf, lock, pen, ring, saw, scissors, screwdriver, spoon, tennis racket, trumpet, watch, whistle
Base objects: baby carriage, bed, bike, boot, cake, candle, chair, couch, cup, desk, harp, helicopter, house, iron, ironing board, kettle, lamp, pitcher, sailboat, table, telephone, ashtray, truck, watering can
Animals: ant, bird, elephant, camel, cat, chicken, cow, dog, donkey, duck, frog, gorilla, grasshopper, horse, monkey, mouse, owl, peacock, pig, rabbit, snail, squirrel, tiger, turtle

Experiment 2

Targets: saw, carrot, leaf, lock, guitar, broom, umbrella, trumpet, book, plug, glove, fork, scissors, french horn, pineapple, hammer, key, violin, butterfly, brush, light bulb, watch, lobster, corn
Fillers: banana, rolling pin, screwdriver, nail file, envelope, cigarette, anchor, mitten, comb, chisel, knife, pliers, toothbrush, watermelon, spoon, ball, whistle, belt, baseball bat, potato, barrel, ruler, tennis racket, spool, wrench, star, football, sun, pen, tomato, pepper, nut

Experiment 3

Targets: alligator, bird, couch, giraffe, desk, house, kettle, plane, piano, snail, squirrel, truck, boot, bike, chicken, chair, helicopter, frog, candle, pants, jacket, duck, wineglass, telephone, bed, baby carriage, cat, car, elephant, dog, mouse, motorcycle, sailboat, rabbit, table, suitcase
Fillers: apple, accordion, axe, anchor, banana, ant, barrel, ashtray, basket, bear, baseball bat, beetle, broom, belt, brush, cake, bottle, cherry, cup, doll, fish, fence, flag, flower, glasses, frying pan, guitar, grapes, gun, hanger, hammer, iron, key, kite, knife, pineapple, pipe, pliers, scissors, sheep, screwdriver, sled, shoe, swing, snake, toaster, trumpet, vase