

# Priming of Depth-Rotated Objects Depends on Attention and Part Changes

Volker Thoma<sup>1</sup> and Jules Davidoff<sup>2</sup>

<sup>1</sup>University of East London, UK, <sup>2</sup>Goldsmiths University of London, UK

**Abstract.** Three priming experiments investigated the role of attention and view changes when common objects were rotated in depth. Objects were shown in prime-probe trial pairs. Experiment 1 extended findings by Stankiewicz, Hummel, and Cooper (1998) showing that attended objects primed themselves in the same but not in a reflected view, whereas ignored objects only primed themselves in the same view. In Experiment 2, depth-rotations produced changes in the visible part structure between prime and probe view of an object. Priming after depth-rotation was more reduced for attended objects than for ignored objects. Experiment 3 showed that other depth rotations that did not change the perceived part structure revealed a priming pattern similar to that in Experiment 1, with equivalent reduction in priming for attended and ignored objects. These data indicate that recognition of attended objects is mediated by a part-based (analytic) representation together with a view-based (holistic) representation, whereas ignored images are recognized in a strictly view-dependent fashion.

**Keywords:** attention, object recognition, priming, structural description, representation, view-dependency

## Introduction

A single 3D object can be encountered from a number of viewpoints each producing a potentially unique 2D projection. How important are these views for object recognition? Many researchers have shown that for common objects recognition performance reliably drops with rotations from a familiar or trained viewpoint and argued that recognition was view-dependent both in the picture plane (e.g., Jolicoeur, 1985; McMullen & Jolicoeur, 1990, 1992; Murray, 1998, 1999; Murray, Jolicoeur, McMullen, & Ingleton, 1993) and in depth (e.g., Hayward, 1998; Lawson & Humphreys, 1996, 1998; Lawson, Humphreys, & Jolicoeur, 2000). View-point dependent effects have also been obtained with novel objects rotated in the picture plane (e.g., Tarr & Pinker, 1989, 1990) and in depth (e.g., Bulthoff & Edelman, 1992; Hayward & Tarr, 1997; Tarr, 1995; Willems & Wagemans, 2001). In consequence view-based (or image-based) theories (e.g., Bulthoff & Edelman, 1992; Edelman & Intrator, 2003; Poggio & Edelman, 1990; Tarr & Bulthoff, 1995; Tarr & Pinker, 1989) propose that we recognize objects by matching their images to specific holistic *views* in long-term memory). However, not all object

recognition is viewpoint-dependent. Biederman and his colleagues have obtained view-invariant effects after image transformations such as translation (Biederman & Cooper, 1991a), scaling (Biederman & Cooper, 1992), left-right (i.e., mirror) reflection (Biederman & Cooper, 1991b), and some rotations in depth (Biederman & Gerhardstein, 1993).

Viewpoint invariant recognition is usually accounted for by theories that propose part-based representations (e.g., Biederman, 1987; Hummel & Biederman, 1992; Marr & Nishihara, 1978; Palmer, 1977). According to this approach the visual system represents objects as structural descriptions, which specify an object's parts, such as generalized cylinders (Marr & Nishihara, 1978) or geons (Biederman, 1987), in terms of their spatial relations to one another. For example, a structural description would represent the shape of a coffee mug as a "curved cylinder" (i.e., the handle) "side-attached" to a "vertical straight cylinder" (i.e., the body, see Biederman, 1987). This description does not specify – and therefore is not affected by – the distance and angle from which the mug is viewed (an exception are "accidental" views, which occlude parts and project volumetric parts as 2D forms). Consistent with the structural de-

scription account of shape perception, there is evidence that the visual system represents the relations among an object's parts both explicitly (Hummel & Stankiewicz, 1996; Palmer, 1978; Tversky & Hemenway, 1984) and independently of the parts themselves (Saiki & Hummel, 1998).

There are properties of object perception, however, that are inconsistent with the concept of structural representations. First, to construct structural descriptions, object recognition would require attention and time to bind parts and spatial relations (e.g., Hummel & Biederman, 1992). However, there is behavioral evidence (e.g., Intraub, 1981) as well as evidence from single-unit recording (e.g., Oram & Perrett, 1992) that indicate that object recognition can operate faster than structural representations would allow (see Hummel, 2001). Second, object recognition can occur without attention. For example, studies have demonstrated that ignored images prime a subsequent corresponding probe in both negative (e.g., Tipper, 1985; Treisman & DeSchepper, 1996) and positive priming paradigms (Thoma, Hummel, & Davidoff, 2004; Stankiewicz et al., 1998; Stankiewicz & Hummel, 2002).

Thus, there is evidence for properties indicating both view-based and structural representations but neither of them seems sufficient to explain human object recognition. Hummel and Stankiewicz (1996; Hummel, 2001) proposed a model in which 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). Unlike a structural description, the holistic representation does not specify an object's parts or their spatial relations independent of each other. Instead, object's parts are represented in terms of their topological positions in a 2-D coordinate system. This representation is sensitive to many variations in viewpoint (such as rotations in the picture plane and substantial rotations in depth) but it is invariant with some changes (such as location in the visual field and with scale). The holistic representation therefore permits rapid, automatic recognition of familiar objects in familiar views (because it does not require visual attention to bind parts and categorical spatial relations), but allows little generalization to novel views or to novel exemplars of known categories. The analytic representation, in contrast, codes an object's shape explicitly in terms of the categorical interrelations among its parts. This representation is largely robust to many variations in viewpoint (such as translation, changes in scale, left-right reflection and some rotations in depth) but it, too, is sensitive to rotations in the picture plane (if categorical spatial relations are changed, 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 independent of each other.

According to the hybrid holistic/analytic model, attending to an object's image activates (and therefore visually primes) both a structural description of the object's shape and a holistic (i.e., view-like) representation of its shape. Ignoring an image activates the holistic representation of its shape but not the structural description. The hybrid model therefore allows novel predictions for the relationship between visual attention and visual priming as a function of variations in viewpoint and other manipulations of an object's image. In general, 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.

The predictions have been tested and confirmed by Stankiewicz et al. (1998) and Thoma et al. (2004). As predicted by the model attended images visually primed themselves, left-right reflected and configurally distorted (split into two displaced halves) versions of themselves, whereas ignored images visually primed only themselves but not their mirror reflections or distorted images. Moreover, the priming advantage for same view prime-probe pairs over left-right reflected or configurally distorted image pairs was equivalent in both attended and unattended conditions (about 50 ms) indicating two independent components. Taken together with previous studies these findings show that the hybrid model can account for a large number of properties in object recognition while providing novel predictions concerning the effects of viewpoint and attention. However, one area in which the model has not yet been tested is rotation in depth (i.e., about the y-axis). These rotations usually affect recognition performance and priming across views (for a review, see Lawson, 1999) but the underlying representations and processes are not clear. Biederman and Cooper (1991b) originally demonstrated the involvement of geons in object recognition by showing that deletion of geon-critical contours (e.g., vertices) have a more detrimental effect on recognition than equal amounts of other contour-deletions. In support of the structural description account, Biederman and Gerhardstein (1993) showed that substantial rotations around the z-axis did not significantly diminish priming for geon-based objects, except for rotations which revealed new geons or occluded previously seen parts. However, Tarr, Williams, Hayward, and Gauthier (1998) have demonstrated that even single geons

(which are by definition represented in a view-independent fashion) incur performance costs after depth rotations in both naming and matching tasks (but see Biederman, 2000). Moreover, Lawson and Humphreys (1998) reported view-dependent priming effects even after depth rotations of only  $10^\circ$  (with a delay of several minutes between prime and probe display). Previously, Lawson and Humphreys (1996) have also shown that view-specificity was reduced with long (2,510 ms, vs. short: 585 ms) interstimulus intervals between prime and probe.

The hybrid theory of object recognition may offer an explanation for the qualitative and metric effects of depth-rotation without making any further assumptions. As with mirror-reflections (Stankiewicz et al., 1998), substantial depth-rotations should produce a view that is significantly different from the holistic representation of the original view. This should lead to a reduction in priming compared to seeing an object in exactly the same view. Therefore, unlike geon theory, the hybrid model would predict that any substantial change in depth orientation from prime to probe trial in a short-term priming (or sequential matching) paradigm should result in recognition costs; this is because the holistic representation will always be affected regardless of whether or not part changes occur. Thus, the model can potentially account for the results of studies that found recognition costs for matching depth-rotated objects (Hayward & Tarr, 1997; Lawson & Humphreys, 1996; Tarr et al., 1998). At the same time, the hybrid model would also predict additional reductions in priming for attended objects after depth-rotations if parts are revealed or occluded between depth-rotated views. According to the model, units representing parts (and their spatial relations) fire maximally on repeated presentation of an object when activated by the same visible parts but less so if parts are missing or new parts are visible. If the number (or type) of visible parts of an object is changed after depth-rotations, then there should be a larger reduction in priming between the attended conditions (reduction in analytic and holistic activation after depth-rotation) relative to the ignored conditions (reduction only in holistic activation after depth-rotation). In other words, for depth-rotations involving part changes the hybrid model would predict an interaction between attention and view. This study aims to explore whether the predictions derived from the hybrid model can account for priming patterns of attended and ignored objects that are presented in the identical view or rotated in depth between prime and probe view.

## Experiment 1: Priming for Mirror Images

The primary goal of this experiment was to replicate the findings of Stankiewicz et al. (1998) in which mirror images were found to prime their original version in the attended conditions, but not in ignored conditions. Mirror images can be considered as depth-rotations if the object has an axis of symmetry that can be aligned with the line of sight, which was the case for most of the objects in this experiment (48 out of 56). The basic procedure followed the paradigm of Stankiewicz et al. (1998). Briefly presented prime displays contained images of two objects, one to the left and one to the right of fixation. One object was pre-cued with a square box in its location (attended prime), and the subject's task was to name the object that appeared within the cueing square; the other (uncued) prime object was to be ignored. Immediately following the prime objects a probe display presented a single image at fixation, which again had to be named.

A novel set of objects was used consisting of grey-level images which were produced by rendering computer generated 3D objects (see Figure 1). The resulting images were near photorealistic and allowed to investigate whether the findings of priming found for black-and-white line drawings as employed in previous experiments (Stankiewicz et al., 1998; Thoma et al., 2004) generalize to ecologically more plausible images. According to the predictions of the hybrid model (Hummel & Stankiewicz, 1996; Hummel, 2001) and the findings of Stankiewicz et al. (1998) attended images prime themselves and mirror-reflections, but ignored images only prime themselves, not their reflected versions. The effects of view and attention are predicted to be additive: The reduced priming component resulting from the view change should be equivalent for both attended and ignored images.

## Method

### Participants

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

### Material

Fifty-six common everyday objects were used (see Appendix). The objects were obtained from various open

sources on the internet in 3D Max (Autodesk) format. Each object was oriented in a standard  $0^\circ$  orientation, in which the main axis of elongation and/or the symmetry axis coincided with the line of sight. An object was then rotated slightly between  $5^\circ$  and  $10^\circ$  in azimuth to give it a more canonical view (Banz, Tarr, & Bulthoff, 1999) as if the observer's vantage point was slightly elevated. Each object was then rotated  $30^\circ$ ,  $60^\circ$ , and  $90^\circ$  in depth from the standard view. All objects were rendered in 3D Max Studio (R3) using a  $25^\circ$  field of view, which gave an impression of perspective without drastically changing the perceived relative size of objects' parts. The objects were surface rendered with overhead lighting but without cast shadows. The size of the images was then standardized. In Experiment 1, objects were only shown in the  $60^\circ$  viewpoint and their mirror reflection (see Figure 1) which was equivalent to a  $120^\circ$  rotation in depth.

The objects were counterbalanced across participants by placing each object in one of fourteen clusters of four objects. Each object from one cluster served as prime or probe object (or both in non-baseline conditions) in 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-pair for a given participant, and all objects appeared equally often in each condition as a probe, an ignored prime or an attended prime.

### Procedure

The ordering of the trials and the pairing of attended and ignored objects on prime trials were randomized for each participant. After reading and paraphrasing the instructions, the participant read the names of the objects on the screen and then received 12 practice trials. Each experimental trial began with an unfilled circle (subtending  $0.032^\circ$  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 cuing square ( $4.57^\circ \times 4.57^\circ$ ) either to the left or right of the fixation cross at a distance of  $4.0^\circ$ . 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 2). Both images were centered  $4.0^\circ$  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 ( $15.6^\circ$  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 1,995 ms followed by a fixation cross ( $0.032^\circ$ ) 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, 3,015 ms elapsed between the end of the prime display and the beginning of the probe display (495 ms for the prime mask, 1,995 ms for the blank screen, 495 ms for the probe fixation dot and 30 ms blank). Following the probe display, a single pattern mask ( $4.57^\circ$ ) was shown in the centre of the screen for 495 ms. The probe object was either identical to the attended object (attended conditions), the ignored object (ignored conditions), or it was a third object not seen previously in the experiment (unprimed baseline condition). Half of the probe images were shown in the view identical to the corresponding prime object and half in the depth-rotated view. The participant's task was to name the probe as quickly and as accurately as possible, after which a display with the names of the attended prime and the probe along with the probe response time appeared. 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.

A departure from Stanekiewicz et al. (1998) was that immediately after the last trial participants were asked if they had been able to identify the to-be-ignored object and if so whether they could name it. This was done to establish whether the participants were following the instructions or if they were paying attention to the ignored (uncued) objects.

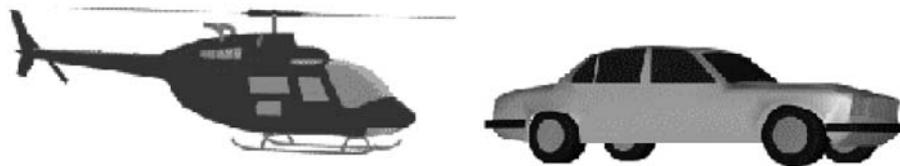


Figure 1. Examples of objects shown in Experiment 1.

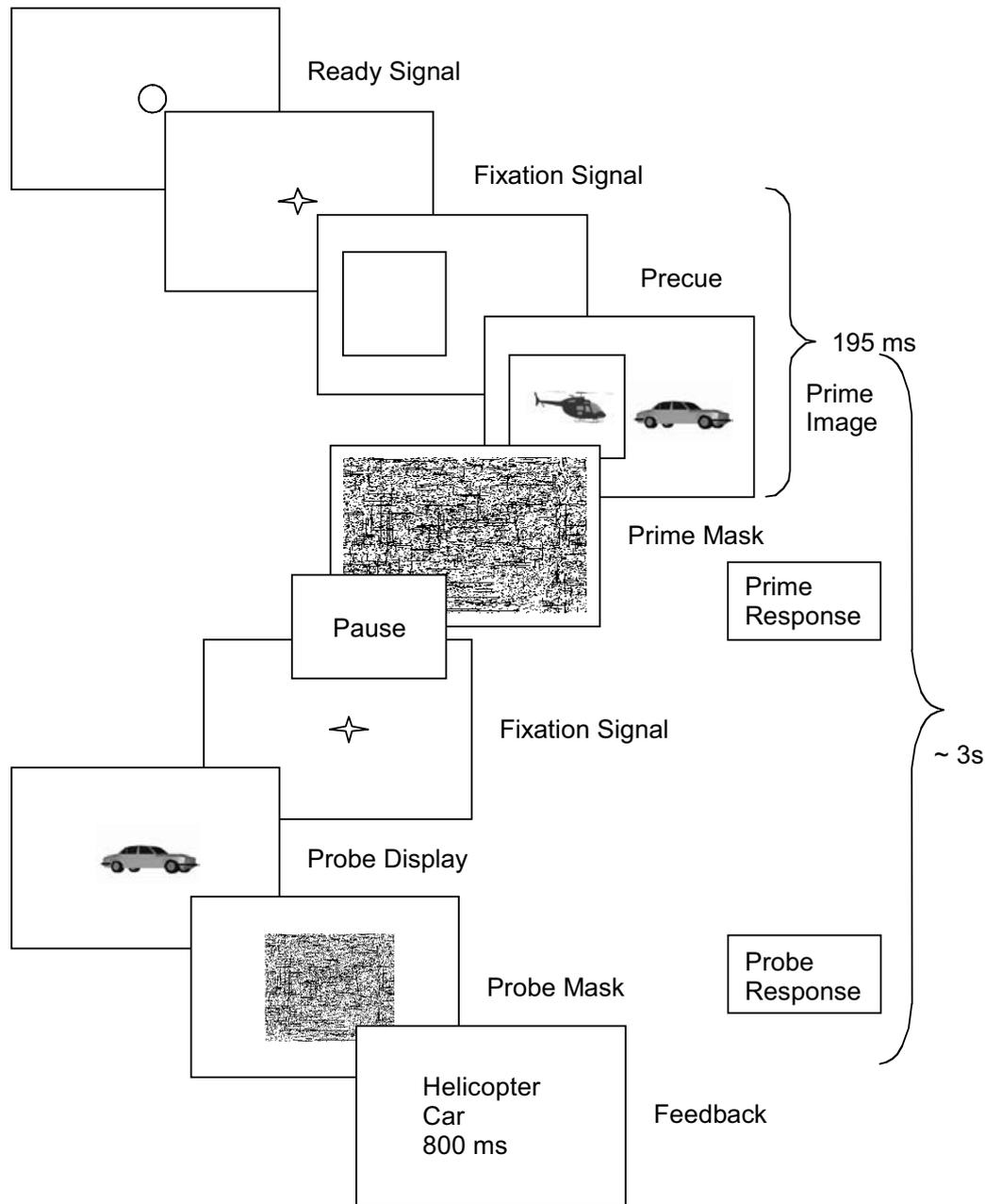


Figure 2. Sequence of displays in Experiments 1–3.

## Results

Trials on which either the prime or probe responses were incorrect were excluded from the analysis of latencies (18.6%) as were voice key errors (6.1%). 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 3).

A 2 (Attention: attended vs. ignored)  $\times$  2 (View: same view vs. reflected view) within-subjects ANOVA was performed on priming latencies. The analysis revealed a reliable main effect of attention,  $F(1, 27) = 42.92, p < .001$  and a main effect of rotation,  $F(1, 27) = 4.77, p < .05$ . The interaction between attention and rotation was not reliable,  $F(1, 27) < 1$  (see Figure 3). A Friedman ANOVA on probe errors revealed no significant effect,  $\text{Chi Sqr. (3)} = 2.56, p > .46$ .

Table 1. Mean response times (RT, in ms), standard errors, and percentage errors for probe objects Experiment 1.

Variable	Attended		Ignored		Unprimed	
	Same	Reflected	Same	Reflected	Same	Reflected
RT	630	671	769	814	835	799
SE	20	23	23	22	32	30
% errors	5	6	10	11	4	4

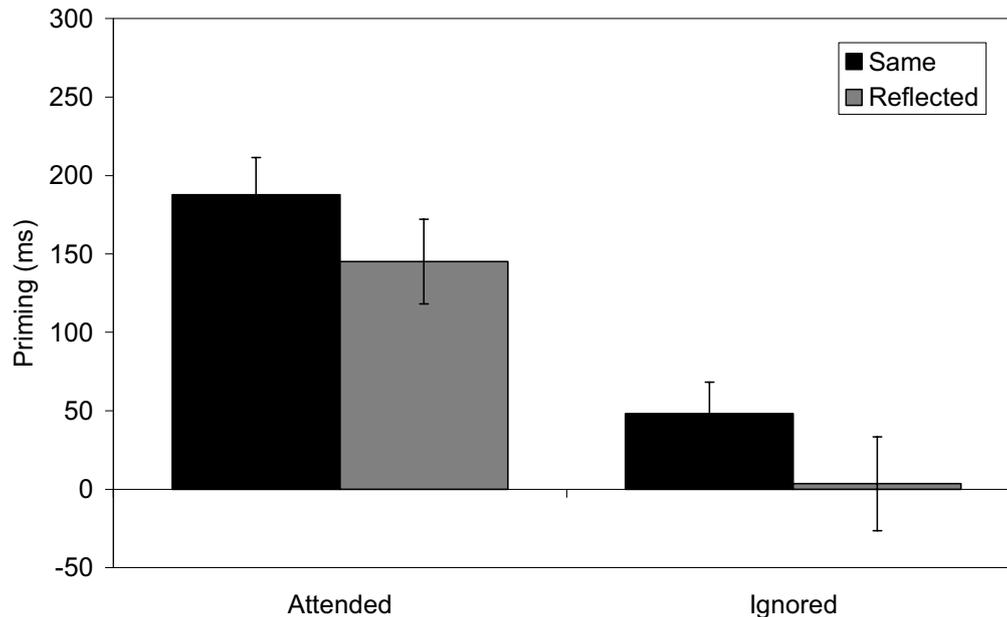


Figure 3. Priming means in ms and standard errors for Experiment 1 as a function of whether the object was attended or ignored in the prime display and whether the probe objects were presented in the same orientation or in a mirror-reflected view ( $n = 28$ ).

Matched pairs  $t$  tests showed priming reliably greater than zero in the attended-same,  $t(27) = 7.93$ ,  $p < .001$ ; attended-reflected condition,  $t(27) = 5.37$ ,  $p < .001$ ; and ignored-same condition,  $t(27) = 2.41$ ,  $p < .05$ , but not in the ignored-reflected condition,  $t(27) < 1$ ,  $p > .05$ . Thus, attended images in the prime display primed the probe image in both the same and the reflected view but ignored images primed the probe object only when it was presented in the same view.

An additional ANOVA on priming RTs was performed excluding the 8 objects that were not strictly symmetrical and consequently whose mirror-reflected views were not equivalent to a  $120^\circ$  depth-rotation (e.g., boot, desk, etc.). As with the complete set, there was a reliable main effect of attention,  $F(1, 27) = 31.02$ ,  $p < .001$  and a main effect of rotation,  $F(1, 27) = 4.37$ ,  $p < .05$ , but no reliable interaction,  $F(1, 27) < 1$ .

The last trial was immediately followed by the question of whether the participant had recognized the ignored object in the prime display. Twenty-six observers responded with “no”; two responded with “yes”, but could not name the object correctly.

## Discussion

Experiment 1 replicated the results of Stankiewicz et al. (1998) that attended objects prime both themselves and their reflected versions, whereas ignored objects only prime themselves but not their mirror versions. The effects of attention and viewpoint were additive, meaning that the advantage for same versus reflected images was equivalent in both the attended and ignored conditions. These results are in line with other studies finding visual priming for ignored ob-

jects (Stankiewicz & Hummel, 2002; Thoma et al., 2004; Tipper, 1985) and showing an advantage for same views over mirror images (Lawson & Humphreys, 1996; Stankiewicz et al., 1998).

The results obtained with this stimulus set are remarkable because they almost exactly replicate the findings of Stankiewicz et al. (1998) which were obtained with line drawings. Thus, grey-shaded images of computer-generated common objects and line drawings by an artist (Snodgrass & Vanderwart, 1980) yield the same priming effects, implying the possible involvement of edge-based representations (encoding parts derived from vertices and contours, e.g., geons) in object recognition. Edge-based approaches (Hummel & Biederman, 1992; Lowe, 1987), have been criticized (Sanocki, Bowyer, Heath, & Sarkar, 1998) because in many studies or simulations their proponents mainly use line-drawings that contain no ambiguous contours or edges (e.g., resulting from shading and highlights) as compared to real objects (or photographs). Although Experiment 1 is not a direct test of whether the more realistic computer-generated grey-level images and the idealized line-drawings are treated equally by the visual system, the findings demonstrate that the priming patterns found in previous experiments are not limited to the use of line drawings.

There is an alternative explanation of the additive priming effects, which is that there is no qualitative difference between priming for attended and ignored items but that unattended items simply show less priming than attended ones, which is further reduced to baseline after a view-change. However, not all experiments of this type produce the additive pattern and so the alternative interpretation is challenged by recent studies. In two experiments Stankiewicz and Hummel (2002) showed that priming for ignored images is invariant with view changes such as translation and scale. Thus, ignored stimuli can allow priming after some changes. At the same time ignored stimuli do not always prime even their own identical images – providing further evidence against the alternative interpretation. Thoma et al. (2004; Experiment 3) found that while attended split objects primed their identical split image as much as attended intact images primed themselves, ignored split images did not prime themselves (intact images did prime themselves). This result is in line with the hybrid model as a split object has no stored holistic representation unless it was previously attended and encoded (for more details on the locus of priming see Thoma et al., 2004). Thus, the alternative hypothesis of additive priming pattern (attended-primed-more and same-image-primed-more) has been refuted in three published experiments. Further (though more indirect) evidence of

two qualitatively different object representations comes from another variation of the priming paradigm gauging the involvement of semantic and visual priming components.

In the present experiments, recognition and naming were confounded in the attended conditions because participants had to name the cued (attended) object as well as the following probe object. Thus, not all the priming in the attended condition is visual – the observed priming will contain a semantic or name 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 (e.g., Biederman & Cooper, 1991b) estimated visual priming by substituting the image in the identical conditions (e.g., “grand piano”) with a different object (e.g., “upright piano”) that had the same basic-level name (“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 subtracting the priming for different exemplars from mirror-reflected (or split) objects of the same shape produced a conservative estimate of a visual priming component of about 80 ms, reflecting the visual priming derived from an analytic representation. As Experiment 1 replicates both the priming pattern of Stankiewicz et al. (1998, Experiment 1) and Thoma et al. (2004, Experiment 1), it is reasonable to assume that the priming found in the attended conditions in the current experiment contained a significant and large visual component (note that absolute priming levels were also almost identical across these studies).

## Experiment 2: Priming for Depth-Rotated Objects with Part Changes

In previous studies (and in Experiment 1) the hybrid model was tested by manipulating only the holistic properties of an image: Reflecting or splitting an image was predicted to affect only the holistic component; the analytic component should not be affected because the same parts were visible between prime and probe displays. Within the hybrid model, the differences in priming between the same and the mirror views reflect the missing priming component from the holistic representation after view changes. In consequence, Experiment 1 showed that the effects of viewpoint and attention are additive.

In contrast to mirror reflections, rotations in depth between study and test can affect the analytic repre-

sensation because visible parts may be occluded or new parts may be revealed (Biederman & Gerhardstein, 1993). Depth-rotations different from those in Experiment 1 should therefore provide an opportunity to further test the theory that two representations work in parallel because depth rotation may affect both representational components (analytic and holistic) instead of just one (holistic). The aim of Experiment 2 is to test whether depth-rotation involving part changes affects priming for attended objects (analytic plus holistic representation) more than for ignored objects (holistic representation only). The logic of the experiment is in three parts: First, according to the hybrid model, all viewpoint changes (except translation and scaling) should affect the holistic component. Second, because the holistic representation works with and without attention, changes in viewpoint by depth-rotations should equally decrease the amount of priming in both attended and ignored conditions compared to priming in the identical viewpoint. Third, depth-rotations that affect the perceived part structure of an object should additionally reduce the amount of priming for attended images (because only then the analytic representation will be affected), but not for ignored images. In summary, if a part-based representation is involved for attended images but not for ignored ones, object rotations involving part changes should affect priming for attended images (holistic and analytic change) more than for ignored images (holistic change only).

In Experiment 2, objects were rotated in depth to produce an altered part-structure between views. However, the degree of part change is not always systematically related to the degree of angular rotation for both natural (Lawson, 1999) and novel objects (Willems & Wagemans, 2001). For example, Figure 4 shows a camel in three different views rotated in depth. Although in 4b the camel is shown rotated further away from view 4a than from view 4c, it shares

more visible parts with view 4a, because two of the legs are hidden in view c, while at the same time we can now see a “new” part – the tail. In 4c the camel is shown in an accidental view (Biederman, 1987; Blanz et al., 1999): It is a complete side view that occludes parts of an object or makes the extraction of parts more difficult. The criterion for an accidental view is that small changes in orientation produce considerable changes in the part structure (Biederman & Gerhardstein, 1993). This applies to complete side views because slight rotations would reveal new parts or new contours and surfaces of parts.

To achieve a qualitative change in view orientation, objects in Experiment 2 were depicted in two views. One was a complete side view (or “planar” view, Blanz et al., 1999) that would be primed by a more conventional (depth-rotated view) or vice versa. Srinivas (1995) used a similar logic to create “part-occluded” objects. The effect of part-change was investigated in a pilot study.

## Method

### Participants

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

### Materials

A pilot study was conducted to test whether the pairs of  $+30^\circ$  and  $+90^\circ$  views (Figure 4b and 4c) produced perceived part changes. To assess if these two views of objects differ qualitatively from each other in their



*Figure 4.* Three views of an example object as used in the pilot study for Experiment 2. View b (termed  $+30^\circ$ ) is rotated further away – with an angular separation  $90^\circ$  – from view a (termed  $-60^\circ$ ) than from view c (termed  $+90^\circ$ ) – with an angular separation of  $60^\circ$ , but view b shares more visible parts with view a, because two of the legs are hidden in view c, but the tail is now visible.

part structure compared to other pairs of depth-rotated views a rating study was conducted. Seven independent observers from the Goldsmiths College student community were shown 87 objects in two pairs of views each on a computer screen. Three views were constructed by rotating an object  $-60^\circ$ ,  $+30^\circ$ , and  $+90^\circ$  (views a, b, c in Figure 4) from a standard frontal view where the axis of elongation or the symmetry axis of the object coincides with the viewing direction of the observer (line of sight). The participants had to compare 2 view-pairs of an object. They saw an object in the  $-60^\circ$  versus the  $+30^\circ$  view in one trial as well as in the  $+30^\circ$  versus the  $+90^\circ$  view in another trial. The task of the observers was to indicate whether crucial parts of an object were visible only in one of the two views (by pressing the "P" key for "part change") or whether the two views basically depicted the same parts (by pressing the "S" key for "same part structure"). To provide the participants with a scale of what is meant by a part they were introduced to the concept of parts and corresponding objects by using Biederman's (1987) illustration of geons and geon-based objects. The order of trial (view-pair) presentation was completely randomized. Participants had as long as they wished to press the "P" key if they thought there was a part change or the "S" key if the same parts were visible in both views.

For the subset of the objects used in Experiment 1, two one-way ANOVAs on the factor View-pair were performed, first with objects and then participants as random factor. The factor levels were the two types of view-pairs separated by depth-rotation ( $-60^\circ$  and  $+30^\circ$  vs.  $+30^\circ$  and  $+90^\circ$ ) with the number of "P" (i.e., part change) responses as dependent variable. An ANOVA with participants as random factor revealed a main effect of type of View,  $F(1, 6) = 26.35$ ,  $p < .01$ , as did the ANOVA over items,  $F(1, 86) = 49.07$ ,  $p < .001$ . Thus, objects shown in the  $+30^\circ$  and  $+90^\circ$  view pair were perceived to exhibit more part changes (mean 4.11, SE .29) across the two views than when shown in the  $-60^\circ$  and  $+30^\circ$  view pair (mean 1.91, SE .26).

The priming experiment used the same 56 objects as in Experiment 1. Prime objects were depicted in two different views counterbalanced across two participant groups: Objects in group 1 were shown in an orientation rotated  $+90^\circ$  off the line of sight in prime displays, and then shown in the same view or rotated  $+30^\circ$  off the line of sight in probe displays, and vice versa for group 2. As in Experiment 1, the objects were counterbalanced across participants so that each object would serve in each condition equally often. The general set-up of the experiment was the same as in Experiment 1.

## Procedure

The procedure was the same as in Experiment 1, except that participants were not asked whether they recognized the ignored image in the last trial. There were six priming conditions (attended-same, attended-rotated, ignored-same, ignored-rotated, unprimed-same view and unprimed-rotated view) in which each of the objects appeared equally often.

## Results

In Figure 5 the priming results of Experiment 2 are given as savings in response times relative to the baseline (unprimed) condition. Trials on which either the prime or the probe responses were incorrect were excluded from the statistical analysis of latencies (20.2%), as were voice key errors (4.1%). The baseline latencies for each of the two probe views were 866 ms (*SE* 38.5) for the  $30^\circ$  probe view and 857 ms (*SE* 27.7) for the  $90^\circ$  probe view (collapsed over groups).

For all conditions, priming was calculated as the difference between each participant's mean response time in the relevant baseline (unprimed) condition and the participant's mean response times in each of the corresponding priming conditions. A 2 (Group: prime view  $30^\circ$  vs.  $90^\circ$ )  $\times$  2 (Attention: attended vs. ignored)  $\times$  2 (View: same vs. rotated) mixed analysis of variance (ANOVA) revealed no reliable effect of group (i.e., the two orientations primed their corresponding probe equivalently),  $F(1, 26) < 1$ , a reliable main effect of attention,  $F(1, 26) = 47.15$ ,  $p < .001$ , and view,  $F(1, 26) = 8.37$ ,  $p < .001$ . The only significant interaction was between attention and view,  $F(1, 26) = 5.04$ ,  $p < .05$ . The difference between the attended-same and attended-rotated conditions (collapsed over groups) was statistically reliable,  $t(27) = 3.73$ ,  $p < .001$ , but not the difference between the ignored-same and ignored-rotated conditions,  $t(27) = 1.27$ ,  $p > .05$ . A Friedman ANOVA on probe errors for each priming condition revealed no significant effects, Chi Sqr. (3) = 1.48,  $p > .68$ .

Matched pairs *t* tests were conducted on each priming condition to determine which type of prime display caused savings in response time for the probe display. Priming was reliably greater than zero in the attended-same condition,  $t(27) = 7.78$ ,  $p < .001$ ; attended-rotated condition,  $t(27) = 4.00$ ,  $p < .001$ ; and ignored-same condition,  $t(27) = 2.95$ ,  $p < .01$ , but not in the ignored-rotated condition,  $t(27) = 1.37$ ,  $p > .05$ . Attended images in the prime display primed the probe image in both the same and rotated view, but ignored images primed the probe object only when presented in the same view.

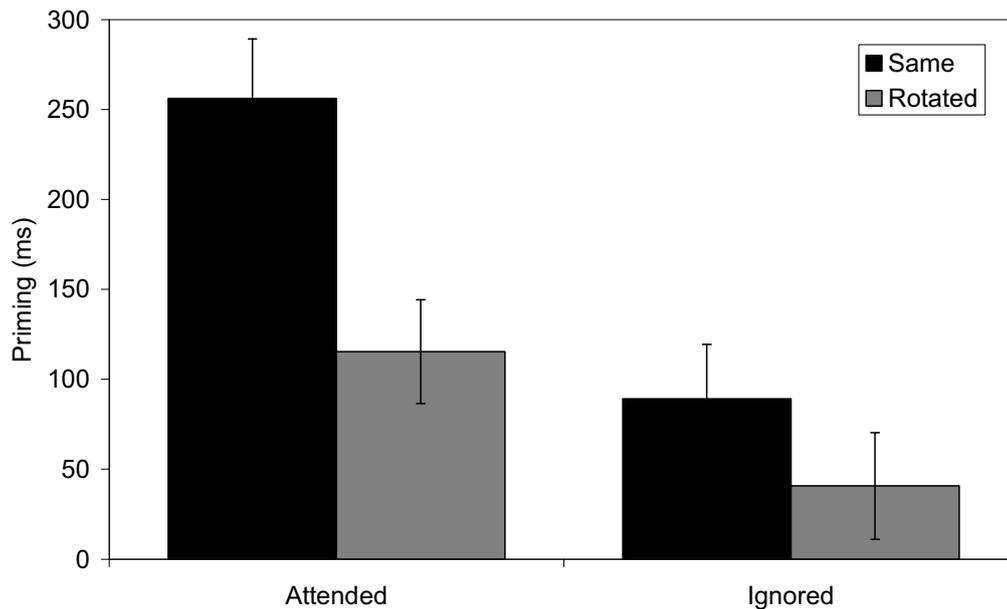


Figure 5. Priming means in ms and standard errors for Experiment 2 as a function of whether the object was attended or ignored in the prime display and whether the probe objects were presented in the same orientation or in a depth-rotated ( $60^\circ$ ) view ( $n = 28$ ).

Table 2. Mean response times (RT, in ms), standard errors, and percentage errors for probe objects Experiment 2.

Variable	Attended		Ignored		Unprimed	
	Same	Reflected	Same	Reflected	Same	Reflected
RT	644	708	810	783	899	824
SE	22	26	28	27	37	27
% errors	7	5	8	9	5	5

## Discussion

The results of Experiment 2 replicate the previous findings of priming for attended images in the same view and in a changed (here: depth-rotated) orientation while ignored objects only primed themselves in the same view. Unlike previous tests of the hybrid model, the data show an interaction between attention and view-change: The difference between identical and depth-rotated views was significantly greater for attended than for ignored images. This novel priming pattern is in line with the prediction of the hybrid model that depth-rotations may cause qualitative changes in analytic representations that depend on attention.

The priming observed for images repeated in the same view was slightly higher in Experiment 2 than in Experiment 1. This is probably due to longer response times for unprimed stimuli, which are likely to be a consequence of Experiment 2 using views (in particu-

lar the  $30^\circ$  orientation) that were slightly less canonical than the ones used in Experiment 1 (see Blanz et al., 1999). These views take longer to recognize and may profit more from priming (see, e.g., Rensink, 2000). Interestingly, priming for unattended objects is now also greater than zero. Although this amount of priming is insignificant, it indicates that the non-additive priming pattern is not just due to a floor effect for ignored images. The difference between ignored images is still  $\sim 50$  ms, but could have been greater if priming for changed views was zero. Rather, the interaction effect was solely due to the reduction between attended primes (presumably the analytic component), as predicted by the model.

The priming effects obtained in Experiment 2 cannot be attributed to difficulties with the objects' specific orientation in depth. The present experiment, as well as other studies (e.g., Hayward, 1998), show that planar views (here:  $+90^\circ$ ) prime themselves as much as other non-planar (canonical) views. The observa-

tion that part changes such as occlusion account for the striking priming differences in the attended conditions is also supported by the findings of Srinivas (1995; Experiment 2). She manipulated part changes across depth-rotations in a similar way (selecting views in which parts of photographed objects were occluded by other parts). Srinivas's participants were shown objects rotated in depth (67°, 130°, and a part-occlusion rotation) for 300 ms in the prime display. Matching object identity did not affect latencies in the view conditions with all parts visible but increased response times in the part-occluded condition.

The finding that depth-rotations affect priming differences for rotated attended objects more than in ignored conditions confirms the prediction of the hybrid model that two qualitatively different representations are employed. Most current view-based accounts do not specify the role of attention and therefore could not have predicted the results described here. However, if attention plays a role in matching representations based on metric properties, one would expect enhanced priming effects for rotated objects in attended conditions relative to ignored conditions because attention would serve to aid the matching process (e.g., Olshausen, Anderson, & van Essen, 1993).

The results also do not fit entirely with theories proposing structural descriptions. These accounts would not have predicted priming in the ignored route because structural descriptions rely on attention to actively bind local features into parts and then parts and relations to objects (see Hummel & Biederman, 1992). However, the results for depth-rotations in the attended conditions are in line with structural description accounts. They would have predicted larger priming costs obtained for rotations that cause part changes compared to mirror reflected views (Biederman & Gerhardstein, 1993). Rotations resulting in part changes will alter the activation pattern of a structural representation between prime and probe view, which means a reduction in priming compared to exactly the same parts being visible in both events (e.g., in mirror images). Experiment 3 investigates priming effects for depth-rotations (other than mirror-reflections) that do not affect visible parts.

### Experiment 3: Priming for Depth-Rotated Objects without Part Changes

Experiment 2 showed that depth-rotations that produce part-changes incurred greater priming costs for previously attended versus ignored objects but it remains to be shown that this is not the case for any other depth

rotation. Thus, Experiment 2 was designed to test whether the priming pattern observed for mirror images (Experiment 1; Stankiewicz et al., 1998) can be replicated with depth-rotated views in which the same parts are visible. The critical assumption is that depth-rotated objects – like mirror images – can be shown in orientations that reveal equivalent part structures but differ substantially in their metric (holistic) similarity.

In Experiment 3 there was a greater degree of angular separation (90°) between the prime and probe view than in Experiment 2 (60°) but the view-pairs were rated as more similar (in terms of shared parts) than the view-pairs of Experiment 2. The two views used were “off-axis”, that is they are shown in orientations that do not fall in the line of sight or perpendicular to it. These “canonical” views have been found to be the easiest to recognize (Boutsen, Lamberts, & Verfaillie, 1998; Palmer, Rosch, & Chase, 1981) and are often rated as the most typical views in which objects appear (Blanz et al., 1999; Palmer et al., 1981; Verfaillie & Boutsen, 1995). Objects were bilaterally symmetric such that a rotation across the line of sight produced a view in which roughly the same parts are visible (see Figure 4a and 4b).

The predictions from the hybrid model of Hummel (2001) are that viewpoint and attention produce additive effects of priming for qualitatively similar views. The analytic representation should not be affected by depth rotations that do not substantially change the part structure, which means that its contribution towards overall priming remains equivalent between the identical and the depth-rotated view. In contrast, the holistic representation should change considerably due to the depth-rotation, because different surfaces of the same object project to different units in the holistic surface map. Its contribution to priming should go towards zero.

## Method

### Participants

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

### Materials

A set of 84 objects was used (see Appendix). All objects were shown in two standard views (see Figure

4a and 4b). These were separated by 90° rotation around the y-axis from each other: The two views were created by rotating objects +30° (i. e., one of the views in Experiment 2) and -60° (the standard view of Experiment 1) from the line of sight (i. e., the line of symmetry). Of the 84 objects, 30 objects were used as target objects, and the rest were used as filler items (i. e., in unprobed conditions). This design was used to reduce the error rates and to boost the statistical power (see Stankiewicz & Hummel, 2002). The target objects were placed into 5 subsets which appeared equally often across participants in all conditions (attended-identical, attended-rotated, ignored-identical, ignored-rotated, and unprimed). The filler items appeared randomly in one of the conditions as unprobed prime object (either attended or ignored). The two standard views for probe displays were counter-balanced across 2 groups (-60° in group 1, and +30° in group 2). The prime objects were displayed in either one of the two views depending on the experimental condition (same view vs. rotated) and intermixed in the unprimed (baseline) condition.

## Procedure

The procedure was the same as in Experiment 2, except that the probe object in each group of participants was always depicted in the same general orientation and the prime object was either in the same orientation or rotated.

## Results

Trials on which either the prime or probe responses were incorrect were excluded from statistical analysis (7.92%), as were voice key errors (3.83%). The latencies for the baseline probe views were similar in both groups: 770 ms (*SE* 31.8; for the -60° orientation) and 803 ms (*SE* 29.7, for 30° orientation), a non-significant difference,  $t(38) < 1$ . Figure 5 shows the priming results of Experiment 3 as savings in response times

*Table 3.* Mean response times (RT, in ms), standard errors, and percentage errors for probe objects Experiment 3.

Variable	Attended		Ignored		Unprimed
	Same	Reflected	Same	Reflected	
<i>RT</i>	601	627	726	770	790
<i>SE</i>	13	14	16	16	22
<i>% errors</i>	2	2	3	2	3

relative to the baseline (unprimed) condition (see also Table 3).

A 2 (Group: probe view 30° vs. -60°) × 2 (Attention: attended vs. ignored) × 2 (View: same vs. rotated) mixed analysis of variance (ANOVA) revealed no reliable effect of group (i. e., priming patterns in the two probe orientation groups did not differ),  $F(1, 38) < 1$ , a reliable main effect of attention,  $F(1, 38) = 105.13$ ,  $p < .001$ , and view,  $F(1, 38) = 10.79$ ,  $p < .01$ . There was no statistically reliable interaction. A Friedman ANOVA on probe errors revealed no significant effects, Chi Sqr. (3) = 1.50,  $p > .68$ .

Matched pairs *t* tests were conducted on each priming condition (collapsed over groups) to determine reliable savings in response time compared to the baseline. Priming was reliably greater than zero in the attended-same condition,  $t(39) = 8.61$ ,  $p < .001$ ; attended-rotated condition,  $t(39) = 7.93$ ,  $p < .001$ ; and ignored-same condition,  $t(39) = 3.47$ ,  $p < .01$ , but not in the ignored-rotated condition,  $t(39) < 1$ . Probe images were successfully primed by attended images in the prime display shown in both the same and rotated view, but ignored images primed the probe image only when presented in the same view. The difference between the attended-same and attended-rotated conditions was statistically reliable,  $t(39) = 2.41$ ,  $p < .05$ , as was the difference between the ignored-same and ignored-rotated conditions,  $t(39) = 2.30$ ,  $p < .05$ .

An additional ANOVA was run contrasting Experiment 2 and 3: The 2 (Experiment: probe view 30° vs. -60°) × 2 (Attention: attended vs. ignored) × 2 (View: same vs. rotated) mixed analysis of variance (ANOVA) revealed no reliable effect of experiment,  $F(1, 66) < 1$ , but a reliable main effect of attention,  $F(1, 66) = 132.05$ ,  $p < .001$ , and view,  $F(1, 66) = 18.92$ ,  $p < .001$ . There was only one significant interaction: this was the 3-way interaction between Experiment × Attention × View,  $F(1, 66) = 7.05$ ,  $p > .01$ . The difference between attended images (same view and rotated) was larger in Experiment 2 than in Experiment 3, whereas the difference between ignored conditions (same and rotated view) remained the same across experiments.

## Discussion

The priming pattern observed with depth-rotated objects in Experiment 3 is the same as in Experiment 1 with mirror-reflections. Probe objects were primed by previously attended images presented in the same view as well as in a changed (here: depth-rotated)

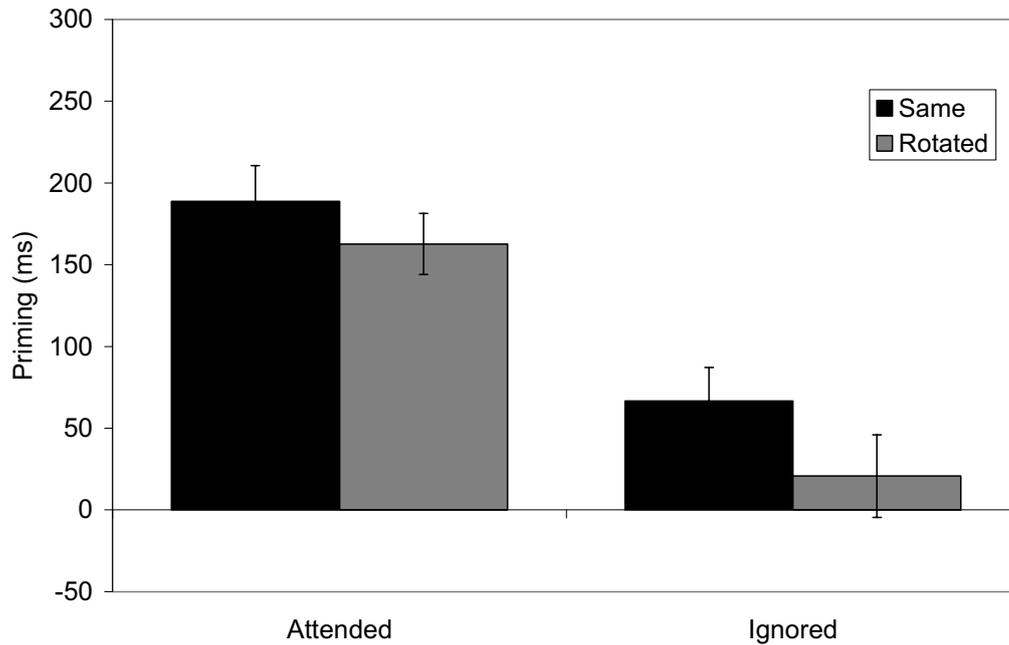


Figure 6. Priming means in ms and standard errors Experiment 3 as a function of whether the object was attended or ignored in the prime display and whether the probe objects were presented in the same orientation or rotated in depth ( $n = 40$ ).

view, whereas a probe image was only primed by an ignored prime if it was presented in the same view. The effects of attention and view were additive: Objects primed themselves more in the same orientation than in a rotated view in both attended and ignored conditions.

The results are in line with other studies employing depth-rotations similar to the ones in Experiment 3. Lawson and Humphreys (1996,1998) studied effects of long-term (over several minutes) priming for line-drawings of common objects rotated in depth. Their results (using very similar objects) show that naming was faster when the prime and probe object were identical relative to when they were rotated. Biederman and Gerhardstein (1993) also used views similar to ours in their Experiments 1 and 2. In long-term (i. e., several minutes) priming studies they, too, found only slight differences for rotations ranging from  $33.75^\circ$  to  $135^\circ$  for common objects if there were no part changes but obtained a marked reduction in priming if different parts were visible across the study and test view.

The priming pattern in Experiment 3 is clearly predicted by the hybrid model of object recognition because the views employed here were equivalent in part-structure and should therefore produce additive priming effects for attention and viewpoint. When a prime

object is attended, the holistic surface representation is activated in parallel with the analytical part-based representation. On subsequent presentation of the same object in the same view, both the units coding parts and spatial relations (i. e., the analytic representation) and the units coding the location of object surfaces (i. e., the holistic component) benefit from the previous presentation resulting in faster recognition. However, when presented with the identical but depth-rotated object, the same parts (and relations) are presented to the analytical unit, resulting in an activation (i. e., analytic priming component) that is equivalent to that of the identical view. At the same time, the locations of surfaces (if projected on a 2D grid) have changed considerably after depth-rotations. Therefore, the activation pattern of the holistic surface units is very different between prime and probe trial and no priming is predicted from the holistic component.

## General Discussion

In three experiments using a short term priming paradigm attended and ignored objects primed themselves in the same view, but only attended objects primed themselves when presented in a depth-rotated (or mirror-reflected) view. Ignored objects never primed their depth-rotated images. View changes caused an equiva-

lent reduction in priming for both attended and ignored images, provided these view-changes did not alter the perceived part structure (Experiments 1 and 3). However, if considerable part changes occur between prime and probe view, the reduction of priming for previously attended objects is greater than for ignored objects. These results indicate that attended objects are treated qualitatively different from ignored objects as predicted by the hybrid model of object recognition (Hummel, 2001). According to this model, an analytic part-based representation mediates recognition only for attended objects, whereas a holistic representation underlies both the recognition of ignored and attended images.

The present data are in accordance with previous studies that show a consistent advantage of same-view priming (Lawson & Humphreys, 1996; 1998) and that priming across views is particularly reduced for planar views (Hayward, 1998). The results for attended images also confirm that object recognition depends on whether the same parts are visible across views (Biederman & Gerhardstein, 1993; Srinivas, 1995). The hybrid's model general notion that object recognition across rotations in depth involves both an analytic and a holistic representation is also corroborated by Foster and Gilson (2002). They used novel 3-D objects that were to be discriminated in matching tasks either by a metric or a non-accidental (i. e., structural) property. Discrimination performance after changing the number of parts was superior to metric changes. Performance dropped for changes in both structural and view-specific features as the objects were rotated in depth but there was no interaction between structural and metric changes. Foster and Gilson proposed that their results show evidence for object recognition based on an additive relationship of two independent representations – one part-based and one image-based. Thus, although their model does not incorporate attention, the work of Foster and Gilson (2002) concurs with the notion of a hybrid representation of object shape.

There is also additional support for the hybrid model from recent functional imaging studies. In a fMRI study, 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. The latter area was sensitive to changes in orientation but not in size, which is directly predicted by the hybrid model (Hummel, 2001) and was confirmed in behavioral studies (Stankiewicz & Hummel, 2002). Neuropsychological evidence also further supports the notion of two components in object recognition. Davidoff and Warrington (1999;

2001) studied patients who were extremely impaired at recognizing object parts. Nevertheless, they were normal in naming intact objects though only when seen in familiar views. In terms of the hybrid model, the patients' holistic components seemed intact, allowing object recognition from familiar views, whereas analytic components were impaired preventing recognition of object parts or from unfamiliar views.

An alternative explanation for the effects of depth rotation has been proposed by Hayward (1998). Hayward (1998; Experiments 3 and 4) employed depth-rotated views almost identical to the ones in Experiment 2. He found that objects in planar views (which were separated by 60° in depth from the study view) were matched and named more slowly than a non-planar view rotated 180° from the standard view. These results are very similar to the present data in the attended conditions. However, Hayward's conclusion was that not part-changes but common outline shape was the crucial factor for object constancy across 3D view changes. This explanation is based on the logic that the 180° rotations employed dramatically changed the part-structure but only marginally varied the outline contour (because it arguably approximated a mirror version of the original object). This explanation cannot account for the data here. First, we found an almost identical priming difference between same and rotated (Experiment 3) and same and reflected (Experiment 1) views, although the two views in Experiment 3 were rotated unevenly from the line-off sight (30° and 60°) resulting in an outline contour that was separated by 30° from its exact mirror version. Second, an "outline-similarity" account does not fit with the priming data in the ignored conditions: Priming reductions across views for ignored objects in Experiment 2 were less reduced than in the attended conditions. Third, similarity of outline shape cannot account for other data very similar to those obtained in Experiments 1 and 3. For example, Thoma et al. (2004) found that split objects (with an outline contour very different from their original versions) produce almost exactly the same priming pattern as mirror-reflected (Stankiewicz et al., 1998) images.

Both our and Hayward's (1998) data are not in line with classic view-based accounts which assume a linear increase (e.g., Tarr & Pinker, 1989) or even an accelerated increase (e.g., Poggio & Edelman, 1990) of latencies with increased orientation changes: For attended conditions, a 60° rotation in Experiment 2 produced greater priming costs than a 90° rotation in Experiment 3 or even a 120° rotation in Experiment 1. In addition, to the best of our knowledge, view-based theories would also not predict priming for ignored objects (see Olshausen et al., 1993). More important,

view-based approaches could not explain the interaction of viewpoint and attention in Experiment 2. Even if view-based accounts could be amended to include priming for ignored objects, they could not explain why a qualitative change in orientation affects attended objects more than unattended ones.

The priming pattern also does not fit entirely with structural description accounts. First, equivalent views in Experiments 1 and 3 incurred priming costs, which is at odds with a central claim of view-independence (Biederman, 1987; Hummel & Biederman, 1992). Geon theory would not have predicted a reduction in priming for depth-rotated objects in the attended conditions (see Biederman & Gerhardstein, 1993) because the two views showed very much the same part structure and were far from accidental views. Of course, it could be argued that every depth-rotation changes the visibility of some parts, and may produce spurious effects of view-dependence. However, the fact that Experiment 3 obtained a very similar priming pattern as Experiment 1 with mirror-images (which by definition show exactly the same parts) make that counterargument less than convincing. Second, structural description models would not predict priming for ignored objects, because without attention part-based descriptions of objects cannot be established. Finally, structural description theories – like view-based accounts – would not be able to account for the interaction of attention and view observed in Experiment 2.

Although the data reported here clearly demonstrate a qualitative difference in priming effects for depth-rotated objects in dependence of attention, it is not necessarily the case that this difference is due to parts defined as geons. Future research needs to employ a stricter criterion for part changes in natural objects to establish whether the effects here are due to representations resembling structural descriptions. Nevertheless, the current results are clearly in line with the notion of a hybrid analytic/holistic representation of shape.

In conclusion, the results show that priming effects after depth-rotations show evidence of both view-specific and part-based representations, which, in turn, depend on attention. The present study is the first to test the hybrid model of object recognition with depth rotations. Previous studies have shown that the model can account for priming patterns after mirror-reflections (Stankiewicz et al., 1998), translation and scaling (Stankiewicz & Hummel, 2002), and changes in configuration (Thoma et al., 2004). Although only a limited range of rotations were chosen here for practical and theoretical reasons, the experiments strikingly show that the hybrid model's complex predictions concerning the interaction of different types of viewpoint changes and attention could not be falsified. The data

contribute to a growing body of evidence for independent holistic (view-like) and analytic (view-insensitive) representations of object shape.

## References

- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.
- Biederman, I. (2000). Recognizing depth-rotated objects: A review of recent research and theory. *Spatial Vision*, *13*, 241–253.
- Biederman, I., & Cooper, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology-Human Perception and Performance*, *18*, 121–133.
- Biederman, I., & Cooper, E. E. (1991a). Evidence for complete translational and reflectional invariance in visual object priming. *Perception*, *20*, 585–593.
- Biederman, I. & Cooper, E. E. (1991b). Priming contour-deleted images: Evidence for intermediate representations in visual object priming. *Cognitive Psychology*, *23*, 393–419.
- Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects – Evidence and conditions for 3-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19* (6), 1162–1182.
- Blanz, V., Tarr, M. J., & Bulthoff, H. H. (1999). What object attributes determine canonical views? *Perception*, *28*, 575–599.
- Boutsen, L., Lamberts, K., & Verfaillie, K. (1998). Recognition times of different views of 56 depth-rotated objects: A note concerning Verfaillie and Boutsen (1995). *Perception & Psychophysics*, *60*, 900–907.
- Bruce, V., Carson, D., Burton, M. A., & Ellis, A. W. (2000). Perceptual priming is not a necessary consequence of semantic classification of pictures. *The Quarterly Journal of Experimental Psychology*, *53A*, 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 United States of America*, *89*, 60–64.
- Davidoff, J., & Warrington, E. K. (1999). The bare bones of object recognition: Implications from a case of object recognition impairment. *Neuropsychologia*, *37*, 279–292.
- Davidoff, J., & Warrington, E. K. (2001). A particular difficulty in discriminating between mirror images. *Neuropsychologia*, *39*, 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. (1998). Effects of outline shape in object recognition. *Journal of Experimental Psychology-Human Perception and 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 and Performance*, *23*, 1511–1521.

- 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, 480–517.
- Hummel, J. E., & Stankiewicz, B. J. (1996). An architecture for rapid, hierarchical structural description. In T. Inui and 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*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory & Cognition*, 13, 289–303.
- Lawson, R. (1999). Achieving visual object constancy across plane rotation and depth rotation. *Acta Psychologica*, 102, 221–245.
- Lawson, R., & Humphreys, G. W. (1996). View specificity in object processing: Evidence from picture matching. *Journal of Experimental Psychology-Human Perception and 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.
- Lawson, R., Humphreys, G. W., & Jolicoeur, P. (2000). The combined effects of plane disorientation and foreshortening on picture naming: One manipulation or two? *Journal of Experimental Psychology-Human Perception and Performance*, 26, 568–581.
- Lowe, D. G. (1987). The viewpoint consistency constraint. *International Journal of Computer Vision*, 1, 57–72.
- Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, Series B*, 200, 269–294.
- McMullen, P. A., & Jolicoeur, P. (1990). The spatial frame of reference in object naming and discrimination of left-right reflections. *Memory & Cognition*, 18, 99–115.
- McMullen, P. A., & Jolicoeur, P. (1992). Reference frame and effects of orientation on finding the tops of rotated objects. *Journal of Experimental Psychology-Human Perception and Performance*, 18, 807–820.
- Murray, J. E. (1998). Is entry-level recognition viewpoint invariant or viewpoint dependent? *Psychonomic Bulletin & Review*, 5, 300–304.
- Murray, J. E. (1999). Orientation-specific effects in picture matching and naming. *Memory & Cognition*, 27, 878–889.
- Murray, J. E., Jolicoeur, P., McMullen, P. A., & Ingleton, M. (1993). Orientation-invariant transfer of training in the identification of rotated natural objects. *Memory & Cognition*, 21, 604–610.
- Olshausen, B., Anderson, C., & Van Essen, D. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *The Journal of Neuroscience*, 13, 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.
- Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, 9, 441–474.
- Palmer, S. E. (1978). Structural aspects of similarity. *Memory and Cognition*, 6, 91–97.
- Palmer, S., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 135–151). Hillsdale, NJ: Erlbaum.
- Poggio, T., & Edelman, S. (1990). A network that learns to recognize 3-dimensional objects. *Nature*, 343 (6255), 263–266.
- Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition*, 7, 17–42.
- 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.
- Sanocki, T., Bowyer, K. W., Heath, M. D., & Sarkar, S. (1998). Are edges sufficient for object recognition? *Journal of Experimental Psychology-Human Perception and Performance*, 24, 340–349.
- 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.
- Srinivas, K. (1995). Representation of rotated objects in explicit and implicit memory. *Journal of Experimental Psychology-Learning Memory and Cognition*, 21, 1019–1036.
- Stankiewicz, B. J., & Hummel, J. E. (2002). Automatic priming for translation- and scale-invariant representations of object shape. *Visual Cognition*, 9, 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, 732–744.
- Tarr, M. J. (1995). Rotating objects to recognize them – a case-study on the role of viewpoint dependency in the recognition of 3-dimensional objects. *Psychonomic Bulletin & Review*, 2, 55–82.
- Tarr, M. J., & Bulthoff, 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, 1494–1505.
- Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape-recognition. *Cognitive Psychology*, 21, 233–282.
- Tarr, M. J., & Pinker, S. (1990). When does human object recognition use a viewer-centered reference frame? *Psychological Science*, 1, 253–256.
- Tarr, M. J., Williams, P., Hayward, W. G., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint dependent. *Nature Neuroscience*, 1, 275–277.
- Thoma, V., Hummel, J. E., and 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, 257–267.
- Tipper, S. (1985). The negative priming effect: Inhibitory effects of ignored primes. *Quarterly Journal of Experimental Psychology Section a-Human 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.
- Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
- Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General*, *113*, 169–193.
- Verfaillie, K., & Boutsen, L. (1995). A corpus of 714 full-color images of depth-rotated objects. *Perception & Psychophysics*, *57*, 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*, 491–499.
- Willems, B., & Wagemans, J. (2001). Matching multicomponent objects from different viewpoints: mental rotation as normalization? *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 1090–1115.

Received December 6, 2004

Revision received February 25, 2005

Accepted March 22, 2005

Volker Thoma

---

School of Psychology  
University of East London  
The Green  
London  
E15 4LZ  
United Kingdom  
Tel. +44 208 223 4966  
Fax +44 208 223 4937  
E-mail v.thoma@uel.ac.uk

---

## Appendix

### Stimuli

#### Experiment 1 and 2

axe, banana, bike, boot, bus, camel, cannon, car, carriage, chair, coffee machine, cow, crocodile, desk, dog, dolphin, drill, eagle, fork, glasses, guitar, hammer, harp, helicopter, hippo, horse, Hoover, iron, ironing board, kettle, key, knife, lamp, microscope, motorbike, phone, piano, pig, pipe, pistol, plane, shark, ship, shoe, shovel, snail, sofa, spoon, stapler, suitcase, toilet, toothbrush, torch, turtle, watch, wrench

#### Experiment 3

Targets:

axe, bed, bike, bird, bus, camel, car, chair, cow, crocodile, dolphin, duck, glasses, gun, helicopter, Hoover, iron, lamp, motorbike, pipe, plane, scissors, ship, snail, spanner, stapler, toilet, truck, turtle, watch

Fillers:

banana, baseball bat, binoculars, boot, camera, cannon, can opener, carriage, chicken, corkscrew, cup, desk, dog, fire extinguisher, fork, guitar, hammer, harp, horse, ironing board, lock, microscope, pen, phone, piano, pig, pincers, plunger, shoe, shovel, skateboard, sofa, spoon, suitcase, tank, toothbrush