Manipulating and Recognizing Virtual Objects: Where the Action Is

KARIN H. JAMES, G. KEITH HUMPHREY, and MELVYN A. GOODALE
University of Western Ontario

Abstract In an earlier report (Harman, Humphrey, & Goodale, 1999), we demonstrated that observers who actively rotated three-dimensional novel objects on a computer screen later showed faster visual recognition of these objects than did observers who had passively viewed exactly the same sequence of images of these virtual objects. In Experiment 1 of the present study we showed that compared to passive viewing, active exploration of three-dimensional object structure led to faster performance on a "mental rotation" task involving the studied objects. In addition, we examined how much time observers concentrated on particular views during active exploration. As we found in the previous report, they spent most of their time looking at the "side" and "front" views ("plan" views) of the objects, rather than the three-quarter or intermediate views. This strong preference for the plan views of an object led us to examine the possibility in Experiment 2 that restricting the studied views in active exploration to either the plan views or the intermediate views would result in differential learning. We found that recognition of objects was faster after active exploration limited to plan views than after active exploration of intermediate views. Taken together, these experiments demonstrate (1) that active exploration facilitates learning of the three-dimensional structure of objects, and (2) that the superior performance following active exploration may be a direct result of the opportunity to spend more time on plan views of the object.

Résumé Lors d'une étude antérieure (Harman, Humphrey et Goodale, 1999), nous avons démontré que les observateurs qui effectuaient eux-mêmes la rotation d'objets tridimensionnels nouveaux, présentés à l'écran d'un ordinateur, affichaient une reconnaissance visuelle de ces objets plus rapide que ne le faisaient les observateurs qui avaient passivement visionné la même suite d'images illustrant ces objets virtuels. Dans l'expérience 1 de la présente étude, nous avons montré que, par rapport à l'observation passive, l'exploration active de la structure d'un objet tridimension-
nel accélère la performance lors d'une tâche de « rotation mentale » qui met en jeu les objets explorés. De plus, nous avons évalué le temps que les observateurs consacraient à différentes vues particulières des objets pendant l'exploration active. Comme dans notre précédente étude, ils ont consacré la majorité de leur temps aux vues de côté et de face des objets plutôt qu'aux vues de trois-quarts. Cette préférence marquée pour les vues « en plan » d'un objet nous a amenés, dans l'expérience 2, à étudier la possibilité que le fait de restreindre les vues lors de l'exploration active à des vues en plan ou à des vues normatives pourrait entraîner un apprentissage différentiel. Nous avons constaté que la reconnaissance des objets se faisait plus rapidement après l'exploration active limitée aux vues en plan qu'après celle ne permettant que les vues normatives. Ces expériences réunies indiquent que 1) l'exploration active favorise l'apprentissage de la structure d'objets tridimensionnels; et que 2) la meilleure performance qui suit l'exploration active peut être une conséquence directe de la possibilité de consacrer plus de temps aux vues en plan de l'objet.

When we encounter objects for the first time, we often walk around them or manipulate them so that we can see them from a variety of perspectives. Yet in contrast to this "active" exploration, most laboratory-based studies of object recognition typically present only a limited set of object views to observers – and do not permit the observers to control the views that they see. Although this "passive" and view-restricted methodology allows the experimenter to control the visual information that is presented to the observer, it does not always reflect what happens in the real world. Moreover, there is a growing body of evidence that active exploration of scenes and objects leads to better learning.

A number of theorists, including the philosopher Merleau-Ponty (1961) and the psychologists Piaget

(1953), Gibson (1962, 1979), Held (1965), and Neisser (1976), have emphasized the importance of motor activity, including exploratory activity, in perception and cognitive development. Some quasi-anecdotal observations lend support to such a view. One such example comes from studies of blind people who learn to use tactile information to "see" the world. It has been shown that moving a video camera across a scene and projecting the "image" onto the surface of the skin through a matrix of vibrators, enables some blind individuals to have an experience of objects in the world. In other words, moving the camera over the scene does not simply result in a sensation on the surface of the skin, but generates instead a percept of objects "out there" in the world. This transformation, in which the stimulation at the skin is referenced to a distal source, is greatly facilitated if the blind person has active control of the camera (Bach-y-Rita, 1974; see also Epstein, Hughes, Schneider, & Bach-y-Rita, 1989).

Held and his colleagues were amongst the first to investigate the relation between self-produced movement and visual abilities. The classic studies of Held and Hein (1963) showed that kittens that were prevented from actively exploring the visual world, even though they received the same visual stimulation as their normal counterparts, failed to show normal behaviour in many visually guided tasks. In addition, Held's experiments on humans adapting to prismatic distortion of the visual world also demonstrated the importance of motor feedback in maintaining accurate visual coordination (e.g., Held & Freedman, 1963; Held & Hein, 1963). Adaptation to prismatic distortions was much better under conditions of active movement in which visual feedback was dependent on a participant's self-generated activity, than it was under more passive viewing conditions, in which the participant was moved about on a trolley or in a wheelchair (for review see Held, 1965; for critical discussion see Dolezal, 1982).

In a similar manner, Tong, Marlin, and Frost (1995) investigated the role of active exploration versus passive viewing in the formation of spatial representations of a 3-D virtual environment. The active participants steered and peddled a stationary bike while they traveled through a virtual visual world presented through a head-mounted LCD display. The passive participants were shown a video recording of what the active participants saw. The active participants developed more accurate representations of the spatial layout of this world than did the passive participants. Tong et al. suggested that the tight coupling that normally exists between motor output and visual input facilitates accurate representations of the environment.

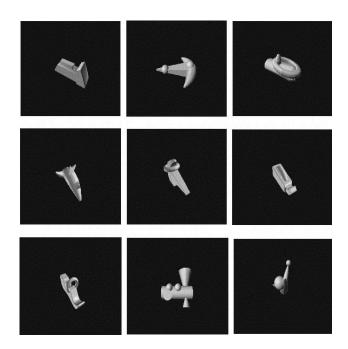Recent studies have also suggested that scene recog-

Figure 1. Examples of the novel, three-dimensional objects that were used in Experiments 1 and 2.

nition can be facilitated by active movement through a virtual environment when compared to passive movement (Christou & Bulthoff, 1999). In these studies, active explorers controlled their own movement through a virtual environment while passive observers watched a playback of the active explorers' route. To make sure that both the active and the passive participants were looking at the display carefully, they were required to respond to markers placed in the different scenes that unfolded on the display. In a recognition test, all participants were required to discriminate snapshots of the environment that they had just seen from snapshots of environments that they had never encountered before. The snapshots of the familiar environment were either scenes that had contained markers or were unmarked scenes. The active explorers were able to identify unmarked scenes in the familiar environment better than the passive observers, but there was no difference between the two groups on the marked scenes. The researchers concluded that spatial encoding may be more complete in active explorers.

The studies outlined above have suggested that recognition of some types of scenes can be affected by whether initial familiarization is achieved via active exploration of the information in the scene or by passive observation of this same information (but see also Simons & Wang, 1998). In a previous study, we found that the recognition of individual objects is also better with active exploration rather than passive observation (Harman et al., 1999). The experiment worked in the

following way. During active exploration, participants studied novel, three-dimensional objects (see Figure 1) by rotating the objects on a computer screen by means of a track ball. The objects could be rotated 360° about any axis. During passive observation, each participant viewed recorded rotations of objects that had been carried out by the preceding participant. Later recognition was tested with an explicit old-new task in which participants had to indicate for every test object whether or not they had seen it before (Harman et al., 1999). Participants were faster at making this decision with objects that they had actively explored as compared with those they had viewed passively.

We speculated that active control allowed participants to test predictions about how changes in viewpoint might affect the appearance of the object. That is, participants could "hypothesize" about how an object might look from different views and then store the trajectories that link one view to another. Although this kind of strategy might also work with passive observation, we argued that the links between views might be stored more effectively when the participants rotated the object from one view to another.

There could be another factor at work as well. It is possible that when the different views are stored under active exploration, later access to those views is accomplished by transforming an internal representation in much the same way as the real object was transformed on the computer screen. In other words, active exploration could make it easier to carry out "mental rotation" of stored object representations. The idea that motor processes and mental rotation are tightly linked has been proposed before. For example, Wolhschlager and Wolhschlager (1998) showed that "manipulations" of mental representations of visual stimuli can be equated with actual manipulations of visual stimuli. In their research Wolhschlager and Wolhschlager used a mental rotation task modeled after that originally used by Shepard and Metzler (1971). In such a task, participants are required to decide if a given object is the same as a rotated version of itself or not (for review see Shepard & Cooper, 1982). During this task, response times increase as a function of the angular difference between the two objects. Wolhschlager and Wolhschlager found that if participants physically rotated one object to match it to the other during their decision, response times increased at the same rate as they do in mental rotation (Wohlschlager & Wohlschlager, 1998). In addition, when translational hand movements were performed during the mental rotation task, these actions interfered with mental rotation – but only if the hand movements were along a different axis from that required during the mental rotation task. Wexler, Kosslyn, and Berthoz (1998) have also provided evi-

dence that motor processes influence transformations of mental representations. In their studies, they required participants to perform a typical mental rotation task using Shephard-Metzler figures while at the same time executing an unseen motor rotation with the hand. They found that motor rotation that was in the same direction as the required mental rotation resulted in faster response times for the mental rotation task than motor rotation in the opposite direction. In addition, they found that the speed of the motor rotation also affected the ease with which participants could perform the mental rotation task (Wexler et al., 1998). These researchers concluded that motor processes are not simply an end product of cognitive processes, but may be an integral part of cognitive operations in general.

Experiment 1

This evidence suggests that motor programs are invoked during the performance of tasks requiring mental rotation. It is not clear, however, whether or not the motor processes invoked during the encoding of an object representation can aid mental rotation at a later time. It was this question that motivated the present study. In other words, we investigated whether or not active exploration of a three-dimensional object would facilitate performance in a subsequent "perceptual match" task that is thought to involve mental rotation. In addition, the use of a perceptual match task provided a measure of how previous experience with an object affects performance without the need to recall specific encounters with that object. Finally, we attempted to replicate an earlier finding showing that participants explored objects in a somewhat stereotyped way, focusing their attention on particular views.

METHOD

Participants. Twenty-four right-handed students volunteered to participate in the present experiment, and were given course credit for their participation. All reported normal or corrected-to-normal visual acuity, and their ages ranged from 18-30 years (mean age = 20 years). Eleven males and 13 females participated.

Materials and apparatus. Study stimuli used during the familiarization phase consisted of 20 computer-rendered images of greyscale, three-dimensional novel objects (see Figure 1 for examples). They were presented on a black background with virtual overhead ambient lighting on a 15-inch computer monitor. Presentation of images and recording of participants' responses were controlled by a Macintosh G3 computer. All objects had a central axis of elongation and "geon-like" (Biederman, 1987) parts attached to a cen-
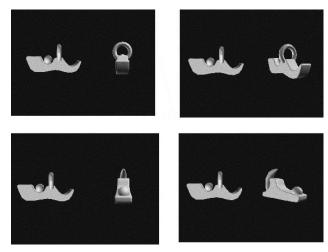
Figure 2. Example of the test session of Experiment 1, a perceptual match task. Two objects are shown on the screen at one time. Upper left: Same object where the target is a front view (90° rotation); Upper right: Same object where the target is a three-quarter view (45° rotation); Lower left: Different objects, where the target is a front view; Lower right: Different objects, where the target is a three-quarter view.

tral body. The object images were viewed from a distance of 60 cm. For the views in which the long axis of the object was perpendicular to the line of sight, the mean image size was 9 cm for the X dimension (8.5° of visual angle) and 6 cm for the Y dimension (5.8° of visual angle). For images in which the axis of elongation of the object was parallel to the line of sight, the mean size was 5 cm for the X dimension (4.8° of visual angle) and 6 cm for the Y dimension (5.7° of visual angle).

Images that were presented during the perceptual match test session were presented two at a time. The left object was termed the "referent" object as it was always presented in a "side" orientation where the axis of elongation was perpendicular to the line of sight. The "target" object was presented on the right-hand side of the screen, and was either a foreshortened view, where the axis of elongation was parallel to the line of sight, or it was a three-quarter view at a 45° rotation from the foreshortened view (see Figure 2).

PROCEDURE
Familiarization phase. Participants studied half of the objects actively and half passively. The active and passive trials were run in separate blocks, and the order of the blocks was counterbalanced. Within the blocks, the order of object presentation was randomized across participants. During active exploration, participants were told to study each object carefully from all angles, so that they had a good idea of the object's three-dimensional shape. After a practice session, participants moved a trackball that was 5 cm in diameter with

their right hand to rotate the object 360° around any axis. Participants were free to rotate each object for 20-s. During the passive viewing condition, each participant viewed a 20-s recording of the previous participant's active exploration of that object. The participants were again told to study each object carefully from all angles, so that they had a good idea of the object's three-dimensional shape. The data from the first participant were not used, but this active study was recorded and used as the passive component of the second participants' study session. Thus, the full yoked design began with the second participant. The order of the active and passive study sessions was counterbalanced across participants, and within each session the study objects were presented in a pseudo-random order.

The onset of the 20-s study period with each object was initiated by the experimenter, and the inter-trial interval was about 7 s. After studying all 20 objects (10 actively and 10 passively), the test session began.

Test phase. Each test trial was composed of a 500-ms fixation cross followed by a 100-ms blank screen and then the presentation of the test image. When the test image appeared, the participant was required to respond as quickly and as accurately as possible by pressing one of two keys on a keyboard to indicate whether or not the two images depicted the same object, or depicted different objects. Response latency and accuracy were recorded by a Macintosh G3 computer. There were four combinations of the test images: a) the two objects were the same object and the target object was rotated 90° from the referent (20 images); b) the two objects were the same and the target object was rotated 45° from the referent (20 images); c) the two objects were different and the target was rotated 90° from the referent (20 images); and d) the two objects were different and the target was rotated 45° from the referent (20 images). Therefore, 80 test images in total were presented to the participants. Participants were told in advance that the two objects would be the same or different, but that they would always be presented at different angles from one another. After the participant's response, there was an interval of 100 ms before the next trial began. This procedure continued until the participants responded to the 40 "same" images and the 40 "different" images.

RESULTS
Perceptual match task. Two separate repeated-measures 2 x 2 x 2 (active or passive study condition; same or different decisions; foreshortened or three-quarter target views) ANOVAS were run on the resultant data, one for the response latency data on the correct trials
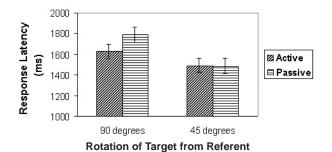
Figure 3. Interaction between study condition and degree of rotation of target from referent on response latency. Note that actively studying objects facilitates matching only a front view with a side view (90° rotation). Error bars indicate ± standard error of the mean.

and one for the accuracy data. Data removal was based on a sensitivity measure (Snodgrass & Corwin, 1988), which in our case was a correct decision rate ([hits+ correct rejections]/total number of trials) of 60% or greater. The data from three participants failed to reach this criterion and were therefore removed prior to the analysis.

Response latencies. The ANOVA revealed that there was a main effect of whether objects were studied through active exploration or passive observation, $F(1,20) = 4.85$, $p < .05$. Participants were able to perform a perceptual match task faster when objects were studied via active exploration ($M = 1,562$ ms) than via passive observation ($M = 1,649$ ms). There was also a main effect of target rotation, $F(1,20) = 42.62$, $p < .0001$; participants were able to match a target rotated by 45° ($M = 1,490$ ms) from the referent faster than a target rotated by 90° ($M = 1,711$ ms). In addition, there was also an interaction between study condition (active or passive) and target rotation (45° or 90°), $F(2,20) = 7.64$, $p < .01$. As depicted in Figure 3, active exploration significantly facilitated responses in the perceptual match task only when the target object was rotated 90° from the referent object (simple effects, $p < .05$). That is, active exploration facilitated performance on the more difficult matching task but not on the easier matching task.

Accuracy. Analysis of the accuracy data revealed no significant main effect for the active/passive study condition. There was, however, a main effect for the target-rotation condition, $F(1,20) = 103.8$, $p < .0001$, with the 45° target rotations being matched more accurately ($M = 96.5$) than the 90° rotations ($M = 82.7$) (Figure 4). Also, not suprisingly, "same" decisions ($M = 87.2$) were less accurate than "different" decisions ($M = 92.0$), $F(1,20) = 8.0$, $p < .01$.
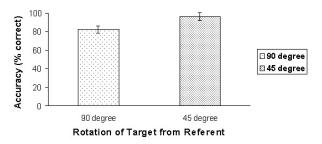

Figure 4. The effect of degree of rotation of target from referent on accuracy in the matching task. The 90° rotation is more difficult than the 45° rotation.
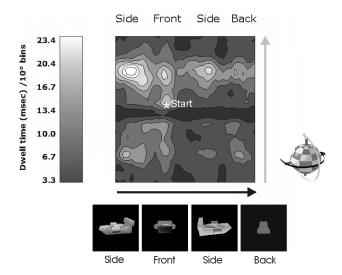

Figure 5. A contour map depicting dwell times during the exploration (study) phase of Experiment 1. The map is a representation of the flattened viewing sphere (right). This particular map is a mean of all actively explored objects and all participants. White areas represent higher dwell times and black areas represent lower dwell times. The top half of the map depicts dwell times about the vertical axis when objects are upright. (Most objects had a flat "bottom" allowing us to determine upright and inverted orientations.) The "start" orientation is left of centre and is a view of the object from the top. Thus the object required a rotation before it was in an "upright" orientation. Therefore, the pattern in this figure could not be an artifact of starting position. The spatial resolution of the dwell time calculation was 10°.

Exploration data. An analysis of the exploration data revealed that participants spent a majority of their study time focusing on four particular views of the objects. Specifically, participants focused on the front (foreshortened), the back, and the two side views of the object in an upright orientation, while virtually ignoring intermediate or three-quarter views (Figure 5). Participants spent significantly more time than expected on these "plan" views (in which the axis of elongation of the object was either parallel or perpendicular to the line of sight) than they did on the intermediate views, $X^2 (9) = 54.98$, $p < .001$.

DISCUSSION

When participants studied novel three-dimensional objects actively by rotating them on a computer screen, they performed faster on a later perceptual matching task with the same objects than they did with objects they had simply studied passively. This result replicates our earlier findings using a related paradigm (Harman et al., 1999), but suggests further that encoding an object via active exploration can facilitate performance on a task that is thought to involve mental rotation. The facilitation observed in the present experiment was evident, however, only when the test stimuli were 90° apart.

The difference in performance on the more difficult match task following active exploration vs. passive viewing could have occurred for two reasons. One possibility is that because the angular difference between the side view and the front views of an object was relatively large, the "mental rotation" necessary to match the two views was more demanding. But if the participants had the experience of physically rotating the same object (in the active exploration condition), then this experience could have facilitated a later mental rotation between the relevant views. Certainly, there is evidence that mental rotation might involve some of the same neural machinery as actual physical rotation (Wexler et al., 1998). In short, the early active rotation of the object could have "primed" the later mental rotation.

But there is a second possibility as well. Perhaps the particular views of the objects to be matched were both recognized more readily after active exploration than after passive viewing – and as a consequence were associated more readily with the same object representation. Our earlier report showed that both side and front views of objects were certainly recognized more quickly after active exploration than after passive viewing, whereas for the three-quarter view of the object that was most similar to the one used in the present experiment there was no effect of study condition on recognition latency (Harman et al., 1999). Thus, rather than facilitating mental rotation from one view to another, the active exploration condition may have created stronger associations between some sort of object template and different views of that same object.

Of course, these two explanations are not mutually exclusive. It could be the case that mental rotation is enhanced by virtue of the fact that different views of the object are "linked" more efficiently when the objects are actively rotated from one view to another. Mental rotation then would re-activate the same links between that which had been set up earlier when the object had been physically rotated. The participants who simply looked at the object rotation passively would not have had this opportunity to set up these links. This idea is reminiscent of an earlier proposal about mental rotation put forward by Edelman and Weinshall (1991).

The idea that the front and side views of the object might be directly linked by active exploration is reinforced by patterns of exploration exhibited by the participants. In both the present experiment and the earlier study, participants spent more time studying the front and side views than they did the intermediate or three-quarter views. All of this begs the question as to why participants used this particular strategy; in other words, why did they spend more time exploring the "plan" views than they did the intermediate views? As a first step towards answering this question we carried out another experiment.

Experiment 2

Perrett and his colleagues (1992) have proposed that when observers explore objects they concentrate on "plan" views, like the front and side views, because these views are "unstable" and can be thought of as singularities in the viewing space of an object. In other words, these are the views where there is the greatest amount of change in the visibility of the object features as the object is rotated by a small amount. Inspection strategies that concentrate on such views might facilitate the encoding of the object's three-dimensional structure. We can see now why observers would not dwell on any particular intermediate views. The intermediate views are all perceptually similar: all the major features of the objects are visible over a wide range of image projections. Thus, observers do not need to concentrate on one particular intermediate angle because of the high similarity among many of the successive images. This might explain why, in the present experiment and our earlier study, the participants deviated only a little from side to side when exploring a plan view; larger excursions would not have produced much more information than was already available.

Perrett and his colleagues (1988; Harries, Perrett, & Lavender, 1991) tested whether or not preferential inspection of the plan views of three-dimensional objects correlates with later recognition. They found that even though participants inspected plan views more than intermediate views, the amount of time spent on these views did not predict later performance on a test of recognition. But in these experiments, the emphasis was on how people inspected objects rather than on how different inspection strategies would affect later recognition. In our previous study, we also found no correlation between the time spent inspecting the plan views and the reaction time in the recognition test (Harman et al., 1999), but again the experiment was
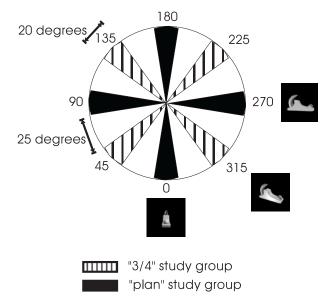
Figure 6. A representation of the study phase of Experiment 2. Objects were either studied by viewing three-quarter angles, or by viewing "plan" angles. The objects that were studied via "plan" views were only seen from the front (0° rotation), sides (90° and 270° rotations) and back view (180° rotation). We allowed the participants to have 10° of movement around each study angle. The objects that were studied via three-quarter views were only seen from the 45°, 135°, 225°, and 315° angles, again with 10° of movement around each angle. Thus, if the participants moved the object slowly, the object would appear to "jump" from view to view; however, most participants moved the objects fast enough that no jumps were apparent.

not designed to pull out the effects of inspection strategies on later recognition. Rather than relying on individual differences in the time spent on particular views to explore the relationship between viewing strategies and later recognition, we decided to examine quite directly whether or not limiting the availability of particular views would affect later recognition of the objects. In Experiment 2 then, we limited the views that participants were able to explore to either only plan views or only intermediate views of the study objects. We then tested their recognition of both plan and intermediate views of these objects.

METHOD
Participants. Participants were 24 undergraduate students (16 females, 8 males) who participated for course credit. Ages ranged from 18 to 28 with a mean age of 20. All participants reported normal or corrected-to-normal visual acuity and were naïve to the research design and to the appearance of the objects.

Materials. The stimuli used in the present experiment were the same three-dimensional images that were used in our previous studies. The experimental equipment and setup were identical to those used in

Experiment 1. In this experiment, however, the views that participants studied were limited. For half of the study objects, participants moved the objects on a computer screen around the vertical axis only and were able to explore only the intermediate views of the objects. For the other half of the objects, participants again rotated the objects about the vertical axis, but were able to explore only the plan views of the objects. The plan views were the 0°, 90°, 180°, and the 270° views (see Figure 6); however, we allowed 10° of movement around each of these views. Therefore, the visible angles became 350-10°, 80-100°, 170-190°, and 260-280°. When exploring the objects' intermediate views (see Figure 6), participants were able to explore views between 35-55°, 125-145°, 215-235°, and 305-320°. When the end of any range of movement was exceeded, the view of the object "jumped" to the closer angle of the neighbouring range. Image size and virtual lighting were the same as those used in Experiment 1. The plan and intermediate views study sessions were blocked and their order of presentation was counterbalanced, but stimuli within each block were presented in a different random order for each participant.

The test stimuli were four static views, each presented in isolation, of all 20 previously studied or "old" objects and four static views of 20 distracter or "new" objects. The four test angles in both cases were a front (0°) view, a side (90°) view, and two intermediate views, a 45° and a 225° view. Therefore, a total of 160 test images were presented in random order. Again, image size and virtual lighting were the same as those used in Experiment 1. Thus, for the objects that were studied by rotating to the plan views, the front and side test views were studied, but the two intermediate test views were not studied. Conversely, for the objects that were studied by rotating to the intermediate views, the two intermediate test views were studied, and the front and side test views were not studied.

Procedure. After instructions and a practice session, participants were presented with the first object to explore. Depending on the condition, participants would rotate either the plan-view objects in the first block, or the intermediate-view objects in the first block. Participants were able to rotate each object about the vertical axis for 20 s. After a 5-s intertrial interval, a new exploration trial was begun. Note that all objects were studied actively; there was no passive observation condition in the present study. There was no break between the two blocks, and participants were not told that the viewing conditions differed. After studying all 20 objects, the test session began.

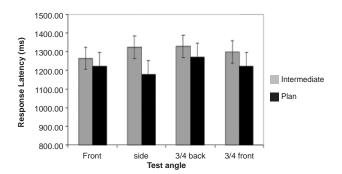During the test session, an image of an object would appear and participants were required to decide

Figure 7. The effect of study views on response latency in Experiment 2. Studying only plan views resulted in faster object recognition for each test view in a subsequent test task than did studying only three-quarter views.

whether or not they had seen the object during the study session. They indicated their "old-new" decision by pressing one of two keys on a keyboard. Following their response, a fixation-cross appeared for 500 ms followed by the next test image. An image remained on the screen until participants made a response. Participants were instructed to respond as quickly and accurately as possible.

RESULTS

Two separate 2 X 4 (Exploration Condition: plan views or intermediate views and Test Angle: front, side, intermediate front, and intermediate back) repeated-measures ANOVAS were run on the resultant data, one for response latency scores and one for accuracy scores. Data from one participant was removed because he scored less than 60% in total correct decisions.

Response latency. The ANOVA revealed a significant main effect of Exploration Condition. Objects whose exploration had been limited to plan views were recognized faster than objects whose exploration had been limited to intermediate views, *F*(1,22) = 8.78, *p* < .01 (Figure 7). There was no main effect of test angle, nor was there an interaction between Exploration Condition and Test Angle.

Accuracy. The overall mean accuracy was 70.7%. There were no significant effects of Exploration Condition or Test Angle on accuracy. Sensitivity (d') was also calculated and found to be not significantly different for the two study groups.

DISCUSSION

By limiting the views of objects that participants could explore, we were able to demonstrate that studying only plan views of objects results in better learning

than studying only intermediate views. Thus, when participants studied the plan views of objects they showed faster recognition of test objects than they did when they studied intermediate views of objects. This difference was found with all test views of the objects – no matter whether those test views were plan views or intermediate views, as reflected by the lack of interaction between study condition and test view (see Figure 7). In other words, studying only the plan views of an object appears to lead to a better representation of the three-dimensional structure of the object than does studying only the intermediate views.

As we discussed earlier, Perrett and his colleagues (1988, 1992) have proposed that the reason observers concentrate on "plan" views is because these views offer the greatest amount of change in the visibility of the object features as the object is rotated by a small amount. Inspection strategies that concentrate on such views would be important in the encoding of these particular views. In contrast, moving around the intermediate views would provide little new information about object features. The results of the present study support Perrett's conjecture and show that this "natural" viewing strategy is an efficient way of encoding the important object features. In other words, when viewing is artificially limited to the plan views, observers perform better on later discrimination tasks.

General Discussion

The results of Experiment 1 again show, like our earlier study (Harman et al., 1999), that active exploration of novel objects leads to better performance on later tests of object recognition. But Experiment 1 is the first study to demonstrate that active exploration can improve performance on a task thought to involve "mental rotation" of object representations. This perhaps is not surprising given recent research suggesting that mental rotation may involve motor processes (Wexler et al., 1998; Wohlshlager & Wohlshlager, 1998). Our results, however, extend this idea of a motor theory of mental rotation by suggesting that earlier experience manipulating objects that one is actually viewing may facilitate later mental rotation of representations of the objects in the "mind's eye."

These findings may have important implications for education and training – particularly in situations where individuals have to learn about the geometrical structure of complex objects. Although there is a long tradition of research and educational policy that has emphasized the importance of "learning by doing," this idea has not often been applied to perceptual learning in vision. One tradition that has emphasized the role of active exploration in perceptual learning is the Gibsonian tradition, starting perhaps with J.J. Gibson's

classic paper on "active touch" (Gibson, 1962). The present study underscores Gibson's original notion but shows that active control over the different views of an object that one is learning about can also improve later recognition. Thus, with the development of computer-based virtual displays of organic molecules, anatomical structures, architectural models, and other complex three-dimensional forms used for training and education, it might be useful to allow the student to control the rotation of these objects on the computer screen. Indeed, it has been claimed, albeit anecdotally, that having control over the way in which a four-dimensional object, such as a hypercube, is presented in a three-dimensional display allows a mathematician to develop a strong intuition about the structure of the hypercube (Davis & Hersh, 1981; Kellert, 1994).

The results of Experiment 2 show that the exploration strategies used by observers to study novel objects (i.e., concentrating on the plan views) in fact leads to better learning. When participants were limited to exploring only certain views of novel objects, they did better on later tests of recognition for objects that they had explored around the plan views than for objects that they had explored around the intermediate views. This result supports Perrett's suggestion that movement around plan views offers the most salient information about object features. But there is an apparent paradox here. One might have predicted, based on the research of Palmer, Rosch, and Chase (1981), that observers would have done better with the intermediate views than with the plan views. After all, there is a large body of research showing that individuals find it easiest to recognize the intermediate or "canonical" views of familiar objects that have a principal axis of elongation (e.g., Humphrey & Jolicoeur, 1993; Palmer et al, 1981; Warrington & Taylor, 1978; for review, see Jolicoeur & Humphrey, 1998). It is important to keep in mind, however, that in all those studies that have shown an advantage with canonical views, common objects have been used and the observers were presumably already familiar with their structure. In Experiment 2 in our study, participants were still learning about the object and their knowledge about the object's structure was being assembled from either a set of plan views or a set of intermediate views. It appears that having access to the plan views leads to a better representation of the object than having access to only the intermediate views. This perhaps explains why all the studies that have examined the way in which observers explore objects have shown that observers concentrate on the plan views (Harman et al., 1999; Locher, Vos, Stappers, & Overbeeke, 2000; Perrett & Harries, 1988; Perrett et al., 1991, 1992).

Address correspondence to Melvyn A. Goodale, PhD, C.R.C., Director, CIHR Group on Action and Perception, Department of Psychology, University of Western Ontario, London, Ontario N6A 5C2 (Tel: (519) 661 2070; Fax: (519) 661 3961; E-mail: mgoodale@uwo.ca).

References

Bach-y-Rita, P. (1974). Visual information through the skin – A tactile vision substitution system,. *Transitions of the American Academy of Opthalmology and Otolaryngology, 78*, 729-739.

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

Christou, C. G. & Bulthoff, H. H. (1999). View dependence in scene recognition after active learning. *Memory & Cognition, 27:6*, 996-1007.

Davis, P. J., & Hersch, R. (1981). *The mathematical experience*. Boston, MA: Birkhauser.

Dolezal, H. (1982). *Living in a World Transformed: Perceptual and Performatory Adaptation to Visual Distortion*. New York: Academic Press.

Edelman, S., & Wienshall, D. (1991). A self-organizing multiple-view representation of three-dimensional objects. *Biological Cybernetics*, 64, 209-219.

Epstein, W., Hughes, B., Schneider, S. L. & Bach-y-Rita, P. (1989). Perceptual learning of spatiotemporal events: Evidence from an unfamiliar modality. *Journal of Experimental Psychology: Human Perception and Performance, 13:1*, 28-44.

Gibson, J. J. (1962). Observations on active touch. *Psychological Review, 69:1*, 477-491.

Gibson, J. J (1979). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Harman, K. L., Humphrey, G. K., & Goodale, M. A. (1999). Active manual control of object views facilitates visual recognition. *Current Biology, 9,* 1315-1318.

Harries, M. H., Perrett, D. I., & Lavender, A. (1991). Preferential inspection of views of 3-D model heads. *Perception, 20*, 669-680.

Held, R. (1965). Plasticity in sensory-motor systems. *Scientific American, 213:5*, 84-94.

Held, R., & Freedman, S. J. (1963). Plasticity in human sensorimotor control. *Science*, *142(3591)*, 455-462.

Held, R., & Hein, A. (1963). Movement-produced stimulation in the development of visually guided behaviour. *Journal*

of Comparative & Physiological Psychology, 56(5), 872-876.

Humphrey, G. K., & Jolicoeur, 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(3)*, 137-159.

Humphrey, G. K., & Khan, S. C. (1992). Recognising novel views of three-dimensional objects. *Canadian Journal of Psychology, 46(2)*, 170-190.

Jolicoeur, P., & Humphrey, G. K. (1998). In V. Walsh & J. Kulikowski (Eds.), *Visual constancies: Why things look as they do.* Cambridge, MA: Cambridge University Press.

Kellert, S. H. (1994). Space perception in the fourth dimension. *Man and World, 27,* 161-180.

Locher, P., Vos, A., Stappers, P. J., & Overbeeke, K. (2000). A system for investigating 3-D form perception. *Acta Psychologia, 104,*17-27.

Merleau-Ponty, M. (1962). *Phenomenology of perception.* London, UK: Routledge & Kegan Paul.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology.* W.H. Freeman & Company.

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

Perrett, D. I., & Harries, M. H. (1988). Characteristic views and the visual inspection of simple faceted and smooth objects: "tetrahedra and potatoes." *Perception, 17,* 703-720.

Perrett, D. I., Harries, M. H., & Looker, S. (1992). Use of preferential inspection to define the viewing sphere and charac-teristic views of an arbitrary machined tool part. *Perception, 21,* 497-515.

Piaget, J. (1953). *The origins of intelligence in the child.* London: Routledge and Kegan Paul.

Shepard, R. N., & Cooper, L. (1982). *Mental images and their transformations.* Cambridge, MA: MIT Press.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,* 701-703.

Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science, 9:4,* 315-320.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117:1,* 34-50.

Stein, J. F. (1986). Role of the cerebellum in the visual guidance of movement. *Nature, 323,* 217-221.

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

Tong, F. H., Marlin, S. G., & Frost, B. J. (1995). *Cognition map formation in a three-dimensional visual virtual world.* Poster presented at the IRIS/PRECARN Workshop, Vancouver, BC.

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

Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition, 68,* 77-94.

Wohlschlager, A., & Wohlschlager, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology: Human Perception and Performance, 24:2,* 397-412.