

Comparing view sensitivity in shape discrimination with shape sensitivity in view discrimination

REBECCA LAWSON

University of Liverpool, Liverpool, England

and

HEINRICH H. BÜLTHOFF

Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany

In three picture–picture matching experiments, the effects of a view change on our ability to detect a shape change (Experiments 1 and 2) were contrasted with the effects of a shape change on our ability to detect a view change (Experiment 3). In each experiment, both view changes and shape changes influenced performance. However, shape changes had more influence than did view changes in the shape change detection task. Conversely, view changes were more influential when the task was to detect view changes. Thus, the participants could often discriminate between the effects of shape changes and the effects of view changes. The disruptive effect of task-irrelevant changes (view changes in the first two experiments; shape changes in the final experiment) does not support Stankiewicz's (2002) claim that information about viewpoint and about shape can be estimated independently by human observers. However, the greater effect of variation in the task-relevant than in the task-irrelevant dimension indicates that the observers were moderately successful at disregarding irrelevant changes.

When we see a three-dimensional (3-D) object, two factors have an important influence on the image of the object that is projected onto our retina: first, the 3-D shape of the object and, second, the viewpoint from which that object is seen. For everyday object recognition, the object's shape is important, whereas its viewpoint is usually semantically irrelevant. Our visual system, therefore, needs to be able to distinguish between the important effects of shape changes and the irrelevant effects of view changes when recognizing objects. In the present experiments, we investigated how effectively our visual system achieves this goal. People were shown two pictures sequentially, each depicting a novel 3-D object. They decided whether the two objects had the same or similar shapes (Experiments 1 and 2) or were depicted from the same view in depth (Experiment 3; see Figure 1). Performance should be based on the 3-D shape similarity of the two objects in the first task and on the similarity in the depicted view of those objects relative to the observer in the second task. We investigated how well people could extract, from the two-dimensional (2-D) appearance of objects, either 3-D shape (while ignoring view changes in the first task) or viewpoint (while ignoring shape changes in the second task). If our visual object recognition system can

detect shape changes while ignoring view changes and vice versa, there should be no view sensitivity in the first task and no shape sensitivity in the second task.

The theoretical interpretation of findings of view sensitivity in studies of object recognition has been the source of heated debate (e.g., Biederman & Gerhardstein, 1993, 1995; Hayward, 2003; Lawson, 1999; Tarr & Bülthoff, 1995). The main dispute has been over why performance on object recognition tasks is often view sensitive, since there are a number of reasons why view sensitivity may be found even if shape and view information can be extracted independently.

1. *Poor visual information.* Some views may be harder to recognize because of relatively low-level, initial processing difficulties for these images. For example, some views may occlude much important shape information or may make the main axis of the object difficult to extract (such as highly foreshortened views; see Humphrey & Jolicœur, 1993; Lawson, 1999; Marr, 1982). Any such general view effects are uninformative about the recognition of particular objects. Analogously, if people were found to be slower to identify tiny pictures of animals, this would probably be due to their problems seeing detailed information for these pictures (and would be expected to be similar for any other tiny pictures). It would not provide evidence for scale specificity in either the processes or the representations involved in identifying animals.

2. *Use of explicit tasks.* Biederman and Gerhardstein (1993) and others (see Lawson, 2004a) have suggested that view sensitivity may be observed with some tasks

The first author gratefully acknowledges the support of the Leverhulme Trust and the Max-Planck-Institut für biologische Kybernetik in Germany. Correspondence concerning this article should be addressed to R. Lawson, School of Psychology, University of Liverpool, Eleanor Rathbone Building, Bedford Street South, Liverpool L69 7ZA, England (e-mail: rlawson@liverpool.ac.uk).

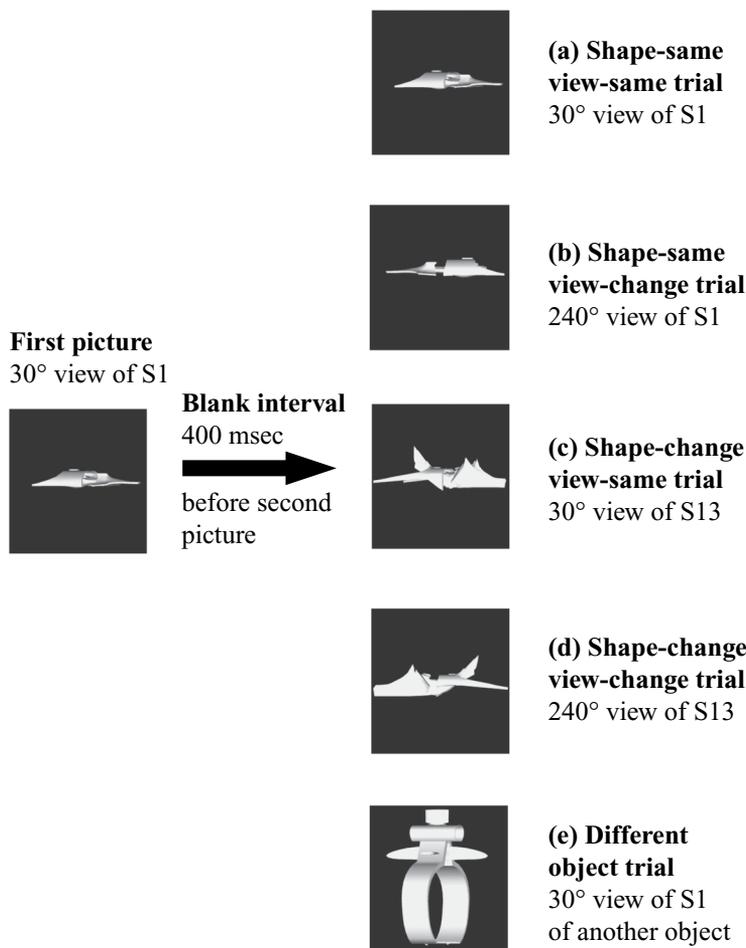


Figure 1. Illustrations of the five types of condition used in the three picture-picture matching experiments reported here. A 150° view change is shown in (b) and (d), an S12 shape change is shown in (c) and (d), and a category change is shown in (e). Experiments 1 and 2 both tested shape-change detection. For the object recognition task used in Experiment 1, the correct response was “shape same” for (a) and (b) and “shape change” for (c) and (d); (e) was not shown. For the object categorization task used in Experiment 2, the correct response was “category same” for (a), (b), (c), and (d) and “category change” for (e). For the view-change detection task used in Experiment 3, the correct response was “view same” for (a) and (c) and “view change” for (b) and (d); (e) was not shown.

(such as explicit recognition), but not with other tasks (such as implicit memory measures of object recognition; e.g., Biederman and Gerhardstein, 1993, tested the priming of object naming and sequential picture matching). Biederman and Gerhardstein (1993) further argued that only implicit tasks tapped the structural description representations that they assumed subserved everyday object recognition. Similar arguments were made by Seamon and Delgado (1999), who reported view sensitivity in an explicit *old/new* recognition task but view invariance in an affective preference task that they proposed as a measure of implicit memory representations. Here, the visual system is assumed to store both view-specific (episodic) object representations that are accessed in explicit memory tasks and view-invariant (structural description) represen-

tations that are accessed in implicit memory tasks and, importantly, that are used when we normally recognize objects. This proposal has been strongly critiqued both on theoretical grounds (e.g., Tarr & Bühlhoff, 1995) and because it cannot explain the many cases of view sensitivity reported for the implicit memory tasks that Biederman and others have argued tap view-invariant object representations. For example, Lawson and Humphreys (1998) reported view-sensitive priming of object naming, Lawson and Humphreys (1996) reported view sensitivity in a sequential picture-matching task, and Lawson (2004a) reported view sensitivity in an affective preference task.

3. *Viewpoint aiding object recognition.* Observers may respond differently to different views of an object if knowledge of viewpoint provides information about that object's

shape either prior to testing or during testing (Jolicœur & Humphrey, 1998). For example, Stankiewicz (2002) suggested that many reports that object recognition is sensitive to depth rotation (e.g., Bühlhoff & Edelman, 1992; Edelman & Bühlhoff, 1992) and plane rotation (e.g., Tarr & Pinker, 1989) may have resulted from viewpoint's being informative either within the task tested or from people's prior experience, and not because shape processing is necessarily influenced by viewpoint (as would be the case if view information and shape information were inextricably bound together). Any such view effects should improve recognition of an object from the most common views from which it has previously been seen. Some objects, such as keys, have been seen at most views approximately equally. For such objects, knowing its view would not help to decide its identity. However, we almost always see televisions from an upright, front-on view, and tables, although commonly seen at all views in depth, are typically seen only upright or near-upright. Here, view information may aid object recognition and could explain the relative ease of identifying canonical over unusual views (e.g., Palmer, Rosch, & Chase, 1981) and upright over plane-oriented views (e.g., Jolicœur, 1985, 1990). This proposal assumes that information about both the view and the shape of a particular object is stored. However, these two types of information could be stored separately and could be accessed independently, so information about object shape could still be view invariant.

4. *Shape and view information bound together.* Finally, view-sensitive object recognition may occur because view information is inextricably bound up with shape information, so these two sources of information cannot be processed independently. Here, different shape information would be activated by different views of the object, causing our visual system to respond differently. Only in this case would view and shape information be so closely integrated together that, in general, shape information could not be extracted independently of view information. In this article, we argue that only this final account can explain the view sensitivity that we observed in the first two experiments reported here and the shape sensitivity that we found in the final experiment.

Most current theories of object recognition acknowledge that view sensitivity may occur for this final reason (unlike the structural description account proposed by Marr, 1982). However, there is considerable debate as to the conditions under which such view sensitivity would be predicted. View sensitivity due to the activation of different, view-sensitive representations of the same object is expected to be the norm by view-based models (e.g., Tarr & Bühlhoff, 1995), whereas structural description accounts (e.g., Biederman & Gerhardstein, 1993, 1995) emphasize that view invariance is predicted in most cases in everyday object recognition. It is important to note that these structural description accounts do, though, acknowledge that multiple views of view-sensitive representations are necessary to represent most objects, since different views of an object often reveal different parts and the spatial relations between parts change as view changes

(Biederman & Gerhardstein, 1993). Controversy has thus centered around the circumstances under which view sensitivity is typically found—for example, whether view sensitivity normally occurs only for subordinate-level categorization (e.g., Biederman & Gerhardstein, 1995; Hamm & McMullen, 1998) or whether it is also usual in everyday, basic-level categorization (e.g., Lawson & Humphreys, 1996, 1998; Lawson & Jolicœur, 2003). Extending this research, some recent studies have examined how sensitivity to view changes is influenced by the difficulty of shape discrimination (Cutzu & Edelman, 1998; Foster & Gilson, 2002; Lawson, 2004b; Lawson, Bühlhoff, & Dumbell, 2003). Recently, a consensus appeared to be forming about the interpretation of the many findings of view-sensitive object recognition (Hayward, 2003; Lawson, 2004b) that acknowledged that both view sensitivity and view invariance can be found under different circumstances and that some findings of view sensitivity can be explained only by the final reason listed above: that view information and shape information are necessarily bound together.

This consensus was not, though, supported by Stankiewicz (2002) on the basis of his results, using a noise-masking paradigm. Stankiewicz presented single pictures of a simple, curved cylinder (shaped like a thick eyelash) to trained observers. In his third experiment, the observers had to detect shape changes due to varying the cylinder's diameter (so that it became fatter or thinner), or in other blocks, they had to detect changes in the view in depth from which the cylinder was shown. Noise was added to the task-relevant dimension (cylinder thickness for shape change detection; view in depth for view change detection), to the task-irrelevant dimension, or to both dimensions. For both the shape-change and the view-change detection tasks, Stankiewicz found that adding task-irrelevant noise did not affect performance, whereas adding task-relevant noise disrupted performance. He therefore argued that information about an object's 3-D shape can be estimated independently of information about its view in depth and vice versa.

If this conclusion was generally valid, it would provide strong evidence against view-based models of recognition (e.g., Poggio & Edelman, 1990; Tarr & Pinker, 1989). However, it would not support Biederman and colleagues' structural description model of recognition either (Biederman & Gerhardstein, 1993, 1995). Changes in the thickness of the curved cylinders could be distinguished only by using fine, metric measurements. Biederman and colleagues argued that only relatively coarse relations are coded, so Stankiewicz's (2002) stimuli would all activate the same structural description representation and hence be indistinguishable. Instead, Stankiewicz concluded that his results supported Hummel and Stankiewicz's (1998) structural description model of object recognition, which proposes that given sufficient processing time and attention, fine, metric discriminations can be made. Stankiewicz's specific finding from his third experiment was not only that his two participants could, in the shape change detection task, discriminate the thin from the thick cylinders, but also that their performance was independent of view changes (and vice versa for the view change detection task). It is

important to determine the generality of this result, since Stankiewicz tested just one shape dimension (thickness) for just one shape (a cylinder) for only a limited range of views (two sets of views for each observer). For these particular stimuli, 2-D image-based information may have been sufficient to perform accurately on the task. For example, for the view-change detection task, since cylinder curvature was fixed in the third experiment, curvature in the 2-D image could be used to estimate the view in depth of the cylinder. In the present experiments, we examined whether, in general, shape information can be extracted independently of view information.

Stankiewicz's (2002) conclusion that 3-D shape information can be estimated independently of information about an object's view in depth appears to be inconsistent with the reports of view-sensitive priming found when noncanonical, foreshortened views of familiar objects were shown as targets in priming studies (e.g., Lawson & Humphreys, 1996, 1998). In both Lawson and Humphreys's (1998) first naming experiment and Lawson and Humphreys's (1996) fourth picture-picture shape-matching experiment, responses to foreshortened targets were faster if they had been primed by foreshortened, rather than canonical, views. Conversely, responses to canonical targets were faster if they had been primed by canonical, rather than foreshortened, views. The first of these findings was the most important: faster responses to foreshortened targets primed by foreshortened views. This occurred despite foreshortened views being harder to recognize *per se* than canonical views, so the first (poor visual information) reason for view sensitivity does not hold. It occurred despite the use of an implicit priming task, so the second (use of explicit tasks) reason for observing view sensitivity did not hold. Also, it occurred despite foreshortened views being less familiar prior to the study and participants being shown equal numbers of foreshortened and canonical views within the study, so the third (viewpoint aiding recognition) reason for view sensitivity did not hold. Instead, any effect of prior experience should have benefited the identification of canonical over foreshortened views. Only the fourth reason, that shape information and view information are bound together, can explain why view sensitivity was found in this critical condition.

Recently, this work was extended by Lawson et al. (2003) and Lawson (2004b). They reported systematic changes in view sensitivity in picture-picture matching tasks that manipulated the difficulty of detecting shape changes. In these studies, novel objects were presented (so there should have been no prior expectations linking particular views to specific objects) and extensive training was not provided to participants (so there was little opportunity for the participants to learn correlations between object shape and object view). In addition, in Lawson et al.'s (2003) experiments, the two objects shown on mismatches always had similar shapes and were shown at the same view as the objects presented on matches. Here, the views of the two objects shown on a trial provided no information as to whether there was a shape change. Nevertheless, responses were very different on view-same

than on view-change trials in all the experiments. Furthermore, view sensitivity ranged from strong to nothing, depending on the difficulty of shape discrimination (from hard for matches to easy for mismatches presenting two dissimilar objects). Finally, view-sensitive priming to depth rotation has also recently been reported in an affective preference task that exploited the mere exposure effect (Lawson, 2004a).

Together, these findings reveal that there is normally a cost to achieving object constancy over depth rotation, particularly when similar-shaped objects must be discriminated. This indicates that view information and shape information are not processed separately. However, this leaves open the question of why Stankiewicz's (2002) third experiment showed evidence that object shape and object view could be estimated independently. In Stankiewicz's experiments, views of single curved cylinders were presented, and he noted that his results might not generalize to more complex objects. The restriction of Stankiewicz's stimuli to single part objects is surprising, given that one of the most important differences between structural description and view-based models of object recognition is that only the former proposes that parts and the spatial relations between parts are explicitly represented. In addition, Stankiewicz tested just one class of shapes (curved cylinders) and one shape dimension (cylinder thickness; although a second shape dimension, cylinder curvature, was manipulated in his two other experiments). He also tested only a difficult, subordinate-level shape discrimination task.

In the present experiments, this issue was reexamined using more complex objects that self-occluded and had multiple parts. Twenty classes of shapes were tested that were very different to each other, and both difficult (Experiment 1) and easy (Experiment 2) shape discriminations were tested. These experiments extend our previous work investigating the effects of view changes on the detection of shape changes (as in Lawson, 2004b, and Lawson et al., 2003, as well as in Experiment 1 here) to two new tasks: object categorization (Experiment 2) and the detection of view changes (Experiment 3). The experiments used a picture-picture matching task and presented views of morphs of novel, complex, 3-D objects, similar to such objects as dogs, chairs, and bicycles that we categorize in the real world. Incrementally morphing these novel objects in 13 stages (from an S1 to an S13 morph for each object) allowed the shape similarity of the morphs to be manipulated systematically.

EXPERIMENT 1

Experiment 1 tested four shape-change and three view-change conditions. These same conditions were then also used in Experiment 3 to permit a direct comparison of the results for the two tasks—detecting shape changes and detecting view changes—tested in Experiments 1 and 3, respectively. From the results of our previous studies (Lawson, 2004b; Lawson et al., 2003), in Experiment 1, we predicted that view changes would influence people's

ability to detect shape changes even though variation in the view in depth was irrelevant to the task (cf. Stankiewicz, 2002). We further predicted that greater view sensitivity would be found on shape-same matches than on shape-change mismatches, supporting our hypothesis that view sensitivity increases as the difficulty of shape discrimination increases.

Method

Participants. Eighteen undergraduate students from the University of Liverpool took part in the experiment for course credit.

Materials. Twenty textured, shaded, 3-D models of familiar objects were individually morphed in varied and arbitrary ways to generate a set of 20 endpoint S1 morphs. These were the 18 objects presented in Experiments 3 and 4 of Lawson et al. (2003), plus Objects 5 and 12 from that article. Views of two of these objects are shown in Figure 1, and a third is depicted in Figure 2. For a given object, its S1 morph was changed incrementally to produce a series of 12 more morphs labeled S2 to S13, where S13 had the least similar shape to S1. Pictures of 12 different depth-rotated views of each morph were produced. The 0° view was assigned to be a foreshortened view if the object had an elongated shape and was otherwise arbitrarily assigned. Foreshortened views can be particularly difficult to identify (Lawson, 1999; Lawson & Humphreys, 1999), so we avoided presenting them. From the 0° view, each successive view was rotated by 30° about the vertical axis running through the midpoint of the

morph. These views were labeled as 30°, 60°, 90°, and so on up to 330°. Figure 1 in Lawson et al. (2003) shows the 30° view of the S1 and S13 morphs of all the objects used here. Altogether, there were 156 pictures (13 morphs, each shown from 12 views) of each of the 20 objects. Grayscale images of the views of the morphs were generated using the SoftImage rendering software with the source light at the camera. These images were presented against a black background inside a window measuring 45 × 45 mm on the computer screen and subtended a visual angle of approximately 5°. In Experiment 1, each object was represented by four morphs: S1, S3, S7, and S13. Each morph was, in turn, depicted from three views in depth: at 30°, 60°, and 240°. Thus, in total, 12 pictures of each object (4 morph shapes × 3 views) were presented (see Figure 2 and Table 1).

Design. All the participants completed one block of 360 experimental trials. In this block, each of the 20 objects were presented on nine matches and nine mismatches. On all the trials, the first picture presented was the 30° view of S1. The second picture was always S1 on matches, and it was S3, S7, or S13 on mismatches. Relative to the first picture, the second picture could show a given object from the same 30° view or from a 60° view or a 240° view. For each of the 20 objects, in each of these three view-change conditions (i.e., for 0°, 30°, and 150° view changes), there were three S0 matches (for which no shape change occurred, since the second, as well as the first, picture depicted S1) and one each of S2, S6, and S12 shape-change mismatches (for which the second picture depicted S3, S7, and S13, respectively). For generality, two different (30° and 150°) view changes were tested. Note, though, that there is no simple rela-

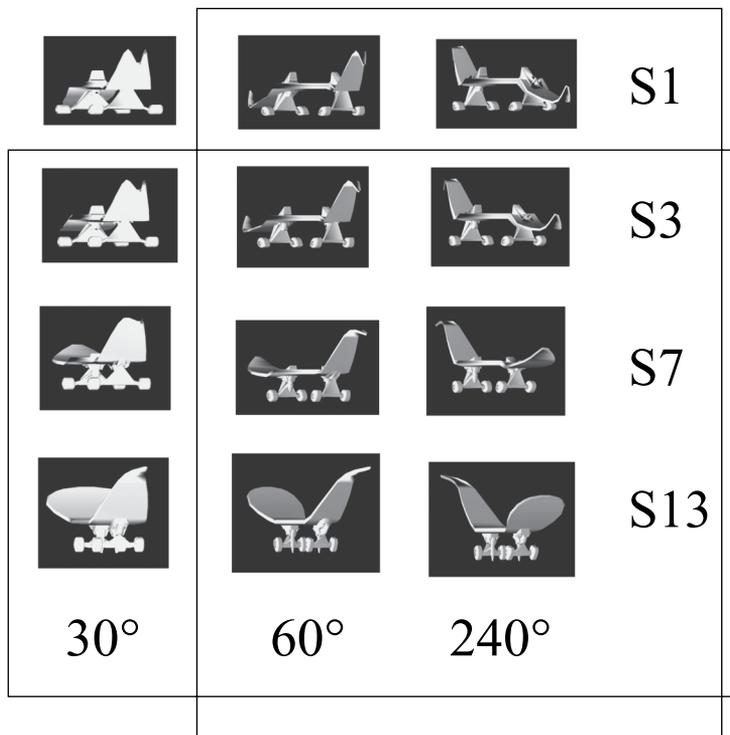


Figure 2. The 12 pictures of Object 15, which was one of the 20 objects used in the present experiments (see Figure 1 in Lawson et al., 2003, for views of S1 and S13 of all the other objects). The left, middle, and right columns show 30°, 60°, and 240° views of the object, respectively, and the first, second, third, and fourth rows show the S1, S3, S7, and S13 morphs, respectively. The leftmost box surrounds the nine stimuli shown as the second picture in mismatches in Experiment 1. The uppermost box surrounds the eight stimuli shown as the second picture in mismatches in Experiment 3. The same stimuli were also presented in Experiment 2, except that S4, rather than S3, morphs were used.

Table 1
Twelve Conditions Tested in Experiments 1 and 3 (in the 3×4 Central, Shaded Matrix), the Correct Responses for the Shape-Change Detection Task Used in Experiment 1 (Right Column), and the Correct Responses for the View-Change Detection Task Used in Experiment 3 (Bottom Row)

	0° same view (VC)	30° view change (VC)	150° view change (VC)	
S0 same shape (SC)	S0 SC 0° VC	S0 SC 30° VC	S0 SC 150° VC	Shape-same match trials (E1)
S2 shape change (SC)	S2 SC 0° VC	S2 SC 30° VC	S2 SC 150° VC	Shape- change mismatch trials (E1)
S6 shape change (SC)	S6 SC 0° VC	S6 SC 30° VC	S6 SC 150° VC	
S12 shape change (SC)	S12 SC 0° VC	S12 SC 30° VC	S12 SC 150° VC	
	View-same match trials (E3)	View-change mismatch trials (E3)		

Note—These 12 conditions map onto the 12 stimuli shown in Figure 2, so that for each condition in the central, shaded 3×4 grid, the stimulus shown in that position in Figure 2 was the second picture shown on the trials in that condition. The first picture shown was always the top left stimulus shown in Figure 2—that is, S1 depicted from a 30° view.

tion between the size of a view change and its perceived effect on the stimulus. Certain view changes, such as those producing foreshortening, make identification much harder. Often, though, there is little effect of increasing the view change above around 30°–45° (Foster & Gilson, 2002; Lawson & Humphreys, 1996). We therefore did not necessarily expect a difference between the two view changes tested here.

Apparatus and Procedure. The experiment was run on a Macintosh PowerPC G4 computer using the PsyScope Version 1.2.5 experimental presentation software. On each trial, a cue saying “Get ready for the next trial . . .” appeared for 750 msec; then, after 500 msec, the first picture was presented 50 pixels above and 50 pixels to the right of fixation for 500 msec. After a blank interstimulus interval of 400 msec, the second picture was presented at fixation until the participant responded. On all trials, there was thus a translation of the stimulus from the first to the second picture so performance could not be based on detecting low-level visual changes. The participants decided whether the two successive pictures showed the same or different-shaped objects and responded with a speeded “m” or “z” keypress, respectively. After they had responded, the correct response was given as feedback for 500 msec, by presenting the letter m or z at fixation. There was an intertrial interval of 750 msec. The participants took a self-timed break after every 120 trials. They were told to ignore any difference in the view depicted in the first and second pictures, and they were warned that on mismatches, the pictures might show two objects with very similar shapes. The participants were encouraged to respond as quickly and accurately as possible. Prior to starting the experimental block, they completed a block of 20 practice trials. These practice trials were selected at

random from the trials used in the experimental block. In the practice and experimental blocks, the trials were presented in a different, random order for each participant.

Results

ANOVAs were conducted on the mean correct reaction times (RTs) and on the percentages of errors for matches and mismatches separately (see Figure 3). Here and in Experiments 2 and 3, the results for the F values in the by-participants and by-items analyses are reported using subscripts: F_p and F_i , respectively. On matches, shape-same “m” responses were correct, whereas on mismatches, shape-change “z” responses were correct. Response latencies less than 300 msec or exceeding 2,300 msec were discarded as errors (fewer than 2% of the trials). No participants were replaced. There was one empty cell in the by-participants ANOVAs and five empty cells in the by-items ANOVAs. These were replaced by the mean for that condition. All ANOVAs included the within-participants factor of view change (0°, 30°, or 150°). Mismatch ANOVAs included a second within-participants factor of shape change (S2, S6, or S12). Unless specified otherwise, all the differences noted were significant ($p < .05$) in both by-participants and by-items post hoc Newman–Keuls analyses.

Shape-same matches. View change was significant for both RTs [$F_p(2,34) = 61.372, p < .001; F_i(2,38) = 114.63, p < .001$] and errors [$F_p(2,34) = 98.960, p < .001; F_i(2,38) = 82.085, p < .001$]. On view-same trials (676 msec, 4% errors), responses were much faster and more accurate than those for 30° view changes (902 msec, 36% errors), which in turn were somewhat faster and more accurate than those for 150° view changes (921 msec, 44% errors; for RTs, not significant for participants in Newman–Keuls analyses).

Shape-change mismatches. View change was not significant for RTs [$F_p(2,34) = 1.417, p > .2; F_i(2,38) = 0.161, p > .8$], but it was for errors [$F_p(2,34) = 8.600, p < .001; F_i(2,38) = 4.751, p < .02$]. In contrast to shape-same matches, responses on view-same trials (854 msec, 43% errors) were less accurate than those for either 30° view changes (875 msec, 34%) or 150° view changes (878 msec, 38%; not significant for items in Newman–Keuls analyses).

Shape change was significant for both RTs [$F_p(2,34) = 3.540, p < .05; F_i(2,38) = 4.186, p < .03$] and errors [$F_p(2,34) = 373.588, p < .001; F_i(2,38) = 128.647, p < .001$]. Responses for the largest S12 shape changes (833 msec, 13% errors) were faster and more accurate than those for the S6 shape changes (886 msec, 34%; for RTs, not significant for participants in Newman–Keuls analyses), which in turn were more accurate, although no faster, than those for the smallest S2 shape changes (889 msec, 68%).

The view change × shape change interaction was not significant for RTs [$F_p(4,68) = 0.427, p > .7; F_i(4,76) = 1.033, p > .3$], but it was for errors [$F_p(4,68) = 17.227, p < .001; F_i(4,76) = 13.204, p < .001$]. View change had no effect on S12 or S6 shape changes. In contrast, for the smallest S2 shape changes, responses on view-same trials

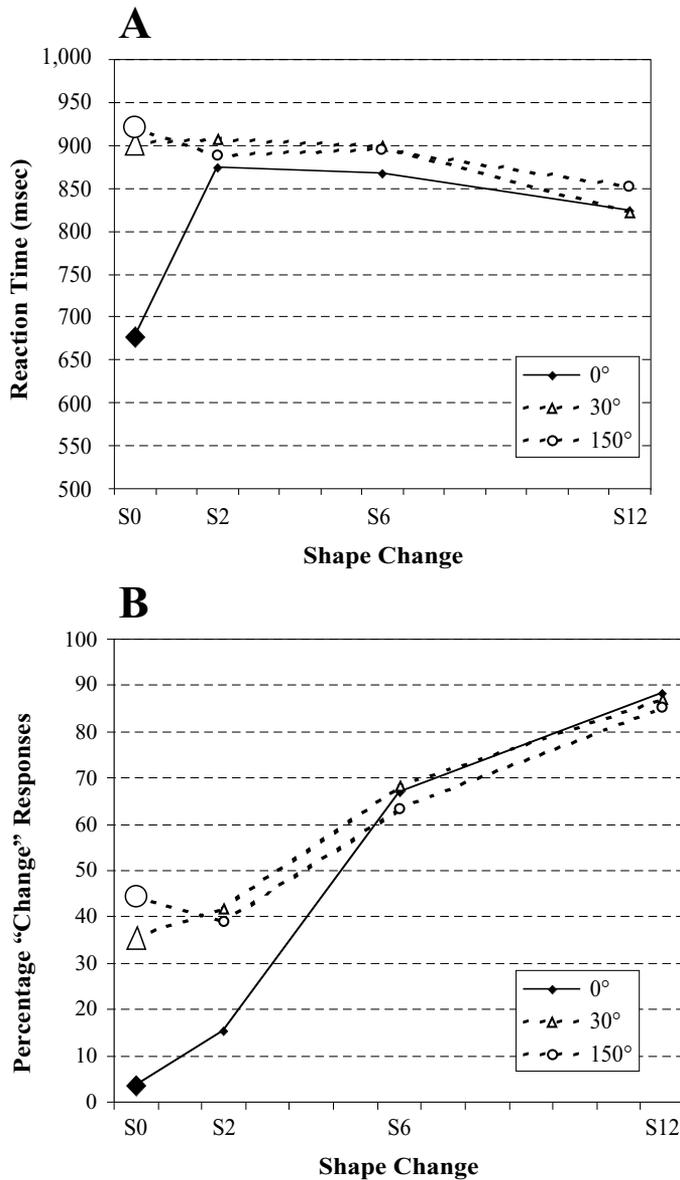


Figure 3. (A) Mean correct reaction times and (B) mean percentage "shape-change" responses in Experiment 1 in a sequential picture-picture matching task for view-same trials (0° view changes) and view-change trials (30° or 150° view changes). The first picture (a 30° view of S1 of a given object) was followed by a second picture of the same object showing either S1 on matches (where "shape change" was the wrong response) or S3, S7, or S13 on mismatches (where "shape change" was the correct response). The larger symbols indicate matches, for which three times more trials contributed to each mean in comparison with mismatches.

(85%) were much less accurate than those for either 30° view changes (58%) or 150° view changes (61%).

Discussion

In Experiment 1, shape-change detection was highly view sensitive on S0 matches, which showed two morphs with the same shape. Responses for view-same matches were much faster and more accurate than those for 30° and 150° view-change matches. In contrast, there was only

weak view sensitivity on S2 mismatches, which showed similarly shaped morphs, and there was view invariance on S6 and S12 mismatches, which showed dissimilar morphs. The latter view insensitivity was unlikely to have been due to ceiling effects, since errors were still high on S6 mismatches (34%).

This variation in the degree of view sensitivity (strong on S0 matches; weak on S2 mismatches; nothing on S6 and S12 mismatches) replicates that which we have re-

ported in similar picture–picture matching studies (Lawson, 2004b; Lawson et al., 2003). The results confirm our prediction that view sensitivity increases as shape discrimination becomes harder (Lawson, 2004b; Lawson et al., 2003; see also Biederman & Gerhardstein, 1993; Hamm & McMullen, 1998; Tarr & Cheng, 2003; but see Hayward & Williams, 2000).

This view sensitivity cannot be explained as being a consequence of viewpoint information being informative (cf. Stankiewicz, 2002). In Experiment 1, there were equal numbers of matches and mismatches for each of the three views tested, yet view sensitivity was observed only on S0 and S2 trials, not on S6 or S12 trials. Also, since only novel, unfamiliar objects were presented, no views of the objects were familiar on the basis of prior experience (cf. Palmer et al., 1981). Neither was there a training phase during which the observers could learn to associate a particular view with a given response (cf. Tarr & Pinker, 1989). Even if the observers had learned to produce view-sensitive responses (e.g., because the first picture on a trial always depicted the same 30° view, so that this view was more familiar than other views), this view sensitivity should have been equal in magnitude (but in the opposite direction) for all mismatches, in comparison with matches. This pattern of performance was not observed. Instead, these results support our claim that our visual system cannot extract shape information about an object independently of view information. As a consequence, changes in viewpoint in depth across the first and second pictures in a matching trial interfere with the detection of shape changes.

EXPERIMENT 2

In Experiment 2, we investigated whether the view sensitivity observed in the object recognition task tested in Experiment 1 would generalize to the more ecologically important task of object categorization (Palmeri & Gauthier, 2004; Poggio & Bizzi, 2004). Picture–picture matching studies usually use an object recognition task in which participants must decide whether two pictures depict either an object whose shape is exactly the same or two different-shaped objects (e.g., Lawson, 2004b; Lawson et al., 2003; as well as Experiment 1 here). Shape similarity may vary between trials for the two objects shown on mismatches (from low shape similarity, such as cat/vase, to medium, such as cat/dog, to high, such as my cat/your cat; see Lawson, 2004b), but typically, it does not vary for the two objects depicted on matches (since both pictures show the same-shaped object).

In contrast, in categorization tasks, an object is identified as a member of a class of different exemplars that may clearly differ in shape but that, nevertheless, are all grouped together. For instance, different exemplars of the category of cat will vary widely in size and shape. Tasks such as object naming and word–picture verification are usually assumed to test categorization because they require participants to use verbal labels that specify categories containing exemplars with different shapes (such as *cat*,

vase) even if, within a particular study, the participants do not have to identify objects with different shapes as belonging to the same category. Note that labeling tasks do not necessarily require categorization; naming the Eiffel Tower requires only recognition, since the label identifies a unique building with a fixed shape.

The ability to categorize is more important than the ability to recognize in most everyday circumstances (Palmeri & Gauthier, 2004; Poggio & Bizzi, 2004). Therefore, if observers use a special strategy in recognition tasks that cannot be used in categorization tasks since the strategy does not permit shape generalization, the results of Experiment 1 and similar recognition studies would be of limited interest. Note that if category members are sufficiently similar in shape, there will be effectively no difference between categorization and recognition tasks, since observers will be unable to detect the shape changes between exemplars that occur only in the categorization task, not the recognition task. For example, although every wine bottle has a slightly different shape, we often cannot perceive these subtle changes, so no difference would be predicted between performance on a wine bottle categorization task and a wine bottle recognition task. The more interesting case is a categorization task for which different members of a category have readily discriminable shapes. This situation was examined in Experiment 2.

In Experiment 2, we investigated whether the view-same advantage on shape-same matches observed in object recognition tasks (as in Experiment 1; see also Lawson, 2004b; Lawson et al., 2003) would also be found for shape-same and shape-similar matches in a task testing object categorization (see Figure 1). In this study, some matches showed two different-shaped objects, whereas other matches showed the same-shaped object in both pictures. Only the latter, shape-same matches were tested in Experiment 1 and in our earlier studies. On the basis of the results of these studies, we predicted that view-same matches would be easier than view-change matches in this categorization task, for both shape-same matches and shape-changed matches.

A second motivation for Experiment 2 was to investigate whether view sensitivity on matches would still be found in a relatively easy shape-change detection task. Overall performance was poor in Experiment 1. The high error rates indicated that the level of shape discrimination required was greater than that necessary in most everyday viewing situations. In Experiment 1, all the participants had to detect subtle S2 shape changes, and even the easiest S12 mismatches presented two morphed versions of the same-shaped object. This was addressed in Experiment 2 by presenting two completely different-shaped objects on mismatches (see Figure 1E), so that the task required only coarse shape discrimination and was much easier. Lawson (2004b) reported reduced view sensitivity when shape-change detection was easier, although importantly, view sensitivity was found in even the easiest shape discrimination task. In contrast, Hayward and Williams (2000) found no such decrease in view sensitivity when comparing the

matching of more versus less discriminable shapes. Both Lawson (2004b) and Hayward and Williams found that overall performance improved when shape-change detection was easier.

In Experiment 2, four groups of participants categorized pairs of sequentially presented pictures of novel objects as both depicting objects from the same category (matches) or from two different categories (mismatches). All four groups saw identical mismatches that depicted two completely different objects. Each group was shown a different combination of S0, S3, S6, and S12 matches. These matches varied depending on the size of the shape change depicted from the first to the second picture of an object. On S0 matches, both pictures depicted an object that was shaped exactly the same, as in the matches in Experiment 1. On S3, S6, and S12 matches, the two objects depicted differed in shape by 3, 6, and 12 levels of morphing, respectively.

Each group was given two types of matches: The S0–S3 group had half S0 and half S3 matches; the S0–S6 group had half S0 and half S6 matches; the S0–S12 group had half S0 and half S12 matches; finally, the S6–S12 group had half S6 and half S12 matches. The four groups, therefore, differed with respect to the difficulty of the categorization required. The S0–S3 group had to categorize together shapes that either were identical or were so similar that they were difficult to discriminate (Lawson et al., 2003). Thus, generalization over shape changes either was not required (S0 matches) or was trivial to achieve (S3 matches). The S0–S3 group, therefore, served as a control for the other three groups for which significant shape generalization was required. The S0–S6 and S0–S12 groups had to categorize together shapes that were moderately and severely distorted versions of each other on S6 and S12 matches, respectively. Finally, the S6–S12 group never saw the same two shapes on a trial (i.e., there were no S0 matches); on every match trial, they had to categorize together either moderately (S6) or highly (S12) dissimilar pairs of shapes. In Experiment 2, unlike in Experiment 1, some shape changes had to be ignored (S3, S6, and S12 changes), whereas other shape changes were important, since they indicated mismatches (different-category changes). In Experiment 1, all shape changes indicated a mismatch trial. In both Experiments 1 and 2, view changes were always task irrelevant.

If the view sensitivity observed in Experiment 1 (see also Lawson, 2004b; Lawson et al., 2003) was an artifact caused by using either a particularly difficult shape-change detection task or an object recognition task that did not require generalization over shape changes, this view sensitivity should be eliminated in Experiment 2. Instead, we expected that the results of the object categorization task tested in Experiment 2 would replicate and extend the results of the recognition task tested in Experiment 1. We predicted that all four groups would find view-same matches easier than view-change matches for both shape-same and shape-change matches. Note that reduced levels

of view sensitivity might occur on S12 and S6 trials, relative to S0 trials, since it generally becomes less meaningful to specify whether two objects are depicted from the same view as the objects become less similar in shape. For instance, consider trying to decide which views should be considered the same for a curled-up cat and a leaping cat. We will return to this issue in Experiment 3.

Method

Participants. Forty-eight undergraduate students from the University of Liverpool took part in the experiment for course credit.

Materials. The materials were identical to those used in Experiment 1, except that no pictures of S3 morphs were shown and pictures of S4 morphs were presented instead.

Design. All the participants completed one block of 240 experimental trials. On all trials, the first picture shown was the 30° view of S1 of a given object. On matches, the second picture depicted S1, S4, S7, or S13 of that object. On mismatches, the second picture showed S1 of a different object. Stimuli from each of the 20 objects were shown as the first picture on six matches and six mismatches. For each of these six mismatches, a different object was depicted in the second picture. One of three view conditions was presented for two of each of these six mismatches: Relative to the first picture, the second picture was depth rotated by 0°, 30°, or 150°.

Twelve participants were assigned to each of four groups, with only the matches differing across the groups. For the six matches for a given object, for the S0–S3 group, three matches showed S1 and three showed S4, in each case with one trial in each of the three view-change conditions (0°, 30°, and 150°); for the S0–S6 group, again three matches showed S1, but three showed S7, in each case with one trial in each of the three view-change conditions; for the S0–S12 group, once again three matches showed S1, but three showed S13, in each case with one trial in each of the three view-change conditions; finally, for the S6–S12 group, three matches showed S7, and three showed S13, in each case with one trial in each of the three view-change conditions. Note that the S6 and S12 matches here showed stimuli identical to those in the S6 and S12 mismatches in Experiment 1 but that the correct response was “change” in Experiment 1 but “same” in Experiment 2.

Apparatus and Procedure. The apparatus and procedure were identical to those in Experiment 1.

Results

ANOVAs were conducted on the mean correct RTs on matches and mismatches and on the percentages of errors separately (see Figure 4). On matches, category-same “m” responses were correct, whereas on mismatches, category-change “z” responses were correct. Response latencies less than 300 msec or exceeding 2,000 msec were discarded as errors (fewer than 1% of the trials). Three participants were replaced, 2 because their overall errors exceeded 20% and 1 because his mean RT exceeded 1,000 msec. There were no empty cells in the ANOVAs. All the ANOVAs included the within-participants factor of view change (0°, 30°, or 150°) and the between-participants factor of group (S0–S3, S0–S6, S0–S12, or S6–S12). Match ANOVAs also included a second within-participants factor of shape change (small or large shape changes; specifically, these were S0 and S3, S0 and S6, S0 and S12, and S6 and S12 for the S0–S3, S0–S6, S0–S12, and S6–S12 groups, respectively). Unless noted, all dif-

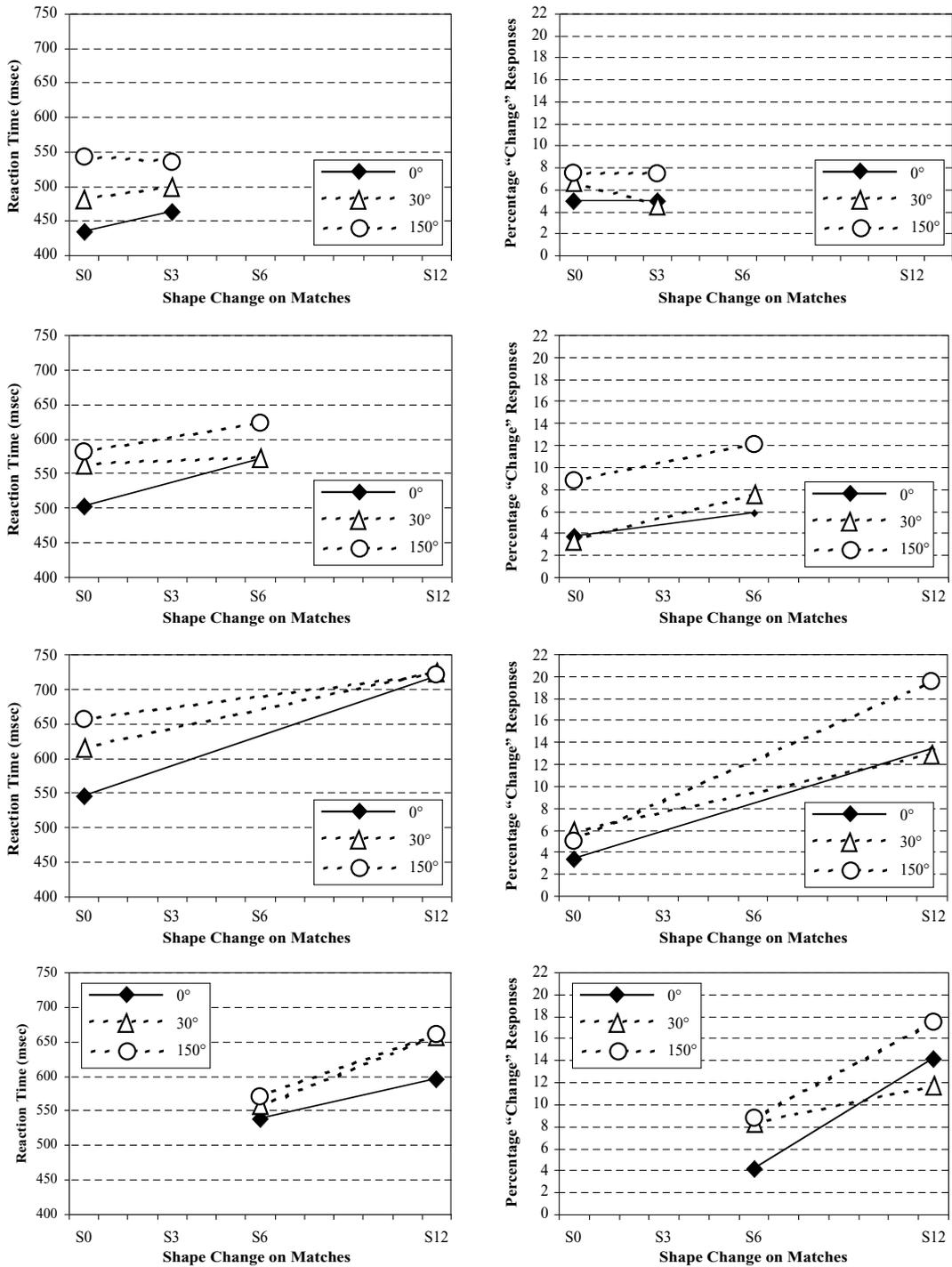


Figure 4. Mean correct reaction times (graphs in left column) and mean percentages of “category-change” responses (right column) in Experiment 2 in a sequential picture–picture matching task for view-same trials (0° view changes) and view-change trials (30° or 150° view changes) for the S0–S3 group (graphs in top row), the S0–S6 group (second row), the S0–S12 group (third row), and the S6–S12 group (bottom row). Each of these four groups saw different sets of match trials, consisting of S0 and S3 matches, S0 and S6 matches, S0 and S12 matches, and S6 and S12 matches, respectively. To achieve this, the first picture (a 30° view of S1 of a given object) was followed by a second picture showing S1, S4, S7, or S13 of the same object on matches (where “category change” was the wrong response) or by a second picture of S1 of a different object on mismatches (where “category change” was the correct response). Only performance for matches is plotted.

ferences described below were significant ($p < .05$) in by-participants and by-items post hoc Newman–Keuls analyses.

Category-same matches. In the initial analyses, the three-way view change \times shape change \times group interaction was significant for RTs [$F_p(6,88) = 2.567, p < .03$; $F_i(6,114) = 2.907, p < .02$], although not for errors [$F_p(6,88) = 1.016, p > .4$; $F_i(6,114) = 0.774, p > .5$]. Separate analyses for each of the four groups were then conducted, with view change and shape change as factors.

For the S0–S3 group, view change was significant for RTs [$F_p(2,22) = 28.430, p < .001$; $F_i(2,38) = 20.397, p < .001$], but not for errors [$F_p(2,22) = 1.009, p > .3$; $F_i(2,38) = 1.275, p > .3$]. Responses on view-same trials (449 msec) were faster than those for 30° view changes (492 msec), which in turn were faster than those for 150° view changes (539 msec). Shape change was not significant either for RTs [$F_p(1,11) = 1.091, p > .3$; $F_i(1,19) = 1.305, p > .2$] or for errors [$F_p(1,11) = 0.193, p > .6$; $F_i(1,19) = 0.271, p > .6$], and similarly the view change \times shape change interaction was not significant either for RTs [$F_p(2,22) = 0.934, p > .4$; $F_i(2,38) = 1.328, p > .2$] or for errors [$F_p(2,22) = 0.370, p > .6$; $F_i(2,38) = 0.314, p > .7$].

First, the lack of influence of shape change on performance for this group confirms our expectation that the S0 and S3 trials were so visually similar that shape generalization was trivial to achieve. The S0–S3 group performed as if the task was object recognition (i.e., as if all matches were S0 shape-same trials). Second, view-same matches were easier than view-change matches, replicating the results in Experiment 1. This group acted as a control to the other three groups for which shape generalization was harder, since clearly detectable (S6 and S12) shape changes had to be ignored on matches.

For the S0–S6 group, view change was significant for both RTs [$F_p(2,22) = 9.497, p < .002$; $F_i(2,38) = 14.725, p < .001$] and errors [$F_p(2,22) = 8.824, p < .002$; $F_i(2,38) = 5.718, p < .007$]. Responses on view-same trials (537 msec, 5% errors) were faster, although not more accurate, than those for 30° view changes (568 msec, 5%) which in turn were both faster and more accurate than those for 150° view changes (603 msec, 10%). Shape change was significant for RTs [$F_p(1,11) = 24.175, p < .001$; $F_i(1,19) = 21.073, p < .001$] and was marginally significant for errors [$F_p(1,11) = 4.805, p < .06$; $F_i(1,19) = 3.525, p < .08$]. Responses for S0 shape-same matches (549 msec, 5% errors) were faster than those for S6 shape changes (589 msec, 8%), with the same trend for errors. Finally, the view change \times shape change interaction was significant for RTs [$F_p(2,22) = 5.025, p < .02$; $F_i(2,38) = 4.616, p < .02$], but not for errors [$F_p(2,22) = 0.209, p > .8$; $F_i(2,38) = 0.261, p > .7$]. For S0 shape-same matches, responses on view-same trials (502 msec) were faster than those for 30° view changes (564 msec) and 150° view changes (582 msec). In contrast, for S6 shape changes, responses on view-same trials (571 msec) and 30° view-change trials (573 msec) were faster than those on 150° view-change trials (623 msec).

For the S0–S6 group, view-same matches were easier than view-change matches for both S0 shape-same matches and, most important, for S6 shape-change matches. This finding replicates the view sensitivity found for the S0–S3 group. Moreover, it extends this finding to object categorization (for S6 shape-change matches), in which shape generalization was nontrivial, as indicated by the significantly slower responses overall on S6 than on S0 matches.

For the S0–S12 group, view change was significant for RTs [$F_p(2,22) = 9.654, p < .002$; $F_i(2,38) = 9.081, p < .001$] and for errors for participants only [$F_p(2,22) = 3.824, p < .04$; $F_i(2,38) = 1.778, p > .1$]. Responses on view-same trials (632 msec) were faster than those for 30° view changes (670 msec) and 150° view changes (690 msec), whereas for participants only, responses on view-same trials (8%) were more accurate than those on 150° view-change trials (12%). Shape change was significant for both RTs [$F_p(1,11) = 25.387, p < .001$; $F_i(1,19) = 64.571, p < .001$] and errors [$F_p(1,11) = 18.600, p < .002$; $F_i(1,19) = 9.606, p < .006$]. Responses for S0 shape-same matches (606 msec; 5% errors) were both faster and more accurate than those for S12 shape changes (722 msec; 15%). Finally, the view change \times shape change interaction was significant for RTs [$F_p(2,22) = 7.533, p < .004$; $F_i(2,38) = 4.585, p < .02$] and for errors for participants only [$F_p(2,22) = 3.687, p < .05$; $F_i(2,38) = 1.212, p > .3$]. For S0 shape-same matches, responses on view-same trials (546 msec) were faster than those on 30° view-change trials (616 msec), which in turn were faster than those on 150° view-change trials (657 msec), but there were no significant differences in accuracy for these three conditions (3%, 6%, and 5% errors, respectively). In contrast, for S12 shape-change matches, responses on view-same trials (13% errors), and 30° view-change trials (13%) were more accurate than those on 150° view-change trials (20%) in the participants analyses only, but there were no significant differences in RTs across these three conditions (717, 725, and 723 msec, respectively).

For the S0–S12 group, view-same matches were easier than view-change matches for both S0 shape-same matches and, most important, S12 shape-change matches. As for the S0–S6 group, this finding of view sensitivity replicates the results for the S0–S3 group. In addition, it extends this result to a task for which object categorization was necessary: Generalization over shape was difficult on the S12 shape-change matches, as evidenced by the clear effects of shape changes. The view-same advantage on S12 matches was significant only in the by-participants post hoc Newman–Keuls analyses of errors. However, first, as was discussed above, weaker view sensitivity on S12 than on S0 matches was expected, and second, we will report below clear view sensitivity on S12 trials for the S6–S12 group.

For the S6–S12 group, view change was significant for RTs [$F_p(2,22) = 5.602, p < .02$; $F_i(2,38) = 5.166, p < .02$] and for errors for participants only [$F_p(2,22) = 4.063, p < .04$; $F_i(2,38) = 2.119, p > .1$]. Responses on view-same trials (567 msec) were faster than those for 30° view

changes (609 msec) and 150° view changes (616 msec), and, for participants only, responses on view-same trials (9% errors) and 30° view-change trials (10%) were more accurate than those for 150° view changes (13%). Shape change was significant for both RTs [$F_p(1,11) = 20.640, p < .001$; $F_i(1,19) = 15.992, p < .001$] and errors [$F_p(1,11) = 17.768, p < .002$; $F_i(1,19) = 21.087, p < .001$]. Responses for S6 shape changes (556 msec, 7% errors) were both faster and more accurate than those for S12 shape changes (639 msec, 14%). The view change \times shape change interaction was not significant for either RTs [$F_p(2,22) = 0.656, p > .5$; $F_i(2,38) = 0.821, p > .4$] or errors [$F_p(2,22) = 1.497, p > .2$; $F_i(2,38) = 2.262, p > .1$].

For the S6–S12 group, view-same matches were easier than view-change matches for both S6 and S12 shape-change matches. There was thus an advantage of view-same over view-change matches for the S6–S12 group (where all matches required generalization over shape, since there were no S0 matches), just as for the S0–S3 group (where little or no shape generalization was necessary to do the task). Object categorization was necessary but was difficult to achieve on every match trial for the S6–S12 group, whereas the task was no harder than an object recognition task for the S0–S3 group. Despite these differences, both groups revealed a similar pattern of view sensitivity.

Category-change mismatches. These trials were not the focus of theoretical interest here but are reported for completeness. Group was significant for both RTs [$F_p(3,44) = 2.871, p < .05$; $F_i(3,57) = 76.937, p < .001$] and errors [$F_p(3,44) = 3.988, p < .02$; $F_i(3,57) = 7.568, p < .001$]. Both the S0–S3 group (526 msec, 5%) and the S0–S6 group (618 msec, 6%) were faster and more accurate than the S6–S12 group (650 msec, 9%; not significant for RTs for participants). In addition, the S0–S3 group was both faster and more accurate than the S0–S12 group (676 msec, 8%; not significant for errors for participants). Finally, for items only, the S0–S3 group was faster than the S0–S6 group, which in turn was faster than the S0–S12 group. View change was significant for RTs [$F_p(2,88) = 16.409, p < .001$; $F_i(2,38) = 6.712, p < .004$] and for errors for participants only [$F_p(2,88) = 10.235, p < .001$; $F_i(2,38) = 1.561, p > .2$]. Responses on view-same trials (607 msec, 6% errors) and 30° view-change trials (603 msec, 7%) were faster and more accurate than those for 150° view changes (642 msec, 9%; not significant for accuracy for items). The group \times view change interaction was not significant for either RTs [$F_p(6,88) = 1.318, p > .2$; $F_i(6,114) = 1.9, p > .08$] or errors [$F_p(6,88) = 1.589, p > .1$; $F_i(6,114) = 1.090, p > .3$].

We did not expect the view manipulation to have much influence on performance on mismatches, since each mismatch showed two entirely different objects. If the two objects had shapes that could not be meaningfully aligned (e.g., as for a dog and a washbasin), no difference would be expected between the view-same and the view-change conditions. However, here, mismatches with 150° view changes tended to be harder than those with smaller or

no view changes. This effect is not surprising, since view was defined with respect to the foreshortened view for most of the objects tested here, so that 150° view-change trials tended to present two less similar stimuli than did 0° or 30° view-change trials. A similar view effect for mismatches of familiar objects was reported by Lawson and Humphreys (1996).

Discussion

The results of Experiment 2 revealed that the view sensitivity observed in Experiment 1 with an object recognition task generalized to an object categorization task. The same pattern of view sensitivity was found across all four groups in Experiment 2: Any differences in RTs or errors benefited view-same matches over view-change matches. The S0–S3 group was shown either identical or highly similarly shaped objects on matches. Although this group performed a categorization task, all the matches were very much like the matches in the recognition task tested in Experiment 1. In contrast, the other three groups had to categorize together objects with perceptually distinct shapes on at least half of the matches. These three groups showed the same pattern of view sensitivity as the S0–S3 group, with responses for view-same matches being faster and/or more accurate than those on changed-view trials. There were some differences across the four groups with respect to the exact pattern of view sensitivity observed (whether effects were observed for RTs, errors, or both and whether only 150° view-change trials or both 30° and 150° view-change trials were harder than view-same trials). However, these differences were not systematically related to the difficulty of the shape generalization required and probably resulted from noise due to the relatively small numbers of participants (12) in each group.

First, the results of Experiment 2 are consistent with our predictions based on the results of our previous object recognition studies testing picture–picture matching (Experiment 1 here; Lawson, 2004b; Lawson et al., 2003). View sensitivity was found even when the participants always had to generalize across shape changes on matches, in the S6–S12 group. Our previous results indicated that S6 and, especially, S12 shape changes were readily discriminable from S0 shape-same trials (Lawson, 2004b; Lawson et al., 2003). Therefore, the S6–S12 group could not rely on simply detecting a shape change to perform accurately, since a readily detectable shape change occurred on all the trials, matches as well as mismatches. Nevertheless, the S6–S12 group was faster and more accurate on view-same than on view-change matches. View sensitivity is thus not an artifact of including only shape-same matches, as in object recognition tasks such as that tested in Experiment 1. Instead, Experiment 2 demonstrated that the same pattern of view sensitivity can be observed in a more ecologically valid object categorization task.

Second, we succeeded in making the task much easier in Experiment 2 than in Experiment 1. There were 33% errors overall in Experiment 1, and mean RTs were around 850 msec. In contrast, there were just 8% errors overall

in Experiment 2, and mean RTs were around 600 msec. This overall improvement in performance in Experiment 2 was achieved by making shape-change detection on mismatches much easier (by having different-object mismatches, rather than S2, S6, and S12 mismatches, as in Experiment 1). This demonstrates that view sensitivity can be found reliably even when mismatches show two objects with entirely different shapes and performance is both fast and accurate (see also Hayward & Williams, 2000; Lawson, 2004b).

The results of Experiment 2 support our claim that our visual object recognition system cannot extract shape information about an object independently of view information. Furthermore, it provides evidence that interactions between the processing of shape and view are not limited to either object recognition tasks or difficult shape discrimination tasks. Instead, the results replicate those of Experiment 1, but in a shape discrimination task that was at least as easy as that of everyday, basic-level object recognition and that required generalization over large shape changes. The reason for the discrepancy between the results of Experiments 1 and 2 here, relative to those reported by Stankiewicz (2002), will be discussed below.

EXPERIMENT 3

In Experiment 3, we investigated whether participants could accurately determine the view of the objects shown in Experiments 1 and 2 regardless of whether a shape change occurred. The same stimuli were presented as in Experiment 1 in a sequential picture–picture matching task. A few studies have tested view-change detection (e.g., Tarr & Kriegman, 2001). However, except for Stankiewicz (2002), almost no research to date has compared the detection of shape changes and of view changes for pictures of the same 3-D objects. The view-change detection task was of interest for two reasons.

First, the results of Experiment 3 could be compared directly with those from Experiment 1 to further test Stankiewicz's (2002) claim that information about the viewpoint from which an object is depicted can be extracted independently of information about its shape and vice versa. This hypothesis predicted no effect of irrelevant view changes in the shape-change detection tasks used in Experiments 1 and 2. Our results did not support that prediction. Instead, in both experiments, we found clear view sensitivity on matches, with much better performance on view-same than on view-change trials. In Experiment 3, we examined the converse prediction: whether irrelevant shape changes would influence the detection of view changes, contrary to the predictions of Stankiewicz's account.

Second, Experiment 3 tested an alternative account of the finding in Experiment 1 of increased view sensitivity when more similar shapes have to be discriminated (see also Lawson et al., 2003). We argued that increased view sensitivity is found when shape-change detection is more difficult (for matches presenting the same shapes and mismatches presenting similar shapes, in compari-

son with mismatches showing dissimilar shapes). However, an alternative account is that view-invariant performance on mismatches presenting dissimilar morphs (such as S12 mismatches) results from people being unable to detect systematic differences between view-same and view-change conditions. To illustrate this, if the two objects shown on a mismatch trial were a dog and a washbasin, comparing view-same and view-change conditions would be almost meaningless. The shapes of the two objects are so dissimilar that views of the objects are difficult to align in any meaningful way, and so view invariance is predicted.

In Experiment 3, if view changes could be detected accurately only when similarly shaped objects were depicted, this would support this alternative hypothesis. If so, the view invariance on mismatches presenting dissimilar objects reported in Experiment 1 here and by Lawson et al. (2003) might not have been due to the shape similarity of the objects per se. Instead, we may have underestimated view sensitivity on such large shape-change trials. If, though, the observers in Experiment 3 could readily decide whether two objects were being shown from the same view, even if the objects had dissimilar shapes (as on S12 mismatches), this would support our original account. This proposes that the human visual object recognition system is disrupted more by view changes when it must discriminate similar objects (as on S0 matches and S3 mismatches) than when shape-change detection is easy (as on S12 mismatches). Note that any support for the former, alternative account would not change our most important finding, that the visual system has difficulty in achieving object constancy over depth rotation when similar shapes must be discriminated. However, it would suggest that our results underestimated the view sensitivity of the visual system.

Method

Participants. Twenty undergraduate students from the University of Liverpool took part in the experiment for course credit.

Materials. The materials were identical to those used in Experiment 1 (see Figure 2 and Table 1).

Design. All the participants completed one block of 320 experimental trials. In this block, each of the 20 objects were presented on eight matches and eight mismatches. On all the trials, the first picture presented was the 30° view of S1. The second picture was always another 30° view on matches, and it was a 60° view or a 240° view on mismatches. Relative to the first picture, the second picture could show the same S1 morph or a different S3, S7, or S13 morph. In each of these four shape-change conditions (i.e., for shape-same S0 trials and for S2, S6, or S12 shape-change trials), there were two view-same (0° view-change) matches, one 30° view-change mismatch, and one 150° view-change mismatch.

Apparatus and Procedure. The apparatus and procedure were identical to those in Experiment 1, except for the instructions given and the fact that participants took a self-timed break after completing every 80 trials. The participants were instructed to make a speeded decision as to whether the two successive pictures showed objects from the same view or from two different viewpoints by making an "m" or a "z" keypress, respectively. The participants were told to ignore any difference in the shapes of the first and second objects. They were warned that on mismatches, the pictures might show two objects from very similar views.

Results

ANOVAs were conducted on the mean correct RTs and on the percentages of errors for matches and mismatches separately (see Figure 5). On matches, view-same “m” responses were correct, whereas on mismatches, view-

change “z” responses were correct. Response latencies less than 300 msec or exceeding 2,300 msec were discarded as errors (fewer than 1% of the trials). No participants were replaced. There were no empty cells in the ANOVAs. All the ANOVAs included the within-participants factor of

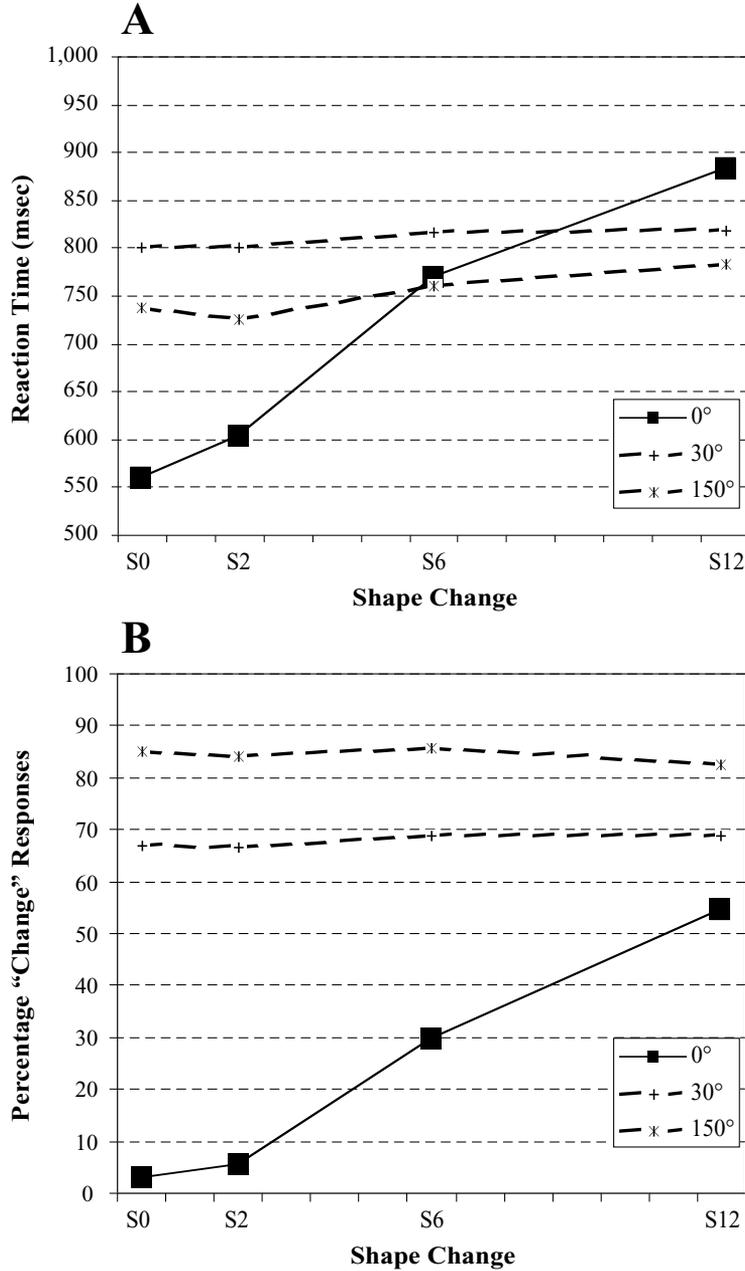


Figure 5. (A) Mean correct reaction times and (B) mean percentages of “view-change” responses in Experiment 3 in a sequential picture–picture matching task for shape-same (S0) and shape-change (S2, S6, or S12) trials. The first picture (a 30° view of S1 of a given object) was followed by a second picture of the same object showing either another 30° view on matches (where “view change” was the wrong response) or a 60° view or a 240° view on mismatches (where “view change” was the correct response). The larger symbols indicate matches, for which twice as many trials contributed to each mean, in comparison with mismatches.

shape change (S0, S2, S6, or S12). Mismatch ANOVAs also included a second within-participants factor of view change (30° or 150°). Unless specified otherwise, all differences noted were significant ($p < .05$) in both by-participants and by-items post hoc Newman–Keuls analyses.

View-same matches. Shape change was significant for both RTs [$F_p(3,57) = 93.949, p < .001; F_i(3,57) = 154.513, p < .001$] and errors [$F_p(3,57) = 221.251, p < .001; F_i(3,57) = 73.853, p < .001$]. Responses on the S0 shape-same trials (559 msec, 3% errors) were faster but no more accurate than those on S2 shape-change trials (604 msec, 5%), which in turn were both faster and more accurate than those on S6 shape-change trials (769 msec, 30% errors), which, finally, were both faster and more accurate than those on the largest, S12 shape-change trials (883 msec, 55% errors).

View-change mismatches. Shape change was significant for RTs [$F_p(3,57) = 3.666, p < .02; F_i(3,57) = 3.452, p < .03$] but not for errors [$F_p(3,57) = 0.322, p > .8; F_i(3,57) = 0.117, p > .9$]. Responses on S0 shape-same trials (770 msec, 24% errors) and S2 shape-change trials (763 msec, 25%) were faster than those on the largest, S12 shape-change trials (802 msec, 24% errors; not significant for items). There were no significant differences involving the intermediate, S6 shape changes (788 msec, 23% errors).

View change was significant for both RTs [$F_p(1,19) = 45.101, p < .001; F_i(1,19) = 22.367, p < .001$] and errors [$F_p(1,19) = 175.872, p < .001; F_i(1,19) = 46.097, p < .001$]. The 150° view changes (752 msec, 16% errors) were detected more quickly and more accurately than 30° view changes (809 msec, 32% errors). Note that Figure 5B plots percentage of “view change” responses, so there was no speed–accuracy trade-off for responses to 30° and 150° view changes. The shape change \times view change interaction was not significant for either RTs [$F_p(3,57) = 0.858, p > .4; F_i(3,57) = 1.308, p > .2$] or errors [$F_p(3,57) = 0.635, p > .5; F_i(3,57) = 0.447, p > .7$].

Discussion

The results from Experiment 3 revealed similarities but also important differences, relative to those of Experiment 1. First, in both experiments, shape changes influenced performance profoundly on view-same trials (solid lines in Figures 3 and 5) but had less effect on view-change trials (broken lines in Figures 3 and 5). As in Experiment 1, in Experiment 3 there was greater view sensitivity on shape-same trials than on shape-change trials and greater shape sensitivity on view-same trials than on view-change trials. This indicates that information about object shape cannot be extracted independently of information about the viewpoint from which an object is depicted (cf. Stankiewicz, 2002). We will return to this issue in the General Discussion section.

Second, in Experiment 3, view sensitivity was found for both RTs and errors for all levels of shape change, and there was also a difference between 30° and 150° view-change

trials at every level of shape change. In contrast, in Experiment 1, view sensitivity was found only on S0 and S2 trials, and this view sensitivity extended only to a difference between 30° and 150° view-change trials on S0 matches. This indicates that the observers were extracting different information in the two tasks. Variation in viewpoint in depth had an effect in Experiment 1 (despite this variation being irrelevant to the shape-change detection task). However, view changes had much more influence when view was the task-relevant dimension in Experiment 3. Statistical evidence for this difference is reported in the joint analysis of the results of Experiments 1 and 3 below.

Third, in Experiment 3, the observers could usually discriminate between view-same and view-change conditions, although their ability to detect view changes was poor when two dissimilarly shaped objects were presented, on S12 mismatches. On around half of the view-same S12 trials, the observers incorrectly responded that the view of the object had changed, whereas there were almost no such errors on view-same S0 trials. This finding provides partial support for the alternative hypothesis outlined above to account for the finding of view invariance on mismatches presenting dissimilar shapes in Experiment 1. Experiment 3 demonstrated that observers find it harder to decide whether two objects are shown from the same view if the objects are also dissimilar in shape (e.g., on S6 and S12 trials). This reduced (but importantly, did not eliminate) the possibility of detecting view sensitivity on such trials in a shape-change detection task such as that used in Experiment 1. Thus, view sensitivity on S6 and S12 shape-change trials may have been underestimated in Experiment 1 and other shape-change detection studies (e.g., Lawson, 2004b; Lawson et al., 2003), since view changes on these trials were harder to detect than those on S0 and S2 trials that presented shapes that were more similar.

JOINT ANALYSIS OF COMPARABLE CONDITIONS IN EXPERIMENTS 1 AND 3

Both shape-sensitive and view-sensitive performance was found in Experiments 1 and 3. These final analyses tested whether this sensitivity to shape changes and to view changes depended on the participant’s task—whether they were detecting shape changes (and ignoring view changes) in Experiment 1 or detecting view changes (and ignoring shape changes) in Experiment 3. We investigated whether people were more sensitive to variation in the task-relevant dimension (shape in Experiment 1; view in Experiment 3) than in the task-irrelevant dimension. If so, this would indicate that people were at least partially successful in selectively attending to the most task-relevant visual information. The critical conditions for this between-experiment comparison are the mismatches in which the same stimuli were presented and the same response (either “same” or “change”) was correct. The trials in the six conditions that satisfied these criteria were identical across the two experiments, with only the task differing between them (see Figure 6).

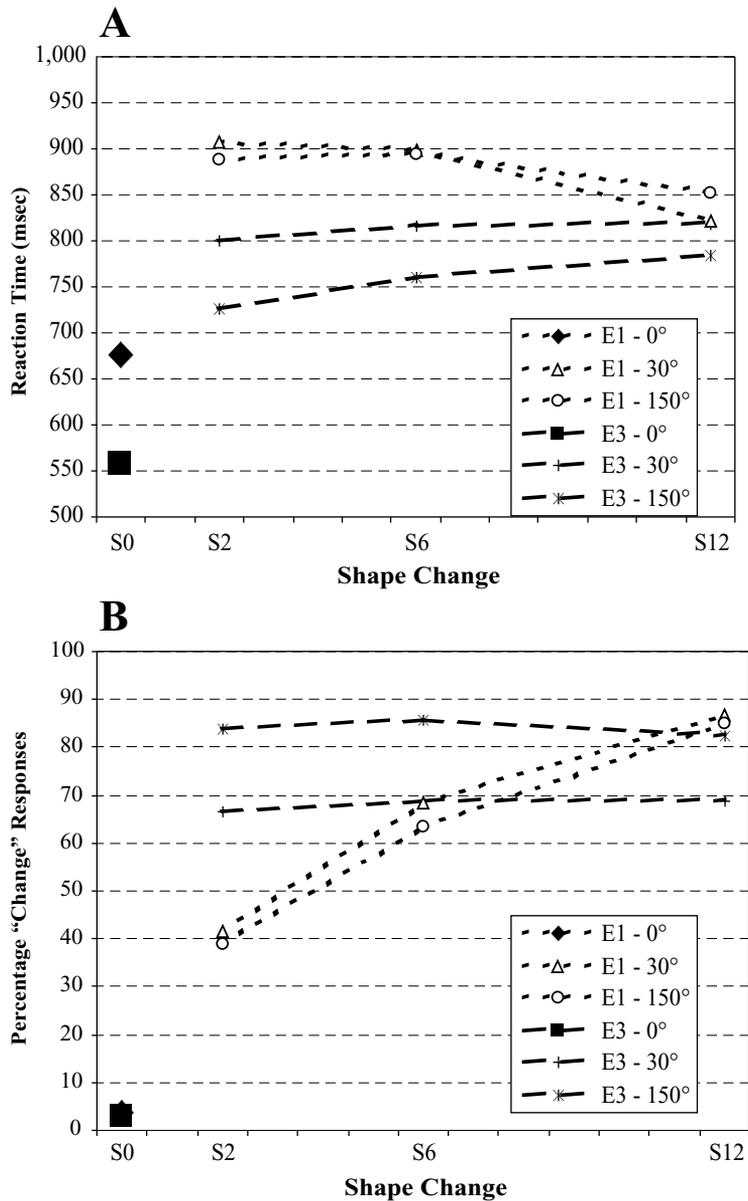


Figure 6. (A) Mean correct reaction times and (B) mean percentages of "change" responses for the shape-same, view-same match condition (with neither a view change nor a shape change) and the six shape-change, view-change mismatch conditions that were directly comparable across Experiments 1 and 3. Only the mismatches were analyzed, since there were both view changes (30° or 150°) and shape changes (S2, S6, or S12) across these trials.

View-Change and Shape-Change Mismatches

For the mismatches for which there was both a view change and a shape change (S2, S6, and S12 shape changes for 30° and 150° view changes), experiment was significant for RTs for items only [$F_p(1,36) = 2.894, p < .1$; $F_i(1,38) = 21.383, p < .001$] and was significant for errors [$F_p(1,36) = 27.515, p < .001$; $F_i(1,38) = 11.456, p < .002$]. Responses for mismatches were faster (for items only) and more accurate in the view-change detection task in Experiment 3 (784 msec, 24% errors) than in the

shape-change detection task in Experiment 1 (877 msec, 36% errors). Shape change was not significant for RTs [$F_p(2,72) = 1.383, p > .2$; $F_i(2,76) = 1.917, p > .1$] but was significant for errors [$F_p(2,72) = 89.866, p < .001$; $F_i(2,76) = 45.835, p < .001$]. Accuracy increased as the size of the shape change increased from S2 (41% errors) to S6 (28%) to S12 (19%). View change was significant for RTs for participants only [$F_p(1,36) = 9.241, p < .005$; $F_i(1,38) = 3.396, p < .08$] and was significant for errors [$F_p(1,36) = 38.244, p < .001$; $F_i(1,38) = 10.737,$

$p < .003$]. Performance improved as the size of the view change increased from 30° (842 msec, 33% errors) to 150° (814 msec, 26% errors).

The most theoretically important results of these analyses were two significant interactions (see Figure 6). First, the experiment \times shape change interaction was significant for both RTs [$F_p(2,72) = 7.046, p < .002$; $F_i(2,76) = 4.921, p < .01$] and errors [$F_p(2,72) = 85.906, p < .001$; $F_i(2,76) = 43.815, p < .001$]. There was a clear effect of shape change in Experiment 1. Responses for S2 mismatches (897 msec, 60% errors) were less accurate than those for S6 mismatches (896 msec, 34%), which in turn were both slower and less accurate than those for S12 mismatches (837 msec, 14%). In contrast, there was no effect of shape change in Experiment 3, with no significant differences between the S2 (763 msec, 25%), S6 (788 msec, 23%), and S12 (802 msec, 24%) mismatches. Thus, there was greater sensitivity to variation in shape when shape was a task-relevant dimension (in Experiment 1), rather than a task-irrelevant dimension (in Experiment 3).

Second, the experiment \times view change interaction was significant for both RTs [$F_p(1,36) = 11.295, p < .002$; $F_i(1,38) = 7.711, p < .009$] and errors [$F_p(1,36) = 86.588, p < .001$; $F_i(1,38) = 24.309, p < .001$]. In Experiment 1, there was no effect of view change, with no significant difference between 30° (875 msec, 34%) and 150° (878 msec, 38%) view changes. In contrast, there was a clear effect of view change in Experiment 3, with responses for 30° view changes (812 msec, 32% errors) being both slower and less accurate than those for 150° view changes (757 msec, 16%). There was thus greater sensitivity to variation in view in depth when view was a task-relevant dimension (in Experiment 3), rather than a task-irrelevant dimension (in Experiment 1).

Discussion of Joint Analysis of Experiments 1 and 3

The mismatch analyses revealed that variation in the task-relevant dimension harmed performance more than did variation in the task-irrelevant dimension. For the six comparable shape-change and view-change mismatch conditions, responses to large shape changes were faster and much more accurate than responses to small shape changes in Experiment 1 with a shape-change detection task. In contrast, in Experiment 3 with a view-change detection task, the participants were relatively insensitive to the size of the shape change. Conversely, in Experiment 1, the participants were insensitive to the size of the view change on mismatches. However, in Experiment 3, responses to large view changes were both faster and more accurate than those to small view changes. These differences show that the observers used different visual information, depending on their task. Shape changes and view changes both influenced performance, but sensitivity to shape changes was greater if the task was to detect shape changes (Experiment 1), whereas sensitivity to view changes was greater if the task was to detect view changes (Experiment 3). Thus, shape information can, to some extent, be distinguished from view information. This may

be because some of the view and shape information that is important in these tasks is coded in separable, independent dimensions, as Stankiewicz (2002) proposed. However, the large residual effects of task-irrelevant changes reported in all three experiments here indicate that shape information cannot be extracted independently of view information and vice versa.

GENERAL DISCUSSION

View changes influenced performance in a shape-change detection task, both for an object recognition task in Experiment 1 and for an object categorization task in Experiment 2. Conversely, shape changes influenced performance for a view-change detection task in Experiment 3. All three experiments thus showed that picture-picture matching performance was disrupted by variation in a task-irrelevant dimension, whether that was due to view changes or to shape changes. These results are inconsistent with Stankiewicz's (2002) hypothesis that information about object shape is extracted independently of information about the viewpoint from which an object is depicted and vice versa. Instead, the present results indicate that view changes do significantly reduce our ability to detect shape changes (Experiments 1 and 2), whereas shape changes harm our ability to detect view changes (Experiment 3).

Further research is needed to clarify why the results obtained here differed from those reported by Stankiewicz (2002). There were a number of potentially important differences across these studies. For example, Stankiewicz presented only simple, single-part curved cylinders. These varied only quantitatively along two shape dimensions (cylinder thickness and cylinder curvature; and only thickness was manipulated in his critical third experiment), and only two sets of views were tested for each of his observers in his third experiment. In contrast, a larger set of more complex, varied and multipart stimuli were used here, each of which was morphed in different ways, producing qualitative, as well as quantitative, shape changes. Also, Stankiewicz tested a difficult (subordinate level) categorization task and only presented single pictures of objects, whereas in the present experiments, people decided whether two sequentially presented pictures of objects matched and shape discrimination difficulty ranged from easy to hard. Finally, Stankiewicz's 2 observers in his third experiment were highly trained, whereas the larger groups of participants tested here received little training.

One possibility is that effects of task-irrelevant variation on performance in the present experiments may have resulted from uncertainty about how to perform the task and, specifically, about what information should be used (see Op de Beeck, Wagemans, & Vogels, 2003). However, Op de Beeck et al.'s own elegant experiments provide evidence against the hypothesis that further training would have reduced the effects of task-irrelevant variation in the present experiments. Pairs of shape dimensions that were integral on initial testing (in Op de Beeck et al.'s first experiment) remained integral following either 1 or 5 h of

category training in their second and third experiments, respectively. Op de Beeck et al.'s discussion of why their results differed from those of Goldstone (1994) and Goldstone and Steyvers (2001) provides further reasons that could explain the discrepancy between the present results and those of Stankiewicz (2002). Further research in this area should seek to provide converging evidence about the conditions under which view information and shape information can be accessed independently—for example, by testing alternative methods, such as the visual search task used by Op de Beeck et al. View sensitivity must also be examined across a wide range of stimuli. Studies testing simple, impoverished objects (e.g., the curved cylinders used by Stankiewicz, 2002, and the 2-D, unshaded shapes tested by Op de Beeck et al., 2003) provide important information, but studies presenting more complex objects closer to the stimuli that we must identify in everyday object recognition (such as the stimuli used in the present experiments) are also essential.

The view sensitivity observed in Experiment 1 here (see also Lawson, 2004b; Lawson et al., 2003) was not due merely to testing a recognition task. A similar view-same benefit for matches was found in Experiment 2 with an object categorization task, even for the S6–S12 group that had to generalize over either moderate or large shape changes on every match trial. In addition, the view sensitivity that we observed in Experiment 1 was not confined to challenging shape-change detection tasks for which shape discrimination was as hard as that required for subordinate-level recognition. Instead, similar view sensitivity on matches was also obtained in a much less difficult task in Experiment 2. Here, shape-change detection was easier than that required routinely in everyday object recognition when, for example, dogs must be distinguished from cats, pens from pencils, and TVs from computers.

In conclusion, the results of the three picture–picture matching experiments reported here indicate that task-irrelevant changes both to view in depth and to shape disrupt performance. However, the joint analysis of comparable mismatches in Experiments 1 and 3 demonstrated that performance was influenced more by changes to the task-relevant dimension, whether shape or view in depth. People thus seem able to partially, but not fully, ignore variation in viewpoint in a shape-change detection task and to partially, but not fully, ignore variation in shape in a view-change detection task (cf. Stankiewicz, 2002). These results are consistent with a large body of evidence using a range of tasks and stimuli that indicate that under most circumstances—including those most similar to everyday object recognition at the entry or basic level—our ability to detect shape changes is influenced by the views from which objects are shown.

REFERENCES

- BIEDERMAN, I., & GERHARDSTEIN, P. C. (1993). Recognising depth-rotated objects: Evidence for 3-D viewpoint invariance. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 1162-1182.
- BIEDERMAN, I., & GERHARDSTEIN, P. C. (1995). Viewpoint-dependent mechanisms in visual object recognition: Reply to Tarr and Bülthoff (1995). *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 1506-1514.
- BÜLTHOFF, H. H., & EDELMAN, S. (1992). Psychophysical support for a 2-D view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences*, **89**, 60-64.
- CUTZU, F., & EDELMAN, S. (1998). Representation of object similarity in human vision: Psychophysics and a computational model. *Vision Research*, **38**, 2229-2257.
- EDELMAN, S., & BÜLTHOFF, H. H. (1992). Orientation dependence in the recognition of three-dimensional objects. *Vision Research*, **32**, 2385-2400.
- FOSTER, D. H., & GILSON, S. J. (2002). Recognising novel three-dimensional objects by summing signals from parts and views. *Proceedings of the Royal Society of London: Series B*, **269**, 1939-1947.
- GOLDSTONE, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, **123**, 178-200.
- GOLDSTONE, R. L., & STEYVERS, M. (2001). The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General*, **130**, 116-139.
- HAMM, J. P., & McMULLEN, P. A. (1998). Effects of orientation on the identification of rotated objects depend on the level of identity. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 413-426.
- HAYWARD, W. G. (2003). After the viewpoint debate: Where next in object recognition? *Trends in Cognitive Sciences*, **7**, 425-427.
- HAYWARD, W. G., & WILLIAMS, P. (2000). Viewpoint dependence and object discriminability. *Psychological Science*, **11**, 7-12.
- HUMMEL, J. E., & STANKIEWICZ, B. J. (1998). Two roles for attention in shape perception: A structural description model of visual scrutiny. *Visual Cognition*, **5**, 49-79.
- HUMPHREY, G. K., & JOLICŒUR, P. (1993). An examination of the effects of axis foreshortening, monocular depth cues and visual field on object identification. *Quarterly Journal of Experimental Psychology*, **46A**, 137-159.
- JOLICŒUR, P. (1985). The time to name disoriented natural objects. *Memory & Cognition*, **13**, 289-303.
- JOLICŒUR, P. (1990). Identification of misoriented objects: A dual systems theory. *Mind & Language*, **5**, 387-410.
- JOLICŒUR, P., & HUMPHREY, G. K. (1998). Rotated objects and visual shapes. In V. Walsh & J. Kulikowski (Eds.), *Perceptual constancy: Why things look as they do* (pp. 69-123). Cambridge: Cambridge University Press.
- LAWSON, R. (1999). Achieving visual object constancy over plane rotation and depth rotation. *Acta Psychologica*, **102**, 221-245.
- LAWSON, R. (2004a). Depth rotation and mirror-image reflection reduce affective preference as well as recognition memory for pictures of novel objects. *Memory & Cognition*, **32**, 1170-1181.
- LAWSON, R. (2004b). View sensitivity increases for same-shape matches if mismatches show pairs of more similar shapes. *Psychonomic Bulletin & Review*, **11**, 896-902.
- LAWSON, R., BÜLTHOFF, H. H., & DUMBELL, S. (2003). Interactions between view changes and shape changes in picture–picture matching. *Perception*, **32**, 1465-1498.
- LAWSON, R., & HUMPHREYS, G. W. (1996). View-specificity in object processing: Evidence from picture matching. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 395-416.
- LAWSON, R., & HUMPHREYS, G. W. (1998). View-specific effects of depth rotation and foreshortening on the initial recognition and priming of familiar objects. *Perception & Psychophysics*, **60**, 1052-1066.
- LAWSON, R., & HUMPHREYS, G. W. (1999). The effects of view in depth on the identification of line drawings and silhouettes of familiar objects: Normality and pathology. *Visual Cognition*, **6**, 165-195.
- LAWSON, R., & JOLICŒUR, P. (2003). Recognition thresholds for plane-rotated pictures of familiar objects. *Acta Psychologica*, **112**, 17-41.
- MARR, D. (1982). *Vision*. New York: Freeman.
- OP DE BEECK, H., WAGEMANS, J., & VOGELS, R. (2003). The effect of category learning on the representation of shape can be biased but not

- differentiated. *Journal of Experimental Psychology: General*, **132**, 491-511.
- PALMER, S., ROSCH, E., & CHASE, P. (1981). Canonical perspective and the perception of objects. In J. [B.] Long & A. [D.] Baddeley (Eds.), *Attention and performance IX* (pp. 135-151). Hillsdale, NJ: Erlbaum.
- PALMERI, T. J., & GAUTHIER, I. (2004). Visual object understanding. *Nature Reviews Neuroscience*, **5**, 291-304.
- POGGIO, T., & BIZZI, E. (2004). Generalisation in vision and motor control. *Nature*, **431**, 768-774.
- POGGIO, T., & EDELMAN, S. (1990). A network that learns to recognise three-dimensional objects. *Nature*, **343**, 263-266.
- SEAMON, J. G., & DELGADO, M. R. (1999). Recognition memory and affective preference for depth-rotated solid objects: Part-based structural descriptions may underlie the mere exposure effect. *Visual Cognition*, **6**, 145-164.
- STANKIEWICZ, B. J. (2002). Empirical evidence for independent dimensions in the visual representation of three-dimensional shape. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 913-932.
- TARR, M. J., & BÜLTHOFF, H. H. (1995). Is human object recognition better described by geon-structural-descriptions or by multiple-views? *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 1494-1505.
- TARR, M. J., & CHENG, Y. D. (2003). Learning to see faces and objects. *Trends in Cognitive Sciences*, **17**, 23-30.
- TARR, M. J., & KRIEGMAN, D. J. (2001). What defines a view? *Vision Research*, **41**, 1981-2004.
- TARR, M. J., & PINKER, S. (1989). Mental rotation and orientation dependence in shape recognition. *Cognitive Psychology*, **21**, 233-283.

(Manuscript received June 18, 2004;
revision accepted for publication July 25, 2005.)