ELSEVIER

Acta Psychologica 102 (1999) 265–292

acta
psychologica

The visual representation of three-dimensional, rotating objects

Zoe Kourtzi [a,b,*], Maggie Shiffrar [a]

[a] *Department of Psychology, Rutgers University, Newark, USA*
[b] *Department of Brain and Cognitive Science, MIT, E-10 238, 79 Amherst St., Cambridge, MA 02139, USA*

Received 8 April 1998; received in revised form 14 September 1998; accepted 10 December 1998

Abstract

Depth rotations can reveal new object parts and result in poor recognition of "static" objects (Biederman & Gerhardstein, 1993). Recent studies have suggested that multiple object views can be associated through temporal contiguity and similarity (Edelman & Weinshall, 1991; Lawson, Humphreys & Watson, 1994; Wallis, 1996). Motion may also play an important role in object recognition since observers recognize novel views of objects rotating in the picture plane more readily than novel views of statically re-oriented objects (Kourtzi & Shiffrar, 1997). The series of experiments presented here investigated how different views of a depth-rotated object might be linked together even when these views do not share the same parts. The results suggest that depth rotated object views can be linked more readily with motion than with temporal sequence alone to yield priming of novel views of 3D objects that fall in between "known" views. Motion can also enhance path specific view linkage when visible object parts differ across views. Such results suggest that object representations depend on motion processes. © 1999 Elsevier Science B.V. All rights reserved.

PsycINFO classification: 2323

Keywords: Object recognition; Motion; Priming

---

[*] Corresponding author. Tel.: +1-617-258-0670; fax: +1-617-253-9767; e-mail: zoe@psyche.mit.edu

PII: S0001-6918(98)00056-0

1. Introduction

As we move through the physical world, we are constantly presented with different 3D views of the objects within our environment. Indeed, both observer movement and object movement can cause the visible parts of an object to become occluded and previously hidden parts to become revealed. As a result, we frequently observe different object views with different part configurations. How does the visual system integrate different object views so that we can recognize the same object across changes in viewpoint?

Traditional models of object recognition addressing this question seem to cluster into two general categories: structural-description and image-based approaches. Structural-description models suggest that objects are represented based on the arrangement of their parts independent of the observer's viewpoint (Biederman, 1987; Hummel & Biederman, 1992; Marr & Nishihara, 1978). Image-based models propose that objects are represented as sets of multiple 2D views (Bülthoff, Edelman & Tarr, 1995; Edelman & Weinshall, 1991; Edelman, 1995; Perrett, Oram & Wachsmuth, 1998).

More specifically, structural-description and image-based models suggest different approaches to the recognition of unfamiliar views of an object rotating in depth. Structural-description models suggest that we can recognize unfamiliar views if they have the same part configuration as familiar views. For example, the "geon structural descriptions" (GSDs) model suggests that objects are represented by viewpoint-invariant volumetric primitives, known as geons (Biederman, 1987; Hummel & Biederman, 1992). Viewpoint-invariance can be achieved when objects can be decomposed into a distinct configuration of 3D parts that does not change with any object transformation (Hummel & Biederman, 1992). When these conditions are satisfied, object naming and matching can be performed in a viewpoint-invariant manner (Biederman & Gerhardstein, 1993).

Image-based models suggest that unfamiliar views can be recognized by extrapolating from "known" object views (Tarr & Pinker, 1989) or by interpolating between "known" object views (Bülthoff & Edelman, 1992; Edelman & Weinshall, 1991; Poggio & Edelman, 1990). Psychophysical studies suggest that both humans and monkeys can interpolate between familiar object views and recognize novel views falling within limited generalization fields that span up to approximately 45° from "known" orientations (Bülthoff & Edelman, 1992; Logothetis, Pauls, Bülthoff & Poggio, 1994; Logothesis, Pauls & Poggio, 1995). Consistent with these behavioural generalization fields, many inferotemporal (IT) neurons seem to respond selectively to the training orientation of an object and more broadly to neighboring orientations (Logothetis, Pauls, Bülthoff & Poggio, 1994; Logothesis, Pauls & Poggio, 1995).

There is evidence that this restricted generalization can be expanded with frequent exposure to multiple object views (Bülthoff & Edelman, 1992; Logothetis, Pauls, Bülthoff & Poggio, 1994; Logothesis, Pauls & Poggio, 1995) especially when these views are qualitatively similar (Cutzu & Edelman, 1994; Edelman, 1995; Liter, 1998). Moreover, the interpolation between "known" views can be facilitated when multiple object views are linked by temporal sequence (Edelman & Weinshall, 1991; Lawson, Humphreys & Watson, 1994; Miyashita & Chang, 1988; Miyashita, Date & Okuno, 1993; Wallis, 1996). As objects rotate in depth, we perceive multiple views in a temporal sequence.

Several studies suggest that temporal contiguity is important for linking multiple object views. For example, viewing structured sequences of multiple briefly presented views of a 3D object facilitates object naming over viewing random view sequences (Lawson, Humphreys & Watson, 1994). When presented with three different sequences of faces, each one consisting of five different faces in a different pose, subjects have greater difficulty discriminating between different faces from the same sequence than different faces from different sequences (Wallis, 1996). Also, viewpoint-invariant performance has been shown for rotating familiar and novel objects in a short-term recognition test when the study and the test views share the same parts (Srinivas, 1995). Finally, orientation priming is observed across blocks when the prime and the target objects are from the same visually homogenous class, but not when they are from different categories (Gauthier & Tarr, 1997). However, viewpoint-dependent performance has been observed for the same objects in a long-term recognition test (Srinivas, 1995). These results suggest that viewpoint-invariance can be achieved in memory for multiple temporally contiguous views of an object.

Furthermore, neurophysiological studies provide evidence for associative mechanisms based on image similarity or temporal contiguity. Cells in the inferotemporal cortex of monkeys become tuned to a small number of dissimilar visual patterns (colored fractal patterns) when these patterns are sequentially paired over many trials in a pair-associate task (Sakai & Miyashita, 1991, 1994) or when monkeys perform a standard or a delayed matching-to-sample task (Miyashita & Chang, 1988; Miyashita, Date & Okuno, 1993). These studies suggest that viewpoint-invariant performance can be observed when viewpoint-dependent object representations are temporally associated.

All of the above studies investigated object recognition across changes in the orientation of an otherwise static object. However, outside of the laboratory, changes in object orientation occur when objects and observers move. As a result, we are presented with continuous sequences of multiple object views that are highly similar to each other and close in space and time. Can the visual system take advantage of motion and integrate different views of 3D objects rather than associating static snapshots based solely on their similarity and temporal contiguity?

Previous research suggests that motion can link 2D object views and thereby facilitate viewpoint-invariance within the object's path of motion (Kourtzi & Shiffrar, 1997). The following experiments investigated whether motion could similarly enhance the integration of 3D object views more readily than temporal sequence and lead to the immediate construction of viewpoint-invariant representations of 3D objects rotating in depth. Can such a motion-based linkage occur even when depth rotations produce different visible part configurations? We were especially interested in how novel object views might be represented. Priming of novel views falling within the object's motion path would suggest the existence of viewpoint-invariant object representation across depth rotation. In Experiment 1, we asked whether priming would occur for novel views of 3D objects when the prime views were linked by motion. In Experiment 2, we asked whether priming would occur for novel views of 3D rotating objects when the prime views had different part configurations. Experiment 3 examined whether priming would occur for novel object views when the prime views were separated spatiotemporally but perceived in motion when an occluder was placed between them. Finally, in Experiment 4, priming for familiar objects across depth rotations was investigated.

2. Experiment

2.1. Unfamiliar objects rotating in depth

Numerous psychophysical studies have suggested that objects rotating in depth are represented in a viewpoint-dependent manner. As a result, observers show higher error rates and longer reaction times when identifying objects from novel views than from "known" views. For example, depth rotation of familiar objects has been shown to result in slower naming performance as the rotation causes an object to diverge from its canonical view (Palmer, Rosch & Chase, 1981). Performance in naming line drawings of familiar objects rotated in depth is best for canonical views and for views sharing similar image structures (Lawson, Humphreys & Watson, 1994; Lawson & Humphreys, 1996). Also, accidental foreshortened views of line drawings are more difficult to identify than non-foreshortened views of the same objects (Humphrey & Jolicoeur, 1993). Consistent with these object identification studies, naming photographs of familiar objects in a priming paradigm shows less priming when the same objects are presented a second time but at a different view (Srinivas, 1993). This priming decrease is larger when objects are studied in familiar views but tested in unfamiliar ones.

Similar performance decrements with changes in viewpoint have been reported for unfamiliar objects rotated in depth, such as wire-form objects (Rock & DiVita, 1987) and photographs of objects made of clay (Humphrey & Khan, 1992). Also, 3D objects rotated in depth have been shown to require more recognition time as the rotation angle increases (Shepard & Metzler, 1971).

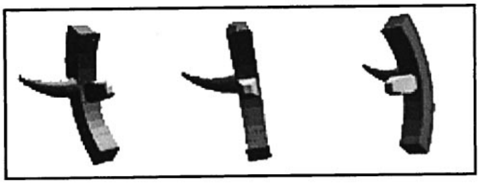
Recent studies have suggested that novel views of depth rotated objects can be recognized by extrapolating from (Tarr, 1995; Tarr & Pinker, 1989) or interpolating between (Bülthoff & Edelman, 1992; Edelman & Weinshall, 1991; Poggio & Edelman, 1990) "known" views. Naming unfamiliar views of 3D asymmetrical objects rotated in depth can be achieved by extrapolating from familiar views (Tarr, 1995). Naming performance with unfamiliar views depends on the distance from the training viewpoint. After practice, performance at the initially unfamiliar views is as good as performance at the training viewpoint, but performance at new unfamiliar views depends on the distance from the closest familiar view. Also, novel views of 3D objects rotated in depth can be recognized by interpolating between "known" object views (Bülthoff & Edelman, 1992; Edelman & Bülthoff, 1992). Specifically, recognition performance is best for "new" views of novel blob-like and tube-like objects within 45° of "known" views.

All of the above studies seem to suggest that objects rotating in depth are represented in a viewpoint-dependent manner. How are these viewpoint-dependent representations integrated so that we can perceive a unique object rotating in depth? Does motion facilitate the linkage of different views of a 3D object? If so, is view linkage confined to the object's path of rotation in depth?

Our previous research suggested that motion can link 2D object views and yield viewpoint-invariant representations of novel objects (Kourtzi & Shiffrar, 1997). Specifically, linking object views by apparent motion facilitated priming of novel object views falling within the path of motion, even for extended paths of motion. However, novel object views in between static "known" views were primed only when they were near "known" views. Moreover, novel view priming was found to depend upon the visual perception of apparent motion rather than on temporal sequence alone. When considered together, these results suggest that observers can recognize novel views of 2D moving objects more readily than novel views of 2D static objects.

In the current experiment, we asked if priming would occur for novel views of 3D objects when the prime views were linked by motion rather than presented statically in a temporal sequence. To investigate this question conservatively, we used novel, 3D geon-like objects (Biederman & Gerhardstein, 1993) as shown in Fig. 1(I). Novel views of these objects are thought to be difficult to recognize especially when they

I. Same part configuration



II. Novel part configuration

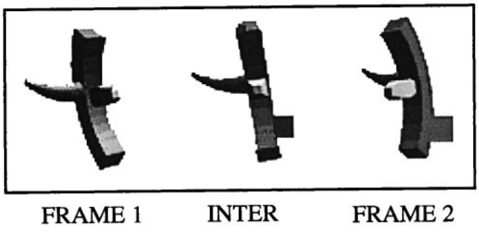


Fig. 1. An example of the novel geon-like objects used as stimuli (I) in Experiment 1, (II) in Experiment 2. The two prime views (FRAME 1 and FRAME 2) and the novel target view (INTER) in between these two prime views.

have part configurations that differ from those of "known" views (Biederman & Gerhardstein, 1993).

An immediate priming paradigm (Sekuler & Palmer, 1992) was used in which a briefly presented prime object is followed by a pair of targets. Subjects were primed with 2 object views linked by apparent rotation in depth (apparent-motion condition) or not linked by apparent motion (static condition). The second view was always a rotated version of the first view in the depth plane. Based on the results of previous studies (Kourtzi & Shiffrar, 1997), only 60° and 120° rotation angles were used. These studies showed that 60° rotations place novel views in sufficient proximity to prime views for priming to be observed under static conditions. However, 120° rotations resulted in priming for novel views only when the prime views were linked by apparent motion. Both rotation angles allowed the same object parts to be visible in the two prime views. In the following experiment, subjects judged if the two targets matched each other. Priming was indicated by faster reaction times when the two targets were the same as the prime object. Any differences in priming between moving and static objects would indicate differences in the manner in which we represent objects rotating in depth.

2.2. Method

2.2.1. Subjects

Forty undergraduate students, recruited from the Rutgers subject pool, participated in this experiment. All subjects had normal or corrected-to-normal vision and were naive to the hypothesis under investigation.

2.2.2. Materials

Stimuli were presented on a 21 in. color monitor with a 1024×768 pixel resolution and 60 Hz refresh rate controlled by a PowerMac 7100. The monitor was positioned 95 cm from a chin rest and the stimuli were drawn within a 4.82×4.82 degree of visual angle (DVA) square area on the screen. Subjects viewed the stimuli through a circular aperture to minimize framing effects from the monitor. This same apparatus was used in all of the experiments reported here.

The stimuli consisted of 20 three-part "geon-like" unfamiliar objects adapted from the novel objects of Biederman and Gerhardstein (1993). The stimuli were designed and smoothly rendered in Swivel 3D that generated displays with 72 dpi. The set of prime objects consisted of 5 objects. The 2 prime views of each object differed only by a rigid rotation of the object around the y-axis. The primes were sequentially presented in the center of the screen while the targets were simultaneously presented 0.6 DVA to the left and right of the screen center.

2.2.3. Procedure

Fig. 2 illustrates the experimental procedure. Each trial began with a fixation point presented for 1500 ms, followed by the first prime frame shown for a variable duration as described below. Then the second prime frame followed for the same duration as the first. A blank screen was then displayed for 500 ms followed by a pair

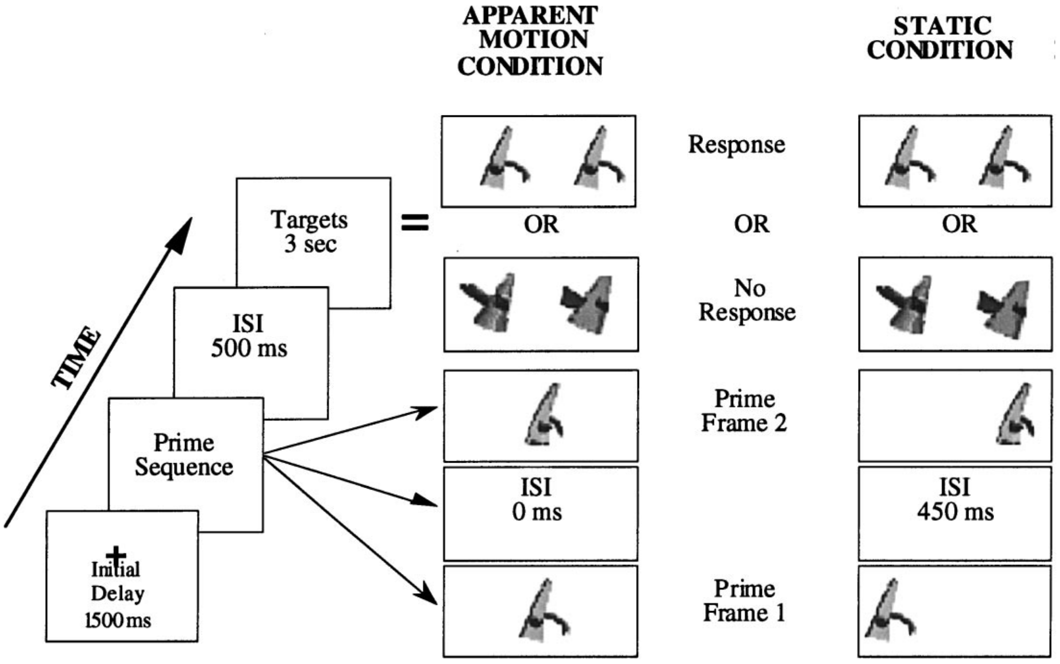


Fig. 2. Experimental design for the apparent motion and static conditions.

of targets presented until the subject responded (with a 3 s maximum). Subjects carefully observed the prime objects and then pressed a key if the two subsequent targets matched each other. This "Go-No Go" task was used to reduce the variability often observed in priming studies that require subjects to select one of two different motor responses (Biederman & Gerhardstein, 1993). Subjects were instructed that both reaction time and accuracy were important. Overall feedback (mean reaction time and percent correct responses) was provided at the end of each block of trials.

The first prime was presented at a view arbitrarily defined as the "starting" view. The second prime was rotated clockwise in depth around the y axis by a 60° or 120° angle relative to the first prime. The duration of the two prime frames varied with the rotation angle between them such that the optimal apparent motion in depth was achieved (Attneave & Block, 1973; Gerbino, 1984; Hecht & Proffitt, 1991; Kolers & Pomerantz, 1971). These durations were selected from pilot studies in which subjects reported the most compelling motion percepts when each prime frame was presented for the duration used by Shepard and Judd (1976) for the corresponding angle plus a constant of 100 ms. This yielded durations of 251 and 309 ms for the two rotation angles, respectively. The interstimulus interval (ISI) between the two prime frames was 0 ms in the apparent-motion condition and 450 ms in the static condition. In the apparent-motion condition, the first and second prime frames were presented so that the prime object appeared to rotate smoothly in depth. In the static condition, the second prime frame was displaced 2.41 DVA to the right of the first. This spatio-temporal separation between the two prime frames eliminated the perception of apparent motion in the static condition.

Before beginning the experimental trials, each subject completed a block of 10 practice trials with objects that differed from those of the experimental trials. Most subjects obtained reaction times less than 1000 ms by the end of the practice block. Subjects having longer reaction times completed a second practice block.

The experimental session consisted of 5 blocks each containing 20 trials. The target objects in each block were presented in 1 of 5 orientations around the y axis: the first orientation of the prime (FRAME 1), the second orientation of the prime (FRAME 2), the orientation half way between the two prime orientations (INTER), an orientation before the first prime orientation (EXTRA 1) or an orientation beyond the second prime orientation (EXTRA 2). The orientation of the INTER target equaled the first orientation of the prime plus half the rotation angle. The EXTRA 1 orientation equaled the first orientation of the prime minus half of the rotation angle. The EXTRA 2 orientation equaled the second orientation of the prime plus half of the rotation angle. Thus, the orientation of the INTER, EXTRA 1 and EXTRA 2 targets all deviated equally from the prime orientations. Table 1 shows all of the target orientations such that zero refers to the "starting" orientation.

Each target orientation was run in a separate block. Block order was counterbalanced across subjects. Stimulus order was randomized within each block. Each block contained 5 trials in which the targets matched each other as well as the prime, 5 trials in which the targets matched each other but differed from the prime and 10 trials in which the targets differed from each other and the prime.

Table 1
Orientations of the two target objects as a function of the prime rotation angle

Prime rotation angle	Target orientations				
	Same as prime		Different from prime		
	FRAME 1	FRAME 2	INTER	EXTRA 1	EXTRA 2
60	0	60	30	−30	90
120	0	120	60	−60	180

In a between-subjects design, two groups of 10 subjects completed the apparent-motion condition and two groups of 10 subjects completed the static condition. Each group of subjects observed stimuli at only one rotation angle so that every subject only viewed objects in novel orientations.

2.3. Results

In all of the experiments reported here, only reaction times to correct responses are reported because all subjects exhibited ceiling levels of performance. Priming is reported as a repeated measurement (within-measure variable), or the reaction time difference between trials in which the prime and targets were identical and trials in which the prime and targets differed. The results are reported on the basis of subjects and collapsed over items, because no systematic pattern of differences was observed between items.

Did priming occur? Repeated ANOVAs with Priming as the within-measure variable indicated significant priming for FRAME 1 ($F(1,39) = 33.1$, $p < 0.001$), FRAME 2 ($F(1,39) = 55.9$, $p < 0.001$), and INTER ($F(1,39) = 13.9$, $p < 0.001$), but not for EXTRA 1 ($F(1,39) = 1.5$, $p = 0.221$) or EXTRA 2 ($F(1,39) < 1$).

Amount of priming. A repeated ANOVA with Priming as the within-measure variable and Condition (apparent motion or static), Rotation Angle (orientation difference between the first and second prime), and Test Frame (FRAME 1, FRAME 2, INTER, EXTRA 1, or EXTRA 2) as the independent variables indicated significant main effects of Priming ($F(1,180) = 86.3$, $p < 0.001$), Rotation Angle ($F(1,180) = 15.0$, $p < 0.001$), and Test Frame ($F(4,180) = 12.6$, $p < 0.001$), but not of Condition ($F(1,180) = 1.7$, $p = 0.188$). A significant interaction was shown between Condition and Rotation Angle ($F(1,180) = 9.8$, $p < 0.01$).

Priming differences between the apparent motion and static conditions are summarized in Fig. 3. A repeated ANOVA with Priming as the within-measure variable and Rotation Angle and Test Frame as the independent variables showed significant main effects of Priming ($F(1,90) = 29.3$, $p < 0.001$) and Test Frame ($F(4,90) = 7.1$, $p < 0.001$) in the apparent-motion condition. No significant effect of Rotation Angle ($F(1,90) < 1$) was observed. Fisher's post hoc comparisons showed that FRAME 1 ($p < 0.001$), FRAME 2 ($p < 0.001$) and INTER ($p < 0.05$) were significantly more

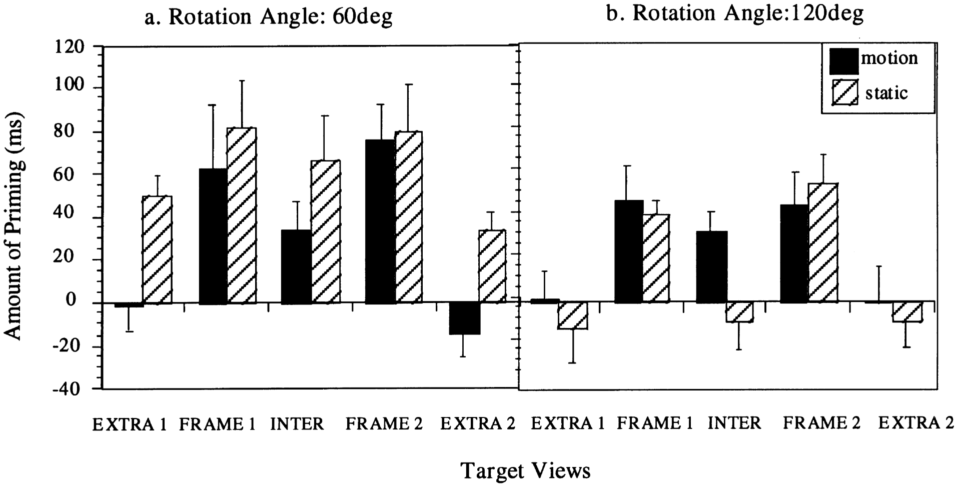


Fig. 3. Experiment 1: Amount of priming for target orientations in the apparent motion and the static condition for: (a) 60° rotation angle and (b) 120° rotation angle.

primed than EXTRA 1 and EXTRA 2. The same analysis in the static condition showed significant main effects of Priming ($F(1,90) = 61.3$, $p < 0.001$), Rotation Angle ($F(1,90) = 26.8$, $p < 0.001$) and Test Frame ($F(4,90) = 5.7$, $p < 0.001$). Fisher's post hoc comparisons showed that FRAME 1 and FRAME 2 were significantly more primed than INTER ($p < 0.05$), EXTRA 1 ($p < 0.01$), and EXTRA 2 ($p < 0.001$).

For the small 60° rotation angles, a one-way ANOVA with Priming as the dependent variable and Test Frame as the independent variable showed a main effect of Test Frame ($F(4,45) = 4.1$, $p < 0.01$) in the apparent-motion condition but not in the static condition ($F(4,45) = 1.4$, $p = 0.231$). Fisher's post hoc comparisons showed that FRAME 1 and FRAME 2 were significantly more primed than EXTRA 1 ($p < 0.01$) and EXTRA 2 ($p < 0.01$) in the apparent-motion condition. INTER was not significantly more primed than EXTRA 1 or EXTRA 2. However, priming for INTER was not significantly different from priming for FRAME 1 or FRAME 2. In the static condition, FRAME 1 was significantly more primed than EXTRA 2 ($p < 0.05$).

For large rotation angles (120°), a main effect of Test Frame was found in both the apparent-motion condition ($F(4,45) = 3.2$, $p < 0.05$) and static condition ($F(4,45) = 6.7$, $p < 0.001$). Fisher's post hoc comparisons showed that FRAME 1 and FRAME 2 were significantly more primed than EXTRA 1 ($p < 0.05$) and EXTRA 2 ($p < 0.05$) in the apparent-motion condition. INTER was not significantly more primed than EXTRA 1 or EXTRA 2. However, priming for INTER was not significantly different from priming for FRAME 1 or FRAME 2. In the static condition, FRAME 1 ($p < 0.01$) and FRAME 2 ($p < 0.001$) were significantly more primed than INTER, EXTRA 1 and EXTRA 2.

2.4. Discussion

Motion seems to link 3D object views more readily than the sequential presentation of static views. More precisely, the results indicate that target views at the same orientation as the primes were primed in both the apparent motion and static conditions for both the 60° and the 120° rotation angles. However, priming differences were observed between the two conditions when the target views differed from either prime orientation. Specifically, priming was observed for all the novel target views in the static condition but only for the 60° rotation angle. In the apparent-motion condition, on the other hand, priming was observed only for the novel target view falling in between the two prime views for both rotation angles. Target views falling outside either end of the rotation path were not primed in the motion condition.

These results suggest that the visual system can generalize across sequentially presented snapshots of a 3D object rotated in depth and link them together but when the angle of rotation separating these snapshots is small. A similar priming effect across only small depth rotations (67°) has been observed for familiar and novel objects when the study and test views shared the same parts (Srinivas, 1995). It is possible that large rotation angles decrease the similarity between 3D object views and make the generalization between temporally linked views more difficult. However, the visual system seems able to overcome this limitation by using dynamic cues such as motion. Motion appears to link 3D object views readily, even when these views are separated by large rotation angles, and to thereby facilitate generalization between "known" views. It is important to note that this generalization is restricted to the path of the object's motion. That is, our results indicate priming for novel views falling within but not outside the path of motion.

Thus, these results seem consistent with studies suggesting that observers can integrate multiple object views and recognize novel views based on restricted generalization fields around "known" object views (Bülthoff & Edelman, 1992). The similarity and temporal contiguity between different object views that are separated by small rotation angles can facilitate this generalization. However, 3D object views may be linked more readily by motion. Motion appears to expand the view-based generalization fields for larger rotation changes and to tune or sharpen them within the path of the object's motion. As a result, novel object views can be recognized readily for extended depth rotations as long as they fall within the object's path of motion. In the following experiment, we investigated the linkage of object views with different part configurations.

3. Experiment 2

3.1. Unfamiliar objects rotating in depth: novel part revealed

Depth rotations can change the apparent part configuration of an object. That is, as opaque objects rotate in depth, visible parts can become occluded and previously occluded parts can become visible.

Object recognition is thought to be particularly difficult when object views reveal novel part configurations (Biederman & Gerhardstein, 1993; Srinivas, 1993). For example, viewpoint invariance across depth rotation has been found in the naming of familiar objects in a priming paradigm (Biederman & Gerhardstein, 1993). However, this priming effect decreases for depth rotations that cause changes in the visible parts, that is, when some of the parts visible in the prime are occluded in the target objects and, symmetrically, when new parts became visible. Moreover, studying usual views of familiar objects does not result in priming for unusual views such as foreshortened views or views with occluded parts of the same objects (Srinivas, 1993).

The perception of a unique object rotating in depth may involve the linkage of object views containing different part configurations. If so, do motion processes underlie this linkage mechanism? Or, is temporal sequence alone sufficient? We investigated whether linking object views with different parts by motion or by temporal sequence would yield differences in the priming for novel object views. Can motion serve to link dissimilar views of objects rotating in depth across larger rotational changes than temporal contiguity alone?

3.2. Method

3.2.1. Subjects

Forty undergraduate students, recruited from the Rutgers subject pool, participated in this experiment. All subjects had normal or corrected-to-normal vision and were naive to the hypothesis under investigation. None had participated in the previous experiment.

3.2.2. Materials and procedure

The stimuli and the procedure used were the same as in Experiment 1. The only difference was that a novel part, which was either a cylinder, a rectangular block or a cone, was added to the second prime view, as shown in Fig. 1(II). This novel part appeared to be revealed by the object's rotation in depth. This novel part also appeared in all of the target views except for the target view at the orientation of the first prime view.

In a between-subjects design, two groups of 10 subjects completed the apparent-motion condition and two groups of 10 subjects completed the static condition. Each group of subjects observed stimuli at only one rotation angle.

3.3. Results

Did priming occur? Repeated ANOVAs with Priming as the within-measure variable indicated significant priming for FRAME 1 ($F(1,39) = 79.2$, $p < 0.001$), FRAME 2 ($F(1,39) = 47.9$, $p < 0.001$), and INTER ($F(1,39) = 23.5$, $p < 0.001$) but not for EXTRA 1 ($F(1,39) < 1$) or EXTRA 2 ($F(1,39) = 3.15$, $p = 0.083$).

Amount of priming. A repeated ANOVA with Priming as the within-measure variable and Condition (apparent motion or static), Rotation Angle (orientation difference between the first and second prime), and Test Frame (FRAME 1, FRAME 2, INTER, EXTRA 1, or EXTRA 2) as the independent variables indicated significant main effects of Priming ($F(1,180) = 131.8$, $p < 0.001$), Condition ($F(1,180) = 3.6$, $p < 0.05$), and Test Frame ($F(4,180) = 17.3$, $p < 0.001$) but no main effect of Rotation Angle ($F(1,180) = 2.5$, $p = 0.110$).

Priming differences between the apparent motion and static conditions are summarized in Fig. 4. A repeated ANOVA with Priming as the within-measure variable and Rotation Angle and Test Frame as the independent variables showed significant main effects of Priming ($F(1,90) = 72.3$, $p < 0.001$) and Test Frame ($F(4,90) = 8.5$, $p < 0.001$) in the apparent-motion condition. No significant effect of Rotation Angle ($F(1,90) < 1$) was observed. Fisher's post hoc comparisons showed that FRAME 1 ($p < 0.001$), FRAME 2 ($p < 0.001$), and INTER ($p < 0.01$) were significantly more primed than EXTRA 1 and EXTRA 2. The same analysis in the static condition showed significant main effects of Priming ($F(1,90) = 60.3$, $p < 0.001$), Rotation Angle ($F(1,90) = 5.8$, $p < 0.05$), and Test Frame ($F(4,90) = 10.1$, $p < 0.001$). Fisher's post hoc comparisons showed that FRAME 1 and FRAME 2 were significantly more primed than INTER ($p < 0.01$), EXTRA 1 ($p < 0.001$), and EXTRA 2 ($p < 0.001$).

For small rotation angles (60°), a one-way ANOVA with Priming as the dependent variable and Test Frame as the independent variable showed a main effect of Test Frame ($F(4,45) = 2.9$, $p < 0.01$) in the apparent-motion condition and in the static condition ($F(4,45) = 5.1$, $p < 0.01$). Fisher's post hoc comparisons showed that FRAME 1 and FRAME 2 were significantly more primed than EXTRA 1 ($p < 0.001$) and EXTRA 2 ($p < 0.01$) in the apparent-motion condition. Also, FRAME 1 and FRAME 2 were significantly more primed than EXTRA 1 ($p < 0.01$)

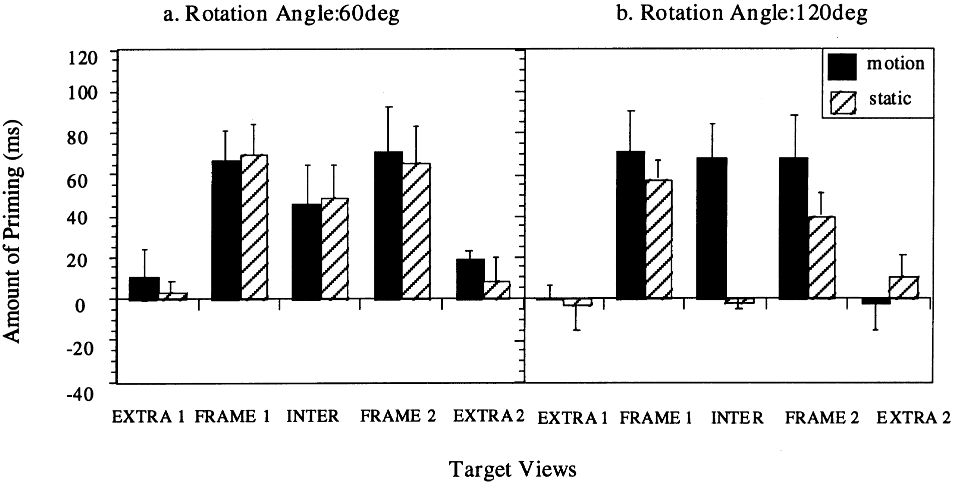


Fig. 4. Experiment 2: Amount of priming for target orientations in the apparent motion and the static condition for: (a) 60° rotation angle and (b) 120° rotation angle.

or EXTRA 2 ($p < 0.01$) in the static condition. INTER was significantly more primed than EXTRA 1 ($p < 0.05$) in the static condition.

For large rotation angles (120°), a main effect of Test Frame was found in both the apparent-motion condition ($F(4,45) = 6.3$, $p < 0.001$) and static condition ($F(4,45) = 6.9$, $p < 0.001$). Fisher's post hoc comparisons showed that FRAME 1, FRAME 2 and INTER were significantly more primed than EXTRA 1 ($p < 0.01$) and EXTRA 2 ($p < 0.01$) in the apparent-motion condition. In the static condition, FRAME 1 ($p < 0.001$) and FRAME 2 ($p < 0.01$) were significantly more primed than INTER, EXTRA 1, and EXTRA 2.

3.4. Discussion

These results suggest that different views of a 3D object which, as the result of a depth rotation show different object parts, can be linked together in a viewpoint-invariant manner within the object's motion path. Specifically, as in Experiment 1, target views at the same orientation as the prime views were primed in both the apparent motion and static conditions for both 60° and 120° rotation angles. In the apparent motion condition, priming was observed for the novel target view falling in between the two prime views for both rotation angles, but not for target views falling outside either end of the rotation path. In the static condition, no priming was observed for novel views with the 120° rotation angle, as in Experiment 1. However, unlike in Experiment 1, for the 60° rotation angle priming was observed only for the novel target view falling in between the two prime views. No priming was observed for novel target views falling outside either end of the rotation path.

These results suggest that motion can facilitate the immediate linkage of object views even when they have different part configurations. Thus, motion may tune generalization gradients between "known" views within an object's rotation path and facilitate such viewpoint-invariance across large depth rotations. However, presenting static object views with different part configurations in spatiotemporal sequence seems to restrict generalization to novel views in between the "known" views for small depth rotations. When the object views share the same parts, as in Experiment 1, spatiotemporal contiguity appears to facilitate generalization not only between but also beyond "known" views for small rotation changes.

Taken together, it is not obvious how these results might be explained by the classic structural-description models of object recognition (Biederman, 1987; Hummel & Biederman, 1992; Marr & Nishihara, 1978) since such models posit that novel object views can be recognized only when they share the same part configuration with "known" views. The current results may be more consistent with the hypothesis that objects are represented as collections of image features (Edelman & Weinshall, 1991; Perrett, Oram & Wachsmuth, 1998), such as image regularities (cusps, T-junctions) and symmetries (Koenderink & van Doorn, 1979). Recent studies have shown that these image features can affect the way in which we recognize different 2D object views (Wagemans, 1992, 1993; Wagemans, Van Gool & Lamote, 1996) and complete 3D objects that undergo self-occlusion as they rotate in depth (Van Lier & Wagemans, 1998). Depth rotations causing changes to structural image features result in viewpoint-dependent performance even when the part configuration remains the same (Hayward & Tarr, 1997; Tarr, 1989; Tarr & Chawarski, 1993; Tarr, Williams, Hayward & Gauthier, 1998). Further evidence for image-based representations sensitive to distinct object features, rather than to part configurations, comes from studies showing that object views with distinct object features are better recognized than other equally familiar views of the same object (Edelman & Bülthoff, 1992). Furthermore, discriminating between objects that differ qualitatively results in viewpoint-invariant representations (Liter, 1998). However, discriminating between objects that have unique part configurations but not distinctive features results in viewpoint-dependent representations (Tarr, Bülthoff, Zabinski & Blanz, 1997).

These viewpoint-dependent image representations can be integrated based on their feature similarity (Perrett, Oram & Wachsmuth, 1998). Temporal contiguity has also been shown to facilitate integration between object views even when the similarity between them is decreased (Wallis, 1996). Thus, it is possible that views of static 3D objects with different object parts can be linked together when presented sequentially. However, the current results in the static condition showed priming only for views falling in between the static prime views separated by small rotation changes. Novel views falling outside the rotation range were not primed unless they shared the same parts as the prime views (Experiment 1). These results suggest that spatiotemporal sequence alone can enhance generalization even between views with decreased similarity for small rotation changes. It is possible that the generalization fields tuned to specific views overlap when small rotations are used. Generalization fields tuned to spatiotemporally linked views seem to expand and facilitate viewpoint-invariance in the area of their overlap. Outside this area, novel views could be recognized only when they are highly similar to the "known" views. Experiment 4 investigated further the generalization between temporally contiguous views with familiar objects.

Motion appears to be a fundamental grouping principle (Wertheimer, 1923) or viewpoint integration tool used in the construction of image-based object representations. Local image features are thought to influence the perception of rigid object motion (Shiffrar, 1994). Yet, even dissimilar targets can be linked by motion (Burt & Sperling, 1981; Kolers & Pomerantz, 1971; Navon, 1976). Apparent motion may integrate information across time for complete objects even when the local features in each frame can not be matched point-by-point. Complementary 2D and 3D contour-deleted images (Biederman & Cooper, 1991) generated so that the line segments deleted from one image would be retained in the other can be linked by apparent motion as a function of the rotation angle in the frontal or in the depth plane (Koriat, 1994).

Consistently, the current results suggest that motion rather than spatiotemporal contiguity alone can readily link object views and facilitate viewpoint-invariance across larger depth rotations, even when these views have different part configurations. These results appear consistent with a recent study showing enhanced priming in the naming of familiar objects when multiple object views are presented in structured sequences linked by apparent motion than in random sequences (Lawson, Humphreys & Watson, 1994). However, this priming advantage was larger for locally similar views rotated by 30° in depth than for locally dissimilar views rotated by 60° in depth, even when the locally similar views were linked by apparent motion in the reverse direction so that the rotation path for the similar views (120°) was longer than for the dissimilar views (60°). Based on these results, Lawson, Humphreys & Watson (1994) concluded that the priming advantage observed for temporally contiguous views was due to local similarity between these views rather than to motion perception. A possible limitation of this study is that the effect of structural similarity and motion are confounded since similarity can facilitate motion perception (Foster, 1972, 1973).

It is also important to note that in the physical world, motion, temporal contiguity, and similarity may be functionally inseparable. A rigidly moving object yields a sequence of similar views. In both Experiment 1 and Experiment 2, temporal contiguity was not the same for the motion and static conditions. The two conditions differed in the perception of apparent motion and ISI duration: 0 ms for the motion condition and 450 ms for the static condition. Thus, it is possible that the apparent motion advantage at large rotation angles was found because the temporal contiguity between the views was higher. It is possible that large rotations result in object views with dissimilar image features that can be linked together, as in the motion condition, only if they are presented sequentially with no temporal separation between them. The following control experiment examined whether the perception of apparent motion, rather than temporal contiguity, caused these priming differences. To that end, an occluder was placed between the two static prime views so that the prime object appeared to be rotating in depth behind the occluder. Under these conditions, the spatiotemporal presentation of the prime views was identical in the static and motion conditions. If motion can link object views more readily than temporal sequence, then the priming effects found in the following experiment should be similar to those of the motion condition rather than to those of the static condition of Experiment 2.

4. Experiment 3

4.1. Linking static 3-D object views with a visible occluder

Taken together, Experiment 1 and Experiment 2 suggest that motion can link 3D object views more readily than temporal sequence between static views. In this experiment, we investigated further whether this advantage of motion over static presentation was due to motion perception or the higher temporal contiguity between the views in the motion condition. To that end, we used the same spatiotemporal separation between the prime views as in the static condition of the previous experiments. However, motion perception was induced by placing an occluder between the two static prime views during the 450 ms blank ISI in the static condition. As a result, the first prime view appeared to rotate in depth behind the occluder and reappear with a novel part added to it. Addition of the novel part to the second prime view, as in Experiment 2, allowed us to investigate the advantage of motion over temporal sequence in the linkage of object views across depth.

Placement of an occluding surface between two display frames has been shown to enhance the perceived spatiotemporal continuity between the elements in motion (Yantis, 1995). As a result, objects in apparent or real motion are perceived as continuing behind occluders over time (Michotte, 1963; Ramachandran & Anstis, 1983; Ramachandran, Inada & Kiama, 1986; Sigman & Rock, 1974; Tipper, Brehaut & Driver, 1990). Psychophysical studies have shown that observers are good at motion discrimination when objects move behind a "picket fence" occluder (Watamaniuk & McKee, 1995). Neurophysiological studies have demonstrated that direction-selective cells in the parietal cortex of a rhesus monkey respond to moving elements when the elements are behind an occluder during the interval when they fall on the cells' receptive field (Assad & Maunsell, 1995).

Moreover, there is evidence that static occluded objects are represented as whole objects rather than fragmented images (Bruno, Bertamini & Domini, 1997; Sekuler & Palmer, 1992) and that they are completed within the context of perceived 3D lay out (He & Nakayama, 1992, 1994; Kellman & Shipley, 1991, 1992; Nakayama & Shimojo, 1992; Shimojo & Nakayama, 1990). For example, a square object occluded partially by a rectangle and placed in stereoscopic depth behind it appears as a complete square (He & Nakayama, 1992). Based on these observations, we asked how objects are represented when they move behind occluders in 3D space.

4.2. Method

4.2.1. Subjects

Twenty undergraduate students, recruited from the Rutgers subject pool, participated in this experiment. All subjects had normal or corrected-to-normal vision and were naive to the hypothesis under investigation. None had participated in either of the previous experiments.

4.2.2. Materials and procedure

The stimuli were the same as those in the static condition of Experiment 2 with a novel object part revealed by the object's rotation in depth. The only difference was that a white bar (4.82×1.60 DVA) appeared as an occluder against the middle gray background. The bar appeared in the middle of the screen simultaneously with the prime object in the first frame and remained on the screen during the 450 ms interval between the two prime views and during the second prime frame. Thus, the first prime view was placed to the right of the bar in screen coordinates, while the second prime view was placed to the left of the bar. The spatial displacement between the two prime views was 2.41 DVA, as in the static condition in the previous experiments. As a result, the prime object appeared to be on the right side of the occluder, then disappear behind the occluder while rotating in depth and reappear to the left of the occluder. Pilot studies demonstrated that all subjects report the perception of an object rotating behind an occluder.

The procedure was the same as in Experiment 2. In a between-subjects design, two groups of 10 subjects completed the static condition with the occluder. Each group of subjects observed stimuli at only one rotation angle so that every subject only viewed objects in novel orientations.

4.3. Results

Did priming occur? Repeated ANOVAs with Priming as the within-measure variable indicated significant priming for FRAME 1 ($F(1,19)=66.2$, $p<0.001$), FRAME 2 ($F(1,19)=22.6$, $p<0.001$), and INTER ($F(1,19)=23.8$, $p<0.001$), but not for EXTRA 1 ($F(1,19)=1.02$, $p=0.324$) or EXTRA 2 ($F(1,19)<1$).

Amount of priming. A repeated ANOVA with Priming as the within-measure variable and Rotation Angle (orientation difference between the first and second prime), and Test Frame (FRAME 1, FRAME 2, INTER, EXTRA 1, or EXTRA 2) as the independent variables indicated significant main effects of Priming ($F(1,90)=60.8$, $p<0.001$) and Test Frame ($F(4,90)=9.2$, $p<0.001$), but not of Rotation Angle ($F(1,90)<1$). Fisher's post hoc comparisons showed that FRAME 1, FRAME 2 and INTER were significantly more primed than EXTRA 1 ($p<0.001$) and EXTRA 2 ($p<0.001$).

Priming differences between small and large rotation angles are summarized in Fig. 5.

A one-way ANOVA with Priming as the dependent variable and Test Frame as the independent variable showed a main effect of Test Frame for both small rotation angles (60°) ($F(4,45)=3.6$, $p<0.01$) and large rotation angles (120°) ($F(4,45)=6.1$, $p<0.001$). Fisher's post hoc comparisons showed that FRAME 1, FRAME 2 and INTER were significantly more primed than EXTRA 1 and EXTRA 2.

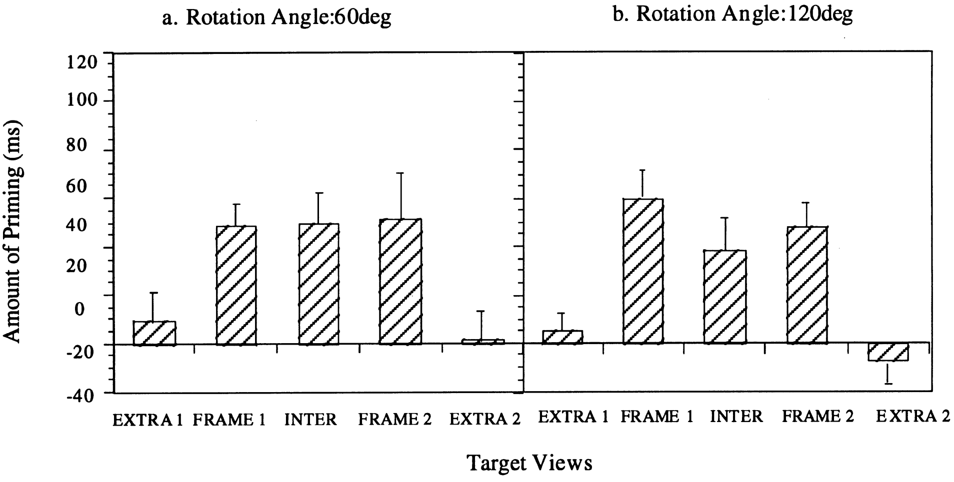


Fig. 5. Experiment 3: Amount of priming for target orientations at: (a) 60° rotation angle and (b) 120° rotation angle.

4.4. Discussion

The priming effect observed when placing an occluder between two spatiotemporally separated views was similar to the priming effect observed in the motion condition of Experiment 2. Specifically, target views with the same orientation as the primes were primed for both 60° and 120° rotation angle. Priming was also observed for the novel target view positioned in between the two prime views for both rotation angles. No priming was observed for target views falling outside either end of the rotation path.

These results provide further evidence that the visual system can take advantage of real world cues, such as motion and occlusion, to integrate more readily multiple object views. Introducing motion perception between spatiotemporally separated object views by placing an occluder between the views seems to facilitate viewpoint-invariance within the object's path of motion. Thus, motion perception rather than spatiotemporal contiguity between views appears to be the dominant factor in the integration of object views across rotational changes. Temporal contiguity and similarity between object views can be used by the visual system to link static object views within restricted generalization fields.

The following experiment investigated whether this advantage of motion over spatiotemporal contiguity generalizes to familiar objects. Observers are more likely to have experienced multiple views of familiar objects in spatiotemporal sequence. As a result, it is entirely possible that static views of familiar objects will be readily linked when they are presented in temporal sequence across large rotation changes.

5. Experiment 4

5.1. Familiar objects rotating in depth: novel part revealed

All of the above experiments suggest an advantage of motion over temporal sequence in linking object views across depth rotations for novel objects. Numerous studies suggest viewpoint-dependent representations of familiar objects (Humphrey & Jolicoeur, 1993; Lawson, Humphreys & Watson, 1994; Lawson & Humphreys, 1996; Palmer, Rosch & Chase, 1981; Srinivas, 1993). In this experiment, we asked whether motion could link rotated views of familiar objects together and facilitate viewpoint-invariance within the object's motion path more readily than temporal sequence.

It is possible that the integration of views in temporal sequence can be facilitated by object familiarity. Recent studies (Lawson & Humphreys, 1996) have shown that matching rotated views of familiar objects is faster when the objects are presented at a familiar orientation (upright) than an unfamiliar orientation (inverted). Familiarity with an object could increase the probability of having experienced multiple views of the object across depth rotations. As a result, generalization between sequentially presented views may be enhanced when the views are familiar (Bülthoff & Edelman, 1992; Logothetis, Pauls, Bülthoff & Poggio, 1994; Logothesis, Pauls & Poggio, 1995).

Differences between motion and temporal contiguity were investigated by the addition of a novel object part in the second prime view, as in Experiment 2. This manipulation resulted in decreased similarity between the prime views.

5.2. Method

5.2.1. Subjects

Forty undergraduate students, recruited from the Rutgers subject pool, participated in this experiment. All subjects had normal or corrected-to-normal vision and were naive to the hypothesis under investigation. None had participated in any of the previous experiments.

5.2.2. Materials and procedure

The stimuli consisted of a set of familiar 3D objects (Biederman & Gerhardstein, 1993). As the novel objects used in the previous experiments these familiar objects were designed and smoothly rendered in Swivel 3D that generated displays with 72 dpi. There were five prime objects, five pairs of two target objects that matched each other but differed from the prime and ten pairs of two target objects that differed from each other and the prime. To test conservatively object-specific priming across novel views, the five pairs of target objects that matched each other but differed from the prime were selected from the same category as the prime objects. The set of the five prime objects consisted of the following objects: 35 mm camera, ballpeen hammer, cabin house, grand piano, crosscut saw. The set of the five pairs of two target objects that matched each other but differed from the prime consisted of the following objects: video camera, claw hammer, suburban house, piano, hacksaw. Thus, any differences in the reaction times between targets that were the same as the prime and the targets that differed from the prime could not be attributed to categorical or semantic differences between the objects.

As in Experiment 2, a novel part, namely, a cylinder, a rectangular block or a cone, appeared on the second prime view. This novel part appeared to be revealed by the object's rotation in depth. This novel part also appeared in all the target views except for the target view at the orientation of the first prime view.

The procedure was the same as in Experiment 2. In a between-subjects design, two groups of 10 subjects completed the apparent-motion condition and two groups of 10 subjects completed the static condition. Each group of subjects observed stimuli at only one rotation angle so that every subject only viewed objects in novel orientations.

5.3. Results

Did priming occur? Repeated ANOVAs with Priming as the within-measure variable indicated significant priming for FRAME 1 ($F(1,39) = 109.2$, $p < 0.001$), FRAME 2 ($F(1,39) = 98.2$, $p < 0.001$), and INTER ($F(1,39) = 24.1$, $p < 0.001$) but not for EXTRA 1 ($F(1,39) = 3.6$, $p = 0.06$) or EXTRA 2 ($F(1,39) = 2.4$, $p = 0.128$).

Amount of priming. A repeated ANOVA with Priming as the within-measure variable and Condition (apparent motion or static), Rotation Angle (orientation difference between the first and second prime), and Test Frame (FRAME 1, FRAME 2, INTER, EXTRA 1, or EXTRA 2) as the independent variables indicated significant main effects of Priming ($F(1,180) = 152.2$, $p < 0.001$), Rotation Angle ($F(1,180) = 12.2$, $p < 0.001$), and Test Frame ($F(4,180) = 12.7$, $p < 0.001$), but no effect of Condition ($F(1,180) = 2.6$, $p = 0.108$). A significant interaction was shown between Condition and Rotation Angle ($F(1,180) = 7.2$, $p < 0.01$).

Priming differences between the apparent motion and static conditions are summarized in Fig. 6. A repeated ANOVA with Priming as the within-measure variable and Rotation Angle and Test Frame as the independent variables showed significant main effects of Priming ($F(1,90) = 53.3$, $p < 0.001$) and Test Frame ($F(4,90) = 8.2$, $p < 0.001$) in the apparent-motion condition. No significant effect of Rotation Angle ($F(1,90) < 1$) was observed. Fisher's post hoc comparisons showed that FRAME 1, FRAME 2 and INTER were significantly more primed than EXTRA 1 ($p < 0.001$) and EXTRA 2 ($p < 0.001$). The same analysis in the static condition showed significant main effects of Priming ($F(1,90) = 105.6$, $p < 0.001$), Rotation Angle ($F(1,90) = 21.1$, $p < 0.001$) and Test Frame ($F(4,90) = 5.7$, $p < 0.001$). Fisher's post hoc comparisons showed that FRAME 1 and FRAME 2 were significantly more primed than INTER ($p < 0.05$), EXTRA 1 ($p < 0.001$) and EXTRA 2 ($p < 0.01$).

For small rotation angles (60°), a one-way ANOVA with Priming as the dependent variable and Test Frame as the independent variable showed a main effect of Test Frame ($F(4,45) = 8.1$, $p < 0.001$) in the apparent-motion condition and in the static condition ($F(4,45) = 3.1$, $p < 0.05$). Fisher's post hoc comparisons showed that FRAME 1, FRAME 2 and INTER were significantly more primed than EXTRA 1 ($p < 0.001$) and EXTRA 2 ($p < 0.001$) in the apparent-motion condition. Also,

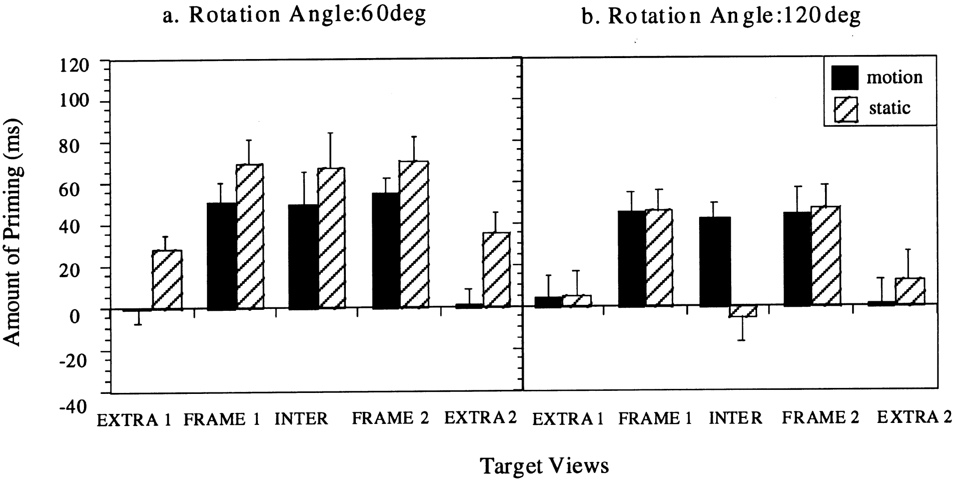


Fig. 6. Experiment 4: Amount of priming for target orientations in the apparent motion and the static condition for: (a) 60° rotation angle and (b) 120° rotation angle.

FRAME 1 ($p < 0.01$) and FRAME 2 ($p < 0.05$) were significantly more primed than EXTRA 1 and EXTRA 2 in the static condition. INTER was significantly more primed than EXTRA 1 ($p < 0.05$) in the static condition.

For large rotation angles (120°), a main effect of Test Frame was found in both the apparent-motion condition ($F(4,45) = 2.3$, $p = 0.05$) and static condition ($F(4,45) = 4.5$, $p < 0.01$). Fisher's post hoc comparisons showed that FRAME 1, FRAME 2 and INTER were significantly more primed than EXTRA 1 ($p < 0.05$) and EXTRA 2 ($p < 0.05$) in the apparent-motion condition. In the static condition, FRAME 1 and FRAME 2 were significantly more primed than INTER ($p < 0.01$), EXTRA 1 ($p < 0.05$) and EXTRA 2 ($p < 0.05$).

5.4. Discussion

The results of this experiment suggest that the visual system can readily link different views of familiar, rotating objects even when these views display different object parts. Moreover, the linkage of object views by temporal sequence alone appears to be enhanced for familiar objects as compared to novel objects. Specifically, as in Experiment 2, target views at the same orientation as the primes were primed in both the apparent motion and static conditions at both the 60° and the 120° rotation angles. In the apparent-motion condition, priming was observed for the novel target view falling in between the two prime views for both rotation angles, but not for target views falling outside either end of the rotation path. In the static condition, no priming was observed for novel views at the 120° rotation angle, as in Experiment 2. For the 60° rotation angle, priming was observed for the novel target view falling in between the two prime views as in Experiment 2. However, priming was also observed for novel target views falling outside either end of the rotation path, as in Experiment 1. Priming for novel target views outside the rotation path was less than priming for the novel view in between the prime views.

These results suggest that the role of motion processes in the integration of object views extends to the perception of familiar objects. Motion facilitates linkage across significant changes in orientation and visible parts of familiar objects rotating in depth. However, when the visual system is deprived of motion cues, object familiarity can facilitate some generalization across dissimilar object views presented in temporal sequence. These results could not be predicted by structural description models suggesting that novel views of familiar objects can be recognized when they share the same part configuration as "known" views (Biederman & Gerhardstein, 1993). These results seem rather to support image-based approaches to object recognition suggesting that practice with multiple object views can facilitate recognition of novel object views (Bülthoff & Edelman, 1992; Logothetis, Pauls, Bülthoff & Poggio, 1994; Logothesis, Pauls & Poggio, 1995). Moreover, the current results seem to provide a different perspective to these approaches by suggesting that generalization across views of familiar objects can be facilitated by motion. That is, we can represent familiar objects in a dynamic and invariant manner as they change while moving. This dynamic representation process seems to be similar for novel and familiar objects. Thus, the visual system seems to utilize motion cues not simply to construct representations of novel objects, but rather to update object representations and "monitor" object changes in the real world.

6. General discussion

In the physical world, motion can result in changes of an observer's image of an object. For example, when opaque objects rotate, some of their visible parts may become hidden and while other hidden parts may be revealed. Such changes are registered as retinal images with different object features. Nonetheless, we perceive unique objects rotating in depth rather than multiple disconnected images. How does the visual system integrate different views of an object as it rotates in depth?

Numerous studies have addressed this question in the laboratory by investigating properties, such as structural similarity and spatiotemporal contiguity between object views, which can facilitate linkage between separate image representations. However, the current results suggest that outside of the laboratory, the visual system may utilize properties of the physical world, such as motion and occlusion, to integrate multiple views of objects rotating in depth. When the visual system is deprived of such cues, then static views presented in temporal sequence can be linked together for small rotations.

More specifically, the results of Experiment 1 suggest that motion facilitates viewpoint-invariant representations of novel objects within the object's path of motion. Experiment 2 suggests that motion facilitates such viewpoint-invariance even when object views have different part configurations. Similarity between views can facilitate generalization between object views linked simply by temporal sequence. Experiment 3 suggests that the perception of motion, rather than temporal contiguity alone, strengthens view integration. Finally, the results of Experiment 4 suggest that object familiarity can enhance generalization between temporally linked views that are structurally dissimilar but separated by relatively small angles of rotation. Motion, however, is needed for the integration of different views of familiar objects that are separated by large rotations.

These results support theories suggesting that objects are represented in sets of viewpoint-dependent images (Bülthoff, Edelman & Tarr, 1995; Edelman, 1995; Edelman & Weinshall, 1991; Perrett, Oram & Wachsmuth, 1998) rather than as structural descriptions (Biederman, 1987; Hummel & Biederman, 1992; Marr & Nishihara, 1978). Some structural-description models suggest that novel object views can be recognized only when they share the same part configuration with familiar object views. However, the current studies suggest that novel views of rotating objects can be recognized even when a new part is revealed.

Importantly, the current experiments propose a different approach to our understanding of image-based object representations. Image-based models suggest that viewpoint-specific images are associated by temporal sequence and similarity in the analysis of three-dimensional object structure (Nakayama & Shimojo, 1992; Sinha & Poggio, 1996). The current experiments suggest that temporal sequence and similarity between static object views can facilitate viewpoint-invariance within restricted generalization fields. However, the visual system may use motion to integrate dissimilar object views across large rotational changes. Thus, motion perception rather than temporal contiguity and/or similarity may serve as the foundation for object recognition across changes in orientation. Object representations may be inherently dynamic (Freyd, 1987) as objects undergo changes in visible structure and viewpoint during their movement and/or the movement of the observer.

Taken together, these results seem consistent with the hypothesis proposed in several of the articles in this issue (see, for example, Lawson, 1998) that the visual system can employ multiple strategies when recognizing rotated 3D objects. As suggested by the 3D object recognition studies presented in this issue, the visual system can take advantage of structural properties such as global or/and local regularities (Van Lier, 1999) or volumetric constraints (Tse, 1999) to integrate static views of 3D objects rotated in depth and to construct complete object representations. However, outside of the frequently static laboratory setting, the visual system may follow a different, inherently dynamic, object recognition process. That is, visual analyses may depend upon dynamic object properties, such as motion, to update object representations continuously. Dissimilar object views created by depth rotation of self-occluded objects may be integrated "on the fly" during the analysis of the object's motion. It is therefore of interest to consider whether the visual system integrates different views of moving objects that violate image regularities and volumetric constraints, as in the case of non-rigidly deforming objects (Kourtzi & Shiffrar, 1998b) and highly complex biological movements (Kourtzi & Shiffrar, 1998a).

Finally, the current studies suggest an important interaction between the "what" or object system and the "where" or motion system (Ungerleider & Mishkin, 1982). The results of these studies suggest that these two systems may interact in important ways. That is, the object system seems capable of reconstructing 3D structure by generalizing between similar, sequential object views that are separated by small rotational changes. However, interactions between the "what" and "where" systems may underlie the visual system's ability to integrate multiple 3D object views within the object's path of motion, even when this path extends across large angles of rotation. Thus, the visual system can use motion cues to tune generalization across viewpoints within the object's path of motion. As a result, we can represent objects in a dynamic and viewpoint-invariant manner within their motion path. These dynamic object representations may facilitate the perception of object constancy and guide human actors to react promptly and successfully to object changes in the physical world. Based on studies discussed in this issue which suggest that observers perform better across object changes that result from self-motion rather than from object motion (Wraga, Creem & Proffitt, 1999), it will be interesting to investigate further how dynamic object representations may affect human-object interactions.

Acknowledgements

We would like to thank Peter Gerhardstein for providing us with his set of familiar objects and geons that we used to generate our stimuli. We also thank Mike

Tarr, Johan Wagemans and two anonymous reviewers for their helpful comments on a previous version of this manuscript. This research was supported by a grant from the Sigma Xi Foundation to the first author and a grant from the National Eye Institute (EY099310) to the second author.

References

Assad, J., & Maunsell, J. (1995). Neuronal correlates of inferred motion in primate posterior parietal cortex. *Nature*, *373*, 518–521.

Attneave, F., & Block, G. (1973). Apparent movement in tridimensional space. *Perception and Psychophysics*, *13*, 301–307.

Biederman, I. (1987). Psychological Review. *Recognition-by-components: A theory of human image understanding*, *94*, 115–117.

Biederman, I., & Cooper, E. E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology*, *23*, 393–419.

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

Bruno, N., Bertamini, M., & Domini, F. (1997). Amodal completion of partly occluded surfaces: Is there a "mosaic" stage?. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1412–1426.

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

Bülthoff, H. H., Edelman, S., & Tarr, M. J. (1995). How are three-dimensional objects represented in the brain?. *Cerebral Cortex*, *5*, 247–260.

Burt, P., & Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review*, *88*, 171–195.

Cutzu, F., & Edelman, S. (1994). Canonical views in object representation and recognition. *Vision Research*, *34*, 3037–3056.

Edelman, S. (1995). Class similarity and viewpoint invariance in the recognition of 3-D objects. *Biological Cybernetics*, *72*, 202–207.

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

Edelman, S., & Weinshall, D. (1991). A self-organizing multiple view representation of 3-D objects. *Biological Cybernetics*, *64*, 209–219.

Foster, D.H. (1972). A method for the investigation of those transformations under which the visual recognition of a given object is invariant. I. The theory. *Kybernetik*, *11*, 223–229.

Foster, D. H. (1973). A hypothesis connecting visual pattern recognition and apparent motion. *Kybernetik*, *13*, 151–154.

Freyd, J.J. (1987). Dynamic mental representations. *Psychological Review*, *94*, 427–438.

Gauthier, I., & Tarr, M. J. (1997). Orientation priming of novel shapes in the context of viewpoint-dependent recognition. *Perception*, *26*, 51–73.

Gerbino, W. (1984). Low-level and high-level processes in the perceptual organization of three-dimensional apparent motion. *Perception*, *13*, 417–428.

Hayward, W. G., & Tarr, M. J. (1997). Testing conditions for viewpoint invariance in object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1511–1521.

He, Z., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature*, *359*, 231–233.

He, Z.J., & Nakayama, K. (1994). Apparent motion determined by surface layout not by disparity or three-dimensional distance. *Nature*, *367*, 173–175.

Hecht, H., & Proffitt, D. R. (1991). Apparent extended body motions in depth. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 1090–1103.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, *99*, 480–517.

Humphrey, G. K., & Jolicoeur, P. (1993). An examination of the effects of axis foreshortening, monocular depth cues, and visual field on object identification. *The Quarterly Journal of Experimental Psychology*, *46A*, 137–159.

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

Kellman, P., & Shipley, T. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.

Kellman, P., & Shipley, T. (1992). Perceiving objects across gaps in space and time. *Current Directions in Psychological Science*, *1*, 193–199.

Koenderink, J. J., & van Doorn, A. J. (1979). The internal representation of solid shape with respect to vision. *Biological Cybernetics*, *32*, 211–216.

Kolers, P., & Pomerantz, P. (1971). Figural change in apparent motion. *Journal of Experimental Psychology*, *87*, 99–108.

Koriat, A. (1994). Object-based apparent motion. *Perception & Psychophysics*, *56*, 392–404.

Kourtzi, Z., & Shiffrar, M. (1997). One-shot view-invariance in a moving world. *Psychological Science*, *8*, 461–466.

Kourtzi, Z., & Shiffrar, M. (1998a). *Dynamic representation of biological movement.* Paper presented at the 5th Annual Meeting of The Cognitive Neuroscience Society, San Francisco, CA.

Kourtzi, Z., & Shiffrar, M. (1998b). *The representation of deforming objects as continuous events.* Paper presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Fort Lauderdale, FL.

Lawson, R. (1999). Achieving visual object constancy across plane rotation and depth rotation. *Acta Psychologica*, 102, 221–245.

Lawson, R., & Humphreys, G. W. (1996). View specificity in object processing: Evidence from picture matching. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 395–416.

Lawson, R., Humphreys, G. W., & Watson, D. G. (1994). Object recognition under sequential viewing conditions: evidence for viewpoint-specific recognition procedures. *Perception*, *23*, 595–614.

Liter, J. C. (1998). The contribution of qualitative and quantitative shape features to object recognition across changes of view. *Memory and Cognition*, *26*, 1056–1067.

Logothetis, N., Pauls, J., Bülthoff, H. H., & Poggio, T. (1994). View-dependent object recognition by monkeys. *Current Biology*, *4*, 401–414.

Logothetis, N., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology*, *5*, 552–563.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of three-dimensional shapes. *Proceedings of the Royal Society of London, Series B*, *200*, 269–294.

Michotte, A. (1963). *The perception of causality*. London: Methuen: Translated by T. & E. Miles from (1946) Edition.

Miyashita, Y., & Chang, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature*, *331*, 68–70.

Miyashita, Y., Date, A., & Okuno, H. (1993). Configurational encoding of complex visual forms by single neurons of monkey temporal cortex. Special Issue: The neuropsychology of animal memory. *Neuropsychologia*, *31*, 1119–1131.

Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science*, *257*, 1357–1363.

Navon, D. (1976). Irrelevance of figural identity for resolving ambiguities in apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 130–138.

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

Perrett, D. I., Oram, M. W., & Wachsmuth, E. (1998). Evidence accumulation in cell populations responsive to faces: An account of generalisation of recognition without mental transformations. *Cognition*, *67*, 111–145.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, *343*, 263–266.

Ramachandran, V. S., & Anstis, S. M. (1983). Displacement thresholds for coherent apparent motion in random-dot patterns. *Vision Research*, *23*, 1719–1724.

Ramachandran, V. S., Inada, V., & Kiama, G. (1986). Perception of illusory occlusion in apparent motion. *Vision Research*, *6*, 1741–1749.

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

Sakai, K., & Miyashita, Y. (1991). Neural organization for the long-term memory of paired associates. *Nature*, *354*, 152–155.

Sakai, K., & Miyashita, Y. (1994). Neuronal tuning to learned complex forms in vision. *NeuroReport*, *5*, 829–832.

Sekuler, A. B., & Palmer, S. E. (1992). Perception of partly occluded objects: A microgenetic analysis. *Journal of Experimental Psychology: General*, *121*, 95–111.

Shepard, R. N., & Judd, S. A. (1976). Perceptual illusions of rotation of three-dimensional objects. *Science*, *191*, 952–954.

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

Shiffrar, M. (1994). When what meets where. *Current Directions in Psychological Science*, *3*, 96–100.

Shimojo, S., & Nakayama, K. (1990). Amodal representation of occluded surfaces: role of invisible stimuli in apparent motion correspondence. *Perception*, *19*, 285–299.

Sigman, E., & Rock, I. (1974). Stroboscopic movement based on perceptual intelligence. *Perception*, *3*, 9–28.

Sinha, P., & Poggio, T. (1996). Role of learning in three-dimensional form perception. *Nature*, *384*, 460–463.

Srinivas, K. (1993). Perceptual specificity in nonverbal priming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 582–602.

Srinivas, K. (1995). Representation of rotated objects in explicit and implicit memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 1019–1036.

Tarr, M. J. (1989). *Orientation dependence in three-dimensional object recognition.* Unpublished doctoral dissertation, Department of Brain and Cognitive Science. Cambridge, MA: MIT press.

Tarr, M. J. (1995). Rotating objects to recognize them: a case study on the role of viewpoint-dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin and Review*, *2*, 55–82.

Tarr, M. J., Bülthoff, H. H., Zabinski, M., & Blanz, V. (1997). To what extent do unique parts influence recognition across changes in viewpoint?. *Psychological Science*, *8*, 282–289.

Tarr, M. J., & Chawarski, M. (1993). *The concurrent encoding of object-based and view-based object representations.* Paper presented at the 34th Annual Meeting of the Psychonomic Society, Washington, DC.

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

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

Tipper, S., Brehaut, J. C., & Driver, J. (1990). Selection of moving and static objects for the control of spatially directed action. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 492–504.

Tse, P. (1999). Complete mergeability and amodal completion. *Acta Psychologica*, 102, 165–201.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. G. Ingle, M. A. Goodale, & R. J. W. Mansfield, *Analysis of Visual Behaviour* (pp. 549–586). Cambridge, MA: MIT Press.

Van Lier, R. (1999). Investigating global effect, in visual occlusion: from a partly occluded square to the back of a tree-formula. *Acta Psychologica*, 102, 203–220.

Van Lier, R., & Wagemans, J. (1998). From images to objects: Global and local completions of self-occluded parts. *Journal of Experimental Psychology: Human Perception & Performance* (in press).

Wagemans, J. (1992). Perceptual use of nonaccidental properties. *Canadian Journal of Psychology*, *46*, 236–279.

Wagemans, J. (1993). Skewed symmetry: a nonaccidental property used to perceive visual forms. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 364–380.

Wagemans, J., Van Gool, L., & Lamote, C. (1996). The visual system's measurement of invariants need not itself be invariant. *Psychological Science*, *7*, 232–236.

Wallis, G. (1996). *Presentation order affects human object recognition learning.* (Technical Report No. 36). Max-Planck-Institut fur biologische Kybernetik.

Watamaniuk, S. N. J., & McKee, P. (1995). Seeing motion behind occluders. *Nature*, *377*, 729–730.

Wertheimer, M. (1923). Untersuchen zur Lehre von der Gestalt, II. *Psychologische Forschung*, *4*, 301–350.

Wraga, M., Creem, S. H., & Proffitt, D.R. (1999). The influence of spatial reference frames on imagined object. and vieweerrotation. *Acta Psychologica*, 102, 247–264.

Yantis, S. (1995). Perceived continuity of occluded visual objects. *Psychological Science*, *6*, 182–186.