Object recognition is mediated by extraretinal information

DANIEL J. SIMONS
Harvard University, Cambridge, Massachusetts

RANXIAO FRANCES WANG
University of Illinois at Urbana-Champaign, Urbana, Illinois

and

DAVID RODDENBERRY
Harvard University, Cambridge, Massachusetts

Many previous studies of object recognition have found view-dependent recognition performance when view changes are produced by rotating objects relative to a stationary viewing position. However, the assumption that an object rotation is equivalent to an observer viewpoint change ignores the potential contribution of extraretinal information that accompanies observer movement. In four experiments, we investigated the role of extraretinal information on real-world object recognition. As in previous studies focusing on the recognition of spatial layouts across view changes, observers performed better in an old/new object recognition task when view changes were caused by viewer movement than when they were caused by object rotation. This difference between viewpoint and orientation changes was due not to the visual background, but to the extraretinal information available during real observer movements. Models of object recognition need to consider other information available to an observer in addition to the retinal projection in order to fully understand object recognition in the real world.

Studies of object representations often ask whether recognition performance is affected by changes to the observer's view of the object. This question is central to our understanding of object perception, in that the degree to which recognition is view dependent may determine the specificity of our representations, thereby constraining models of recognition. Evidence that recognition is slower or less accurate for rotated objects supports the idea that representations are specific to previously studied views. Evidence that recognition is unaffected by view changes is somewhat ambiguous: Either the representation is view independent, or observers can rapidly extrapolate from previously studied views to the new view. Models based on view-dependent representations typically predict that recognition latency and/or error should increase as the magnitude of the view change increases (e.g., Bülthoff, Edelman, & Tarr, 1995; Diwadkar & McNamara, 1997; Edelman & Bülthoff, 1992; Humphrey & Khan, 1992;

Logothetis & Pauls, 1995; Tarr & Pinker, 1989; Tarr, Williams, Hayward, & Gauthier, 1998; Ullman, 1989). Models based on view-independent representations (e.g., structural descriptions) predict no effect of view changes on recognition, provided that the basic structure of the object can be recovered from the new view (e.g., Biederman & Cooper, 1991; Biederman & Gerhardstein, 1993; Cooper, Biederman, & Hummel, 1992).

Both classes of models are based only on the information provided by the retinal projection of the object, either implicitly or explicitly ignoring extraretinal influences on recognition performance. Hence, to our knowledge, almost all previous studies of individual object recognition have examined the effect of view changes by rotating an object in front of a stationary observer (a few exceptions are noted below). Furthermore, in almost all cases, both the object and the view change were simulated on a computer display. Such object orientation changes are assumed to be functionally equivalent to a change in the observer's viewing position (i.e., a viewpoint change) because the change to the retinal projection caused by an object rotation can be made equivalent to that caused by observer movement. In fact, object rotations are frequently referred to as "viewpoint" changes in the literature. This approach has dominated research on object recognition, in part because of this pervasive theoretical assumption, but also because computerized displays are easier to generate and allow parametric variation of the stimuli. In

contrast, real viewpoint changes require a mobile observer, making studies with real viewpoint changes pragmatically more difficult to conduct.

One series of studies has addressed the relationship between observer orientation and object recognition (e.g., Rock, 1973; Rock & DiVita, 1987). These studies used real objects (often novel wire frame objects) to study the effect of rotations in the picture plane. Such rotations typically diminish recognition performance (Biederman & Gerhardstein, 1993; Rock & DiVita, 1987), much as rotations in depth can impair recognition performance. Importantly, Rock (1973) contrasted the environmental orientation of the object with the retinal projection of that object by tilting the observer. When observers were tilted, object recognition performance was still based on the environmental orientation of the object, suggesting that the retinal orientation was not as critical. In other words, an observer rotation was not detrimental in the same way that a display rotation was (Rock, 1973), suggesting some degree of invariance to observer rotation. However, given that models proposing view-independent representations focus almost exclusively on invariance to rotations in depth rather than to rotations in the picture plane (see Biederman & Gerhardstein, 1993, for motivations underlying this practice), these studies do not directly address the degree to which observer position in space affects object recognition.

Although studies of individual object representations have not considered the impact of observer viewpoint changes, the literatures on navigation and perspective-taking strongly suggest that observer movement plays an important role in maintaining and updating spatial representations. Humans and other species can update spatial relationships across changes in position. For example, insects and rodents can return to their nests or feeding locations in a direct path by integrating their body movements over time (e.g., Gallistel, 1990; Wehner & Srinivasan, 1981). Human children and adults can point to a target after walking without vision, suggesting they can update the position of the target relative to themselves using non-visual information (e.g., Landau, Spelke, & Gleitman, 1984; Loomis et al., 1993; Rieser & Rider, 1991). In fact, the receptive fields of some neurons in the parietal cortex move according to intended eye movements, suggesting that the parietal cortex anticipates the consequences of the eye movements on the retinal image and updates the retinal coordinates of remembered stimuli using extraretinal information (Duhamel, Colby, & Goldberg, 1992). In studies of spatial representations in imagined environments, observers are better able to point to targets following actual movement than following imagined view changes (e.g., Farrell & Robertson, 1998; Rieser, 1989). In other words, visual information alone is less useful than visual information in conjunction with actual observer movements. All of these studies suggest that extraretinal information might play an important role in maintaining a representation of spatial position over time.

On the basis of these studies of spatial updating, we predicted that recognition of spatial layouts of objects might similarly benefit from actual observer movement. Performance should be less disrupted if the view change is caused by the movements of an observer relative to a stationary display than if it is caused by the rotation of a display relative to a stationary observer. Over the past few years, we have conducted a series of studies investigating the ability to recognize layouts or arrays of objects across changes in view (Simons & Wang, 1998; Wang & Simons, 1999). We compared change detection performance for layouts of objects following two different types of view change: display rotations (orientation changes), and observer movements (viewpoint changes), while equating for the magnitude of the view change. Observers viewed an array of five real objects on a circular table. After a brief delay during which the table was hidden from view and an object was moved, they viewed the objects again and tried to determine which of the five objects had moved relative to the other objects. On some trials, the table rotated relative to their observation point, causing a display orientation change of about 40°. On other trials, the observer moved and the display remained in its original orientation, causing a viewpoint change of the same size. Across a number of experiments, we found that the ability to identify the moved object in the array was relatively unaffected by the shift in viewpoint. In striking contrast, performance was significantly worse after an orientation change than when observers received the same view of the array. Even though the size of the view change was equated for the viewpoint and orientation conditions, change detection performance was view dependent in the orientation condition and relatively view independent in the viewpoint condition. This basic pattern of results held even when the retinal projections in the two conditions were equated by darkening the room and painting the objects with phosphorescent paint (Simons & Wang, 1998). Because the retinal images were matched precisely in that experiment, the difference in performance for orientation and viewpoint changes must be attributed to the extraretinal information available during observer movement.

Nevertheless, it is not clear whether this distinction between viewpoint and orientation also applies to single object recognition. The difference between viewpoint and orientation changes might be specific to spatial layouts and/or to the specific task used, and not to something fundamental about object recognition. The task used in our previous studies relied on a change detection procedure rather than an old/new recognition task. Observers viewed a layout of objects and were asked to determine what had changed about the layout. Although object recognition studies typically ask observers to discriminate old objects from changed or even completely different objects, most of these studies do not require observers to determine what precisely was different. Instead, they simply need to discriminate whether the object was

the same one or not. Perhaps the change detection task is somehow more sensitive to the contribution of extraretinal information. If so, then object recognition performance may not be differentially affected by the two types of view change.

Furthermore, object and layout recognition might rely on different neural processing. In some prominent models, spatial orientation and updating tasks rely on the dorsal "where/how" pathway, whereas object recognition relies more heavily on the ventral "what" pathway (Goodale & Milner, 1992). For example, evidence from single-cell physiology suggests that cells in the inferior temporal cortex often fire in response to individual objects (e.g., Tanaka, 1993). Furthermore, fMRI evidence suggests a dissociation in the regions that respond to faces and other objects and the regions that respond to spatial layout (Epstein & Kanwisher, 1998; Tong, Nakayama, Vaughan, & Kanwisher, 1998). When observers view faces, areas in the fusiform gyrus are relatively more active (Tong et al., 1998). In contrast, when observers view spatial layouts or landmarks, parahippocampal regions are more active (Epstein & Kanwisher, 1998). Together, these findings suggest that spatial layouts may be represented differently than individual objects. If so, individual object recognition might not benefit from the extraretinal information available during a viewpoint change.

Alternatively, studies of change detection with object layouts and studies of navigation and spatial updating all suggest that extraretinal information plays an important role in representing and recognizing our surroundings. Furthermore, some recent studies suggest that object and layout recognition are similarly affected by display orientation changes. For example, using an old/new recognition task with photographs of object layouts, Diwadkar and McNamara (1997) found that observers were slower to recognize the test array the further it was rotated from the studied view. This result parallels experiments demonstrating view-dependent recognition of individual objects (e.g., Edelman & Bülthoff, 1992; Tarr & Bülthoff, 1995; Tarr & Pinker, 1989). Furthermore, just as in individual object recognition, receiving multiple views of an array of objects improves performance (Shelton & McNamara, 1997). These results suggest the two processes may rely on similar types of representations. If so, then the difference between orientation and viewpoint in change detection for layouts of objects suggests that object recognition performance may similarly be better following viewpoint changes than orientation changes.

The studies reported here tested the possibility that recognition of individual objects across view changes would benefit from the extraretinal information provided by observer movement. No previous studies of object recognition have addressed this issue—previous studies typically have focused on the ability to recognize objects presented in isolation on a computer display. As a result, such studies can shed light only on object recognition

mechanisms that rely solely on the retinal projection of the object. That approach can provide insight into the visual information used for object representations, but models based solely on such information might have limited applicability to cases in which other information is available. The present studies seek to investigate the influence of extraretinal information on object recognition and to test whether the updating mechanisms discovered in studies of spatial navigation and layout representation can facilitate object recognition across viewpoint changes.

In four experiments, we compared the ability to recognize an individual object across view changes caused by display rotations and changes caused by observer movements. In Experiment 1, we compared recognition of real objects from the same and different views when the observer remained in the same viewing position during study and test. On the basis of earlier object recognition research with similar types of objects rendered on computer monitors, we expected performance to be disrupted following the orientation change. However, since most previous studies have relied exclusively on computer displays and have not explored the recognition of real objects, this experiment provides an important test of the generalizability of earlier results. Experiment 2 then tests the possibility that object recognition performance might be differentially affected by viewpoint and orientation changes of the same magnitude. If recognition depends only on the retinal projection of the object, then discrimination accuracy should be comparable for viewer movements and object rotations. However, if observers rely on extraretinal information for individual object recognition, performance should be better across viewpoint changes than across orientation changes. Experiment 3 tested whether the difference in the visual background is sufficient to account for performance differences between the viewpoint change and orientation change conditions of Experiment 2. Experiment 4 tested whether the visual background is necessary for the superior performance in the viewpoint change condition.

EXPERIMENT 1

Method

Apparatus. The experiment was conducted in a small room using an unpainted, wooden circular table (1.22 m in diameter) occluded by a screen with two viewing windows (Figure 1). The viewing windows were separated by 40° as measured from the center of the table. From these viewing windows, observers could see the object and table against the uniformly painted off-white walls of the room. They could also see the side of the computer display and the experimenter. A small transparent plastic rack was fixed at the center of the table, and this rack could hold an object so that it faced the midpoint between the two viewing windows during the study period. Objects were composed of three layers of small wooden blocks (1 cm^3 each), which were mounted together on a square piece of cardboard (Figure 2). The base of each object (closest to the cardboard) was a created by stacking two 3×3 arrays of blocks. On top of this base were 5 blocks that could be arranged onto any of the nine positions of the 3×3 surface. Also, either 1 or 2 blocks
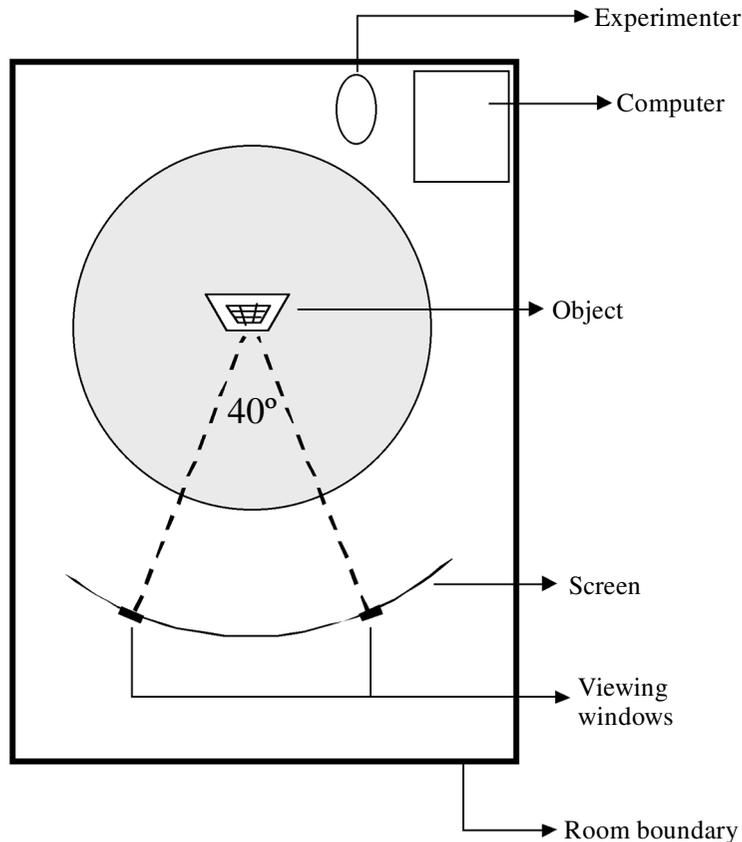
Figure 1. An overhead view of the apparatus for Experiments 1–3.

were attached to the base on each side of the object. Thus, each object shared a base of 18 blocks and differed in the positions of some of the remaining 11 blocks (each object had a total of 29 blocks). The objects were thus fairly similar to each other, but still discriminable. A total of 18 pairs of objects were created by matching objects that were most similar to each other. The two members of a pair shared the same configuration of blocks on the top surface (closest to the observer and farthest from the cardboard base), but the blocks on the sides were in different positions. Thus, the members of a pair were more similar to each other than to other objects in the set. Similar pairs were created in order to increase the difficulty of the object recognition task.[1] In the experiment, each of these pairs was seen four times, with a different edge facing upward for each of these trials. Thus, the study included a total of 72 (18 × 4) trials.

Procedure. Twenty-two undergraduate students at Harvard University participated in exchange for $8. Each subject participated in a total of 72 trials, divided equally into two conditions (orientation and same-view). In both conditions, subjects remained at a single viewing position throughout the experiment. On each trial, observers raised a curtain on the viewing window to view a single object mounted on the rack at the center of the table. After 3 sec, a computer-generated beep signaled the observer to lower the curtain. During the delay interval, the object was either replaced by its paired object (a different trial) or remained unchanged. After 7 sec, the computer beeped to signal the observer to raise the curtain to view the display. Subjects were then asked to write on a response sheet whether the object was the same object viewed initially or whether it was a different object.

For half of the trials, the table rotated by 40° during the 7 sec delay interval (orientation condition). For the other half, the table

remained stationary in front of the observer, so that subjects viewed the object from the same vantage point (same-view condition). Trials were administered in alternating blocks of six trials (e.g., six "orientation" trials and then six "same-view" trials). The block order was counterbalanced across subjects. A different random order of same and different trial types was generated for each subject, and different object pairs were assigned to these same and different trials for each subject.

In order to allow direct comparisons to the performance of observers in Experiment 2, half of the observers in this experiment (in both conditions) were asked to take a step to the side and then back during the delay interval. This movement was designed to make certain that any differences between viewpoint and orientation changes did not result from the absence of movement in the orientation condition (Simons & Wang, 1998; Wang & Simons, 1999).

Results

Subjects with overall accuracy (combining the same/different trials in both conditions) less than chance (50%) or more than 2 *SD*s from the group mean were excluded from the analyses. Data from 2 subjects were eliminated according to this criterion. We measured performance using accuracy (the proportion of correct responses, PC) and *A′* (one measure of the area under the receiver operating characteristic [ROC] curve; see Green & Swets, 1966; Pollack & Norman, 1964).[2]

Subjects performed better when the view was the same at study and test than when there was an orientation

Initial object

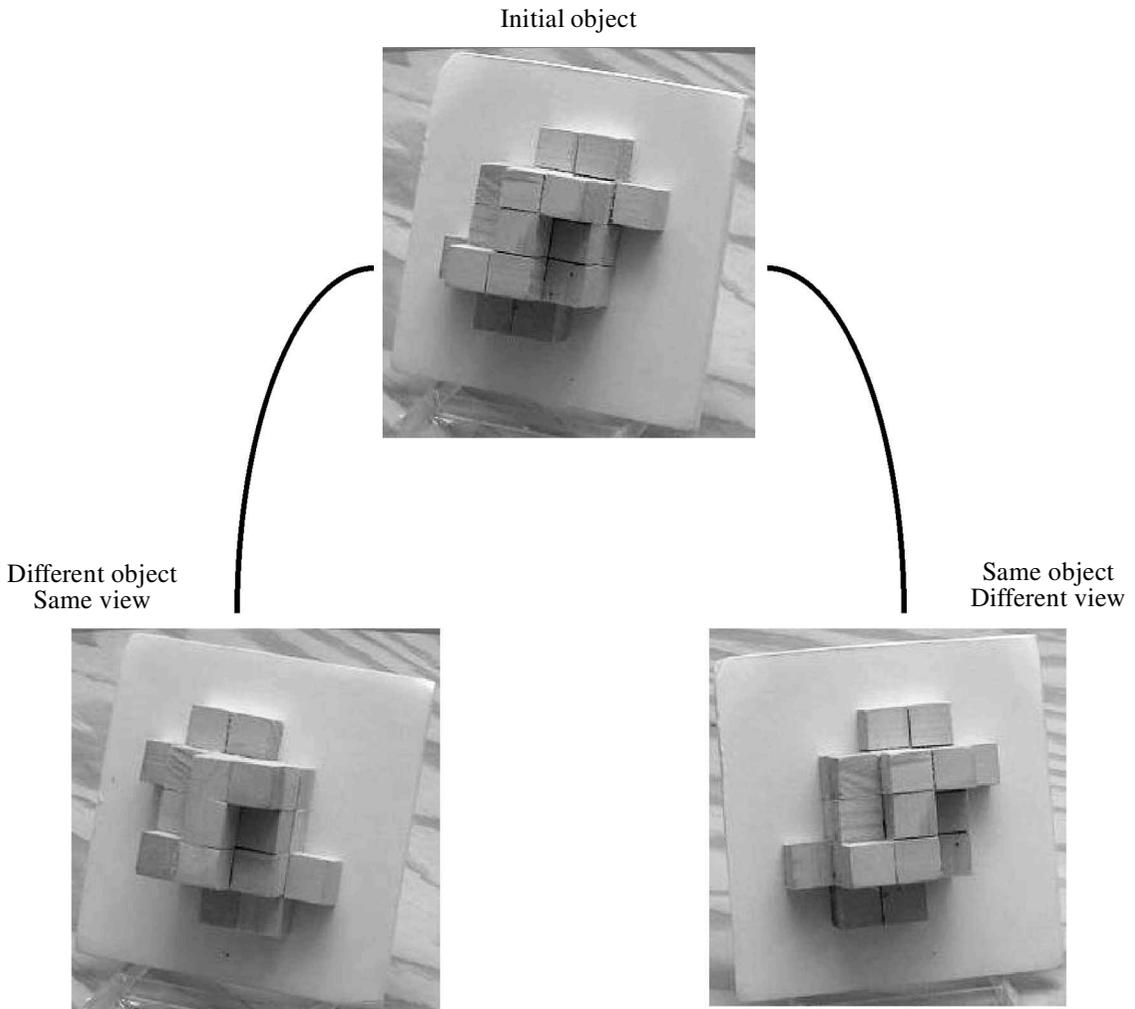Different object
Same view

Same object
Different view

Figure 2. Photographs illustrating the objects used in all four experiments. The top photograph depicts the initial view of an object. The left photograph shows a different object from the same view and the right photograph shows the initial object following an orientation change. The textured wooden table shown in the photographs was used in Experiments 1–3. In Experiment 4, the table was painted white.

change; they were both more accurate and showed better discrimination performance (Table 1). This difference was significant for overall accuracy [$t(19) = 2.70$, $p = .014$] and approached significance for A' [$t(19) = 1.83$, $p = .083$]. There was no significant effect of stepping to the side in the orientation condition (mean accuracy difference = .03; mean A' difference = .04; both ts < 1, p > .5).

This result replicates previous studies illustrating view-dependent recognition performance across orientation changes and extends such findings from computer-rendered objects to the recognition of real, 3-D objects, suggesting that object representations are view dependent. This pattern held even though we used accuracy as a dependent measure and most object recognition stud-

Table 1
Same View and Orientation Change in Experiment 1

Condition	Trial Type	Accuracy	Mean Accuracy	Sensitivity (A')	Criterion (β)
Same view	Same Different	.83 .68	0.76	0.83	1.41
Orientation	Same Different	.85 .55	0.70	0.79	1.75
t-test value (2-tailed)			$t(19) = 2.70$, $p = .014$	$t(19) = 1.83$, $p = .083$	$t(19) = 1.18$, $p = .252$

ies rely on more sensitive response time measures. The similarity between our results and earlier studies of object recognition also suggests that our experimental procedure and stimuli are compatible with those used in the computer-based studies in the literature.

The more interesting question, however, is whether viewpoint changes of similar size to the orientation changes in Experiment 1 would produce the same kind of disruption in object recognition. Studies on the detection of changes to object layouts produce similar orientation-dependent recognition, but performance is relatively unaffected by observer viewpoint changes. However, as noted earlier, recognition of individual objects may differ from the detection of changes to layouts of objects. It remains an empirical question whether the same mechanisms allowing superior performance for viewpoint changes with spatial layouts of objects will apply to object recognition. Experiment 2 addressed this issue by directly comparing the effects of viewpoint changes with orientation changes on individual object recognition.

EXPERIMENT 2

Method

Twenty-one Harvard undergraduate students participated in the experiment in exchange for $8. The apparatus and procedure of the second experiment were similar to those of the first experiment, except that in place of the same-view condition, we included a viewpoint change condition. For half of the 72 trials, observers walked from one viewing window to the other during the delay interval, thereby changing their view of the object by 40° (viewpoint condition). For the remaining trials, subjects stayed at the same viewing position and the table rotated by 40° during the delay interval (orientation condition). The orientation condition in this experiment was identical to that of Experiment 1. To control for the possibility that differences in performance for viewpoint and orientation changes might result from the differing amounts of observer movement in the two conditions, half of the subjects in the orientation condition were instructed to walk halfway to the other viewing position and then to return to their initial position during the delay interval. Thus, their view of the object was always from the same viewing position, but they walked approximately the same amount as subjects in the viewpoint condition.

Results and Discussion

Data from 1 subject were excluded from the analysis because overall accuracy was more than 2 SDs below the group mean. For both accuracy and discrimination measurements, observers were better following a viewpoint change than an orientation change (Table 2). This small

but consistent difference between observer movement and display rotation was significant for both accuracy [$t(19)$ = 2.53, p = .020] and A' [$t(19)$ = 2.54, p = .020]. Most subjects showed better performance in the viewpoint condition than in the orientation condition (15 for A', 12 for accuracy); a few showed no difference between the conditions (1 for A', 5 for accuracy); a few were better in the orientation condition (4 for A' and 3 for accuracy). Performance in the orientation condition was not different when subjects did or did not step to the side (mean accuracy difference = .02 in accuracy; mean A' difference = .01; both ts < 1, p > .4). Furthermore, recognition performance following a viewpoint change was not significantly different from the same-view condition of Experiment 1 in which observers received exactly the same view at study and test (for both accuracy and A', t < 1, p > .4). As in studies of layout change detection, object recognition took advantage of extraretinal information specifying the change in observer viewpoint, and thus was essentially invariant to observer movements.[3] Performance in the orientation change conditions of Experiments 1 and 2 also did not differ (for both accuracy and A', t < 1, p > .3).

However, there is an important difference between the viewpoint and orientation conditions that could potentially account for these results. Although the magnitude of the view change and the target itself (the object and the table) were identical in the two conditions, observers did receive different views of the room background when they moved and when the table rotated. That is, in the orientation condition the background was constant whereas in the viewpoint condition the background changed due to the shift in viewing position. Therefore, the two types of view change were not strictly identical, and the change of visual background in the viewpoint condition could have provided useful information about the magnitude of the view change. Experiment 3 tested whether this difference in the background visual information available for the view change might account for the relatively better recognition performance following viewpoint changes.

EXPERIMENT 3

Method

The design of Experiment 3 precisely replicated that of Experiment 2. However, rather than using real objects on a rotating table, observers viewed photographs of those objects taken from the ac-

Table 2
Viewpoint and Orientation Change in Experiment 2

Condition	Trial Type	Accuracy	Mean Accuracy	Sensitivity (A')	Criterion (β)
Viewpoint	Same	.74	0.77	0.86	1.03
	Different	.80			
Orientation	Same	.72	0.73	0.81	1.18
	Different	.74			
t-test value (2-tailed)			$t(19) = 2.53$, $p = .020$	$t(19) = 2.54$, $p = .020$	$t(19) = .84$, $p = .411$

tual viewing positions. Digital photographs of each object were taken from each viewing position using a tripod-mounted Nikon Coolpix 900 digital camera. Although the difference in the background information available from the two viewpoints was not extensive in this apparatus, as noted earlier, observers could see the experimenter and computer more clearly from one viewing position than the other, and the corner of the room was visible to observers from both positions. Note that the wooden texture of the table rotated with the object so that it provided the same information following orientation and viewpoint changes. Although such images cannot entirely reproduce the information available to an observer looking through the viewing windows, they did capture the difference in the background information available following a viewpoint change.

These 32-bit, 640 × 480 photographs were transferred to a computer and subjects were tested in a different room using an iMac computer (15-in. monitor). To be certain that the conditions in this experiment matched those of Experiment 2, the same randomly generated stimulus orders were used. Prior to viewing the computerized displays, subjects were shown the actual table apparatus and observed how orientation and viewpoint changes occurred. They were given the same instructions as subjects in Experiment 2.

Twenty-two Harvard summer school students participated in this experiment, and each received $5 for participating. On each trial, subjects viewed a photograph of an object (with the visual background visible) for 3 sec, followed by a blank screen for 7 sec. Then another photograph appeared (either the same object or its pair) and subjects were asked to determine whether the object was the same or different. This second object was photographed either from a different viewing window while the table was stationary (viewpoint condition) or from the same viewing window following a table rotation (orientation condition). The second photograph remained on screen until subjects responded same or different. Again we measured accuracy and A' for both the orientation and viewpoint conditions.

Results and Discussion

Data from 5 subjects were excluded from the analysis because their overall accuracy (combining both the orientation and the viewpoint conditions) was below chance (50%). Unlike in Experiment 2, observers were no better following viewpoint changes than orientation changes (Table 2).[4] Even though the visual information was comparable to that of Experiment 2, neither accuracy nor A'

performance was reliably different for the two types of change (both $ts < 1$, $p > .5$). This result suggests that the difference in the visual background does not account for the superior performance in the viewpoint change condition in Experiment 2. Instead, extraretinal information such as vestibular and proprioceptive information seems to contribute to performance in the viewpoint condition. Actual observer movement may be necessary to support improved recognition following a shift in viewpoint.

EXPERIMENT 4

Experiment 3 showed that visual background alone was not sufficient to explain the difference in performance between viewpoint changes and orientation changes. Note, however, that the design of Experiment 3 did not fully eliminate the possibility that visual background information contributed to performance in the viewpoint condition through an interaction between background information and observer movement. To test whether visual background was necessary for better performance in the viewpoint change condition, Experiment 4 replaced the screen used in previous experiments with a uniformly colored, circular curtain that completely surrounded the display. With this curtain, the visual background did not change regardless of the observer's position. The table in this experiment was painted white so that the texture underneath the object would be uniform.[5] If visual background information is necessary to specify the viewpoint change, there should be no difference between the viewpoint and orientation conditions when the background is constant. On the other hand, if visual background is neither necessary nor sufficient in specifying a viewpoint change, then Experiment 4 should replicate Experiment 2, revealing better performance when an observer moves than when the array rotates.

Method

Except as noted, the materials, design, and procedure were identical to those of Experiment 2. The apparatus was modified to provide
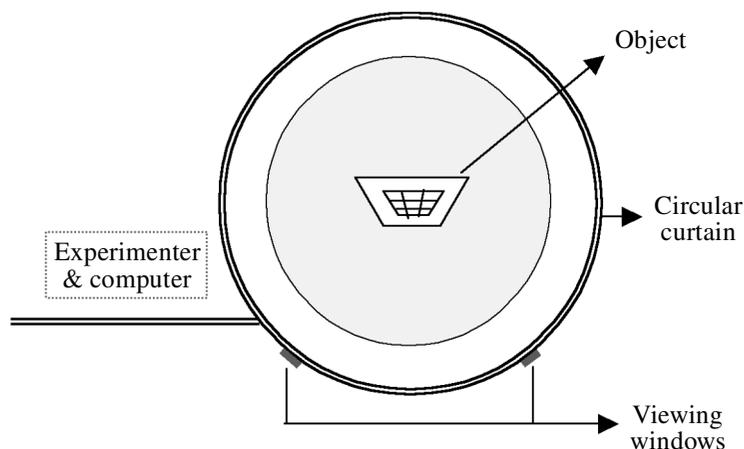

Figure 3. An overhead view of the apparatus for Experiment 4.

Table 3
Viewpoint and Orientation Change in Experiment 3

Condition	Trial Type	Accuracy	Mean Accuracy	Sensitivity (A')	Criterion (β)
Viewpoint	Same Different	.78 .56	0.67	0.76	1.35
Orientation	Same Different	.75 .59	0.67	0.75	1.12
t-test value (2-tailed)			$t(16) = 0$, $p = 1$	$t(16) = .46$, $p = .652$	$t(16) = 2.94$, $p = .010$

a uniform visual background from all viewing positions (Figure 3). The same 15 pairs of the objects used in Experiments 1–3 were employed. Objects were placed in the middle of a wooden rotating table (1 m diameter, 0.8 m high, painted white), surrounded by a gray circular curtain (1.9-m diameter) hung from a ring 2.2 m from the ground. The curtain was made of thick fabric with fine texture, so that any seams were virtually invisible. Two viewing windows (each 7 × 5 cm) were cut from the curtain, 1.1 m from the ground and 75° apart, as measured from the center of the table. The apparatus was illuminated by uniform, indirect light so that the objects would produce no shadows.

Sixteen undergraduate students from an introductory psychology class at the University of Illinois at Urbana-Champaign participated in the study for course credit. Each subject completed two practice trials, one in the viewpoint condition and one in the orientation condition. The goal of these practice trials was to familiarize subjects with the procedure, but not with the objects. Consequently, crumpled paper was used in place of the objects on these trials. Following the practice trials, subjects completed 72 test trials, divided into 12 blocks of 6 trials each. Half of the blocks were assigned to the viewpoint condition. For these blocks, subjects walked to a different viewing window during the 7-sec interval between study and test and the table remained stationary (a viewpoint change of 75°). The other half of the blocks were assigned to the orientation condition. For these blocks, subjects walked halfway toward the other viewing window and then returned to the original viewing position, and during the delay, the table rotated by 75°. Thus, both the change to the retinal projections of the test objects and the appearance of the visual background were identical in the two conditions. In each condition half of the trials involved an object change and the other half had no change. The order of all trials and blocks were randomized for each subject. As in the previous experiments, subjects were informed of the condition before each block.

Results and Discussion

Experiment 4 essentially replicated Experiment 2, with better performance in the viewpoint condition than in the orientation condition. The results were consistent in the face of several changes to the procedure, including the use of uniform visual background, a larger view change (75° instead of 40°), and a randomized order of

blocks rather than alternation (Table 3). As in Experiment 2, subjects were more accurate [$t(15) = 2.39$, $p = .030$] and more sensitive [A': $t(15) = 2.15$, $p = .048$] when the view change was caused by their movements than when the view change was caused by a display rotation. Of the 16 subjects, 11 (69%) performed better in the viewpoint condition than in the orientation condition. These results suggest that a change in the visual background is not needed for better performance in the viewpoint change condition and that the advantage for viewpoint over orientation changes holds even for somewhat larger changes.

GENERAL DISCUSSION

Individual object recognition is better following a viewpoint change than a display orientation change, and differences in the visual background information available following viewpoint and orientation changes do not appear to account for the difference in performance. As in previous studies of layout change detection, this difference appears to depend on an updating process that is available when observers move, but unavailable when the display rotates.

Traditional models of object recognition, at least in their current forms, cannot account for this difference between types of view change. Neither structural description nor alignment models of object recognition would predict that viewpoint changes and orientation changes of identical magnitude would differentially affect individual object recognition. Neither approach has incorporated the potential influence of extraretinal information into the recognition process. In essence, these models have considered only the change to the retinal projection resulting from the rotation of an object in front of a stationary observer. Thus, this difference between orientation and viewpoint leads us to question a

Table 4
Viewpoint and Orientation Change With Uniform Background in Experiment 4

Condition	Trial Type	Accuracy	Mean Accuracy	Sensitivity (A')	Criterion (β)
Viewpoint	Same Different	.83 .62	0.73	0.80	1.52
Orientation	Same Different	.80 .54	0.67	0.75	1.30
t-test value (2-tailed)			$t(15) = 2.39$, $p = .030$	$t(15) = 2.15$, $p = .048$	$t(15) = 1.07$, $p = .302$

basic assumption underlying work on the sensitivity of object recognition to view changes: namely, that object orientation can serve as a proxy for observer viewpoint.

More specifically, difference between orientation and viewpoint changes in individual object recognition are difficult to reconcile with models positing view-independent object representations. Such models could explain orientation-dependent recognition by appealing to the internal similarity of the stimulus set (e.g., Biederman & Gerhardstein, 1993). However, in so doing, they would have difficulty explaining the relatively diminished view-dependence given observer movement. Similarly, if such models note the relative view independence with observer movement as evidence in support of view-independent representations, then they must somehow explain the view-dependent performance with display rotations.

A difference between orientation and viewpoint changes might be somewhat easier to reconcile with models in which object representations are view dependent. However, no existing models incorporate a mechanism that allows representations to be updated for observer movement after a single view. Such models often posit interpolation between views as well as a graded decrease in performance as the test view deviates from the studied view. Consequently, they predict view-dependent performance across orientation changes. However, viewpoint changes include the same deviation from the studied view and consequently should succumb to the same decline in performance. Thus, a difference between viewpoint and orientation either would require additional updating mechanisms to be incorporated into models based on view-dependent representations or would require some means by which view-independent representations could be moderated by movement.

The results of Experiments 3 and 4 suggest that something about observer movement contributes to the difference between performance for viewpoint and orientation changes. This finding is consistent with earlier evidence that layout recognition continues to show a viewpoint advantage even when the retinal projections of the orientation and viewpoint changes are precisely equated (by using phosphorescent objects in a dark room). In Experiment 3, extraretinal influences were eliminated by presenting photographs of the displays on a computer monitor, and the difference in performance between viewpoint and orientation changes was eliminated. In Experiment 4, visual background was precisely matched so that the two conditions differed only in extraretinal information, and performance was higher in the viewpoint condition than in the orientation condition. Together, these findings strongly suggest that the difference between viewpoint and orientation is at least partially due to extraretinal influences on object recognition.

However, these data do not exclude the possibility that visual background information could be used to facilitate object recognition. The visual background in our apparatus was not very salient and did not seem to account for the different performance between viewpoint and orientation change conditions. Nevertheless, when made salient enough, visual background can potentially be a useful source of information to help object recognition (Christou & Bülthoff, 1997). Although the contribution of scene context information to the processing of individual objects is not entirely clear (Hollingworth & Henderson, 1998), our data do not eliminate the possibility that observers can flexibly use other available information in object recognition.

These studies of individual object recognition illustrate the importance of considering the conditions under which object recognition naturally occurs. Studies presenting objects in isolation on a computer display would be unlikely to discover differences between viewpoint and orientation or effects of background information. By looking at object recognition in a real-world context, we can gain a better appreciation for the mechanisms underlying our ability to recognize the same object from varying perspectives.

REFERENCES

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

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

Bülthoff, H. H., Edelman, S. Y., & Tarr, M. J. (1995). How are three-dimensional objects represented in the brain? *Cerebral Cortex*, **3**, 247-260.

Christou, C., & Bülthoff, H. H. (1997). *View-direction specificity in scene recognition after active and passive learning* (Tech. Rep. No. 53). Tübingen: Max-Planck-Institut für biologische Kybernetik.

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

Diwadkar, V. A., & McNamara, T. P. (1997). Viewpoint dependence in scene recognition. *Psychological Science*, **8**, 302-307.

Duhamel, J.-R., Colby, D. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, **255**, 90-92.

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

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, **392**, 598-601.

Farrell, M. J., & Robertson, I. H. (1998). Mental rotation and automatic updating of body-centered spatial relationships. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 227-233.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, **15**, 20-25.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General*, **127**, 398-415.

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

Landau, B., Spelke, E. S., & Gleitman, H. (1984). Spatial knowledge in a young blind child. *Cognition, 16*, 225-260.

Logothetis, N. K., & Pauls, J. (1995). Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cerebral Cortex, 3*, 270-288.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pelligrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: Assessment of path integration ability. *Journal of Experimental Psychology: General, 122*, 73-91.

Pollack, L., & Norman, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science, 1*, 125-126.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15*, 1157-1165.

Rieser, J. J., & Rider, E. A. (1991). Young children's spatial orientation with respect to multiple targets when walking without vision. *Developmental Psychology, 27*, 97-107.

Rock, I. (1973). *Orientation and form.* New York: Academic Press.

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

Shelton, A. L., & McNamara, T. P. (1997). Multiple views of spatial memory. *Psychonomic Bulletin & Review, 4*, 102-106.

Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science, 9*, 315-320.

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

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

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

Tarr, M. J., Williams, P., Hayward, W. G., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint-dependent. *Nature Neuroscience, 1*, 275-277.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron, 21*, 753-759.

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

Wang, R. F., & Simons, D. J. (1999). Active and passive scene recognition across views. *Cognition, 70*, 191-210.

Wehner, R., & Srinivasan, M. V. (1981). Searching behavior of desert ants, genus *Cataglyphis* (Formicidae, Hymenoptera). *Journal of Comparative Physiology, 142*, 315-318.

NOTES

1. The objects used in this experiment are novel and highly similar to each other. It is possible, perhaps likely, that more naturalistic objects would not be subject to the sorts of view dependence explored in these experiments. Objects that are more easily discriminable in terms of their distinctive parts show less view dependence (e.g., Biederman & Gerhardstein, 1993), and familiar objects might show even less view dependence. If we used objects that could be recognized equally well from all views, then we could not test the distinction between orientation and viewpoint. Our goal was to use a set of stimuli that would produce view-dependent recognition when rotated. We intentionally chose novel objects that were highly similar to each other so that recognition performance across views would be difficult. This set of stimuli allows us to explore differences in the degree of view dependence as opposed to simply looking for whether recognition is view dependent at all.

2. Because d' assumes normal distribution and the underlying distribution is unknown, we used the nonparametric A' as a measure of sensitivity; d' gave essentially the same results in all three experiments. For specific comparisons of accuracy and of A' values, we used t tests. However, the nonparametric equivalents of t tests (e.g., Wilcoxon tests) gave the same results.

3. Note that the cross-experiment comparison involves different groups of subjects, so the power to reveal a difference might be somewhat reduced. The between-subjects nature of this comparison somewhat tempers the claim that observer movement has no effect on performance, although it is consistent with some previous work on layout change detection (Simons & Wang, 1998). Note, however, that this comparison is not central to the primary claims of the paper. The comparison only tests whether updating is 100% effective, which we doubt would be true in most cases.

4. The performance was a little worse overall than that in Experiment 2, but the difference in the orientation change conditions of Experiments 2 and 3 was not significant [for accuracy, $t(35) = 1.04$, $p = .305$; for A', $t(35) = 1.41$, $p = .167$]. The difference in the viewpoint conditions of Experiments 2 and 3 was highly significant [for accuracy, $t(35) = 2.66$, $p = .012$; for A', $t(35) = 2.35$, $p = .024$]. The slight decrease in the overall performance may be due to the quality of visual information (e.g., depth, resolution, contrast and luminance, etc.) in a computer display.

5. Note that the table's texture could not have contributed to differential performance in the earlier experiments because the table rotated with the object in both orientation and viewpoint changes.