Cognition, 30 (1988) 1–35

Scene-based and viewer-centered representations for comparing shapes*

GEOFFREY E. HINTON
University of Toronto

LAWRENCE M. PARSONS
Massachusetts Institute of Technology

Abstract

Five studies examined the use of scene-based and viewer-centered (i.e., head- or retina-based) representations for comparing shapes at different orientations. Two objects that resembled either the line-drawings used by Metzler and Shepard (1974), or helices, were placed far apart on a table top, so that the lines of sight from the viewer to each object were 90 or 150 degrees apart. Observers had to decide whether the objects' shapes were identical or mirror images, and were instructed to physically rotate *one object to an orientation that allowed them to make a decision. They often rotated the object until it had the same relationship to the table top (and room) as the other object (thereby achieving scene-based alignment), even though this produced quite different retinal images of the two objects. Responses regressed up to a third of the way toward viewer-centered alignment as principal surfaces of an object were less aligned with significant directions in the scene. Similar patterns of alignments were observed for pairs of objects with very different surface topology. When subjects were instructed to turn one object so that it was perfectly aligned with the other with respect to the scene, responses also regressed toward the viewer-centered alignment, in this case by as much as 15 degrees (as the unmoved standard object's principal surfaces were less aligned with the scene's significant directions). Overall, these results suggest that when comparing shapes in this kind of task people rely more on scene-based representations than on viewer-centered representations.*

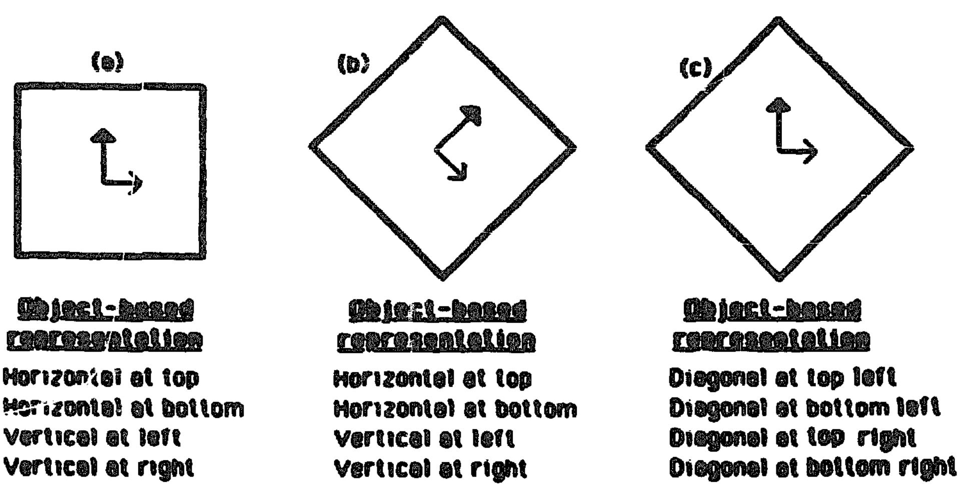
*The first author was supported by grant 87-2-36 from the Alfred P. Sloan Foundation. The second author was supported by a grant from A.P. Sloan Foundation Program in Cognitive Science to MIT Center for Cognitive Science, and by NRSA Fellowship F32 HD6605-03 from National Institute of Health. We thank Martha Farah, Jay McClelland, Steve Pinker, Roger Shepard, and an anonymous reviewer for helpful comments. Requests for reprints should be addressed to: Lawrence M. Parsons, Department of Psychology, Mezes Hall 330, The University of Texas at Austin, Austin, TX 78712, U.S.A.

Introduction

The spatial structure of an object can be encoded in many different ways. We assume, like others, that the human perceptual system represents the positions, orientations, and sizes of an object's constituents relative to some reference frame. Such a frame of reference is often thought to be based or centered on the object (Attneave, 1968; Hinton, 1981a; Hinton & Parsons, 1981; Mach, 1897; Marr & Nishihara, 1978; Palmer, 1975, 1977; Pinker, 1985; Rock, 1973).[1] One disadvantage of using object-based frames of reference to achieve canonical shape representations is that some objects have several different frames, and so they receive different shape representations depending on which frame is selected. A square that is tilted at 45 degrees, for example, can be seen either as a tilted square, or as an upright diamond depending on which object-based frame is used (Figure 1). The fact that people have two quite different ways of seeing this object is evidence in favor of object-based frames.

Different choices of a reference frame yield encodings with very different properties. Early visual processes appear to use a frame defined by the retina: The shape of an object is implicitly encoded by activity in neurons each of

Figure 1. *(a) A square in its standard orientation with its standard representation. (b) A tilted square represented relative to an approximately tilted object-based frame of reference. (c) An alternative representation obtained by using a different object-based frame.*



[1]An alternative, but less common, hypothesis is that spatial structure is encoded by complex relational features such as "contains three points that form a triangle with angles of 30, 60, 90 degrees" or contains "5 line endings."

which represents the conjunction of a type of feature (e.g., a stopped, oriented edge) with a position, orientation, and size relative to the retina. We call this a "retina-based" representation because of the frame that is uses, not because it uses retinal neurons, or because they form a retino-topic map in which nearby neurons represent features at nearby retina-based positions.

The retina-based representation of an object is the easiest to extract from an image, but unfortunately it has the disadvantage that it changes with every change in viewpoint. Retina-based and object-based frames are both useful but they are not sufficient for some of the tasks the visual system must typically perform. At least one other, qualitatively different type of reference frame is probably required.

The visual system must not only perceive the shapes of single objects: It must perceive the spatial relationships between objects, some of which may not be simultaneously visible. When two or more objects are seen simultaneously, it is conceivable that we can represent their spatial relationship as the shape of the configuration they form, so an object-based frame for the larger configuration could be used for encoding the spatial relationships between the objects. But when several objects are perceived sequentially (when looking through a moving peephole, for example) the object-based frame for the whole configuration cannot be known at the time that the first object is encoded. (For relevant discussions of sequential viewing, see Cattell, 1900; Girgus, Gellman, & Hochberg, 1980; Hochberg, 1968, 1982; Morgan, Findlay, & Watt, 1982; Rock, 1981; Shimojo & Richards, 1986.) An object-based reference frame based on the first object alone may be used to encode its intrinsic shape, but some other reference frame must be used to encode the first object's position, orientation, and size in a way that will allow the visual system later to perceive the whole configuration of sequentially perceived objects. The visual system needs to store where each object is in space, but achieving such a representation could mean a variety of things.

In the perception of a sequence of objects, the stored encoding of one object's disposition (i.e., its position, orientation, and size) cannot be relative to the object-based frame of the whole configuration because that cannot yet be determined, so the obvious remaining possibilities are as follows.

(1) The dispositions of objects (those in the scene but not necessarily receiving direct visual attention) are stored relative to the current retina-based frame. Whenever the eye moves the dispositions of all objects in the scene already perceived are updated so that they are correct for the new retina-based frame defined by the new eye-position. This requires a lot of active computation after each movement of the eye, head, or body, and it allows small errors to accumulate after each movement.

(2) The dispositions of objects are stored relative to a head-based frame that is stable across eye-movements. This has the same qualitative disadvantages as the retina-based frame but at least updating is not required when the change in eye position is caused solely by a change in the position of the eye in its orbit. One big advantage of using a head-based frame is that the head contains the vestibular apparatus, so integrating visual and vestibular information is easiest if the visual information is head-based. However, when the head is tilted, the head-based positions and orientations of all recently perceived objects must be changed. Of course, it is conceivable that people use a "mixed" frame of reference that is not based on any single object: the origin and scale of the frame could be defined by the head, but the vertical could be defined by gravity. Thus, the direction of gravity could define two of the three degrees of orientational freedom and the last could be defined by projecting the front-back direction of the head down into the horizontal plane (the head need not be vertical) (see Parsons & Shimojo, 1987, 1988).

(3) Extending the logic that led to a head-based frame, we could postulate a body-based frame that is stable across head movements (Parsons & Shimojo, 1987, 1988). It is hard to be precise about the body-based frame because, unlike the head, the body is not a rigid object. When one moves a shoulder forwards, or bends the spine a little more, does this change the body-based frame?

(4) Finally, the disposition of an object could be stored relative to some larger, "scene-based" frame of reference that is based on some salient object in the scene such as a room, table-top, blackboard, or page (Attneave, 1972; Attneave & Farrar, 1977; Attneave & Pierce, 1978; Biederman, 1981). The disposition of the current object of perception, O, is simply the relationship R_{os} between the scene-based frame (S) and the current object-based frame. The object that is used to define the scene-based frame is not necessarily the object that contains O as an immediate constituent. The advantage of using a scene-based frame is that it leaves the observer free to move around without having to update the dispositions of every object in his or her environment. (Instead, as the observer moves around, he or she will need to update the disposition of any object that has moved, such as his or her own body (Bridgeman, Lewis, Heit, & Nagle, 1979; Noda & Warabi, 1986; Thomson, 1980, 1983), and refer any newly-perceived objects to the scene-based frame.) The relationships of objects to appropriate scene-based frames can also be stored in long-term memory because they do not change unless the object itself moves relative to the scene.[2]

We expect there to be different groups of neurons for each different reference frame that the perceptual system actually uses. Primary visual cortex is retina-based (e.g., Hubel & Wiesel, 1962, 1974; Van Essen, 1985), infero-temporal cortex is probably object-based (e.g., Gross, Rocha-Miranda, & Bender, 1972; Ungerleider & Mishkin, 1982), and there is evidence that the parietal cortex (e.g., Andersen, Essick, & Siegel, 1985; Lynch, 1980; Mishkin, Ungerleider, & Macko, 1983) and Purkinje cells in the cerebellar flocculus may be scene-based (Noda & Warabi, 1986). Besides groups of neurons that encode spatial structure relative to various different frames, we also expect there to be neural apparatus for mapping the information from one frame to another (Ballard, 1986; Hinton, 1981a). For example, a neuron whose receptive field is retina-based but whose activity is gated by the position of the eye in its orbit (or by the position of the fixation point in the scene) could well be involved in mapping between retina-based and head-based (or scene-based) frames. Andersen, Essick, and Siegel (1984) have found such neurons in area 7a of the macaque.

Ultimately, neurophysiological and neuropsychological research may be the most effective way of untangling the reference frames used by monkeys and humans. In the meantime, psychological experiments can provide insight into which frames may be involved in various tasks. For example, in a typical "mental rotation" task (Shepard & Metzler, 1971), observers decide whether two objects are identical by imagining one object rotated into alignment with the other. During mental rotation, the subject presumably changes what aspects of the spatial structure of the rotated object are represented, or changes the position, orientation, and size at which the object is represented (or changes both). It is possible to ask which reference frames are being used without deciding whether the representation that is changing just represents disposition (as suggested by Hinton and Parsons, 1981) or whether it also encodes the object's spatial structure (as assumed by most investigators). We designed our experiments to distinguish between two possibilities that cannot be discriminated in the usual mental rotation tasks:

(1) The internal representation that is changing is relative to a scene-based frame.
(2) The internal representation that is changing is a representation relative to a retina-based or head-based frame. For the sake of brevity, we refer to both these frames as "viewer-centered."

[2]When the relation of the viewer to the scene is very rapidly changing, it is conceivable that the overhead of updating the scene-based frame would make it more practical to use only two frames, the object-based and viewer-based frames.

The second possibility seems plausible because the internal representations that result from different glances at a scene, or from tasks such as mental rotation, seem to be associated with viewpoints (Carpenter & Just, 1978; Keenan & Moore, 1979; Kosslyn, 1980; Metzler & Shepard, 1974; Pinker, 1980; Pinker & Finke, 1980). This hypothesis is consistent with the idea that the objects are internally represented from viewpoints centered on the observer. (Such viewer-centered representations could be continually refreshed from underlying object-based structural descriptions.) This possibility is typically assumed by those working in the mental rotation paradigm.

The experimental paradigm

Our subjects discriminated between pairs of identical or mirror reflected objects under naturalistic viewing conditions. In the test condition, the objects, which were unfamiliar and abstract, were separated by 90 (or 150) degrees of visual angle, at the same distance to the subject's right and left. Both objects were sitting on the same rectangular table, so the local context of one object was continuous with that of the other. The objects differed in orientation by a rotation about each object's major axis, which was aligned with the vertical. We instructed our subjects to rotate one of the objects physically, rather than merely mentally, if a rotation was needed in order to determine whether the two objects were identical or mirror images of each other.

We can preview our findings as follows. Subjects in Experiment 1 spontaneously turned one object beyond viewer-centered alignment with the other object, to an orientation that was near scene-based alignment but that regressed slightly in the direction of viewer-centered alignment. In Experiment 2, the more the unmoved object's principal surfaces were aligned with the significant directions in the scene, the less subjects' alignments regressed away from scene-based alignment toward viewer-centered alignment. In Experiment 3, subjects were instructed to turn the comparison object so that it was at the same exact orientation as the standard with respect to the local environment. They set the object to an orientation that regressed slightly away from scene-based alignment toward viewer-centered alignment. This regression was as much as 15 degrees when there was poor alignment between the principal surfaces of the unmoved object and the scene's significant directions. In Experiment 4, we controlled for, and eliminated, the hypothesis that the compromise alignments used to compare shapes were not due to the use of scene-based frames but were due to subjects turning their bodies slightly toward each object to perceive each object with respect to a body-based frame of reference. In Experiment 5, we found the same pattern of alignments when subjects were comparing objects with quite different kinds of shapes.

Experiment 1: Scene-based and viewer-centered representations of shape

In Experiment 1, subjects had to discriminate between pairs of identical or mirror-reflected wooden replicas of the objects depicted in line drawings in Metzler (1973) and Metzler and Shepard (1974) (see Figure 2 and Shepard & Cooper, 1982, p. 19 ff.). In the test condition, the objects were separated by 90 degrees of visual angle, at the same distance to the subject's right and left (Figure 3).

There are four obvious plausible outcomes. First, subjects could discriminate the objects' shapes without physically (or mentally) reorienting the objects. This would be possible if the perceptual system described each object with respect to its own, object-based frame of reference, and compared the resulting descriptions. Previous work on the discrimination of identical and mirror image shapes indicates that subjects are unlikely to do this (for the real cube-figures used here, Kaushall & Parsons, 1981; for various other stimuli: Cooper, 1975; Cooper & Shepard, 1973; Parsons, 1987a, b, c; Shepard & Hurwitz, 1984). Appendix 1 discusses some reasons why object-based frames may be insufficient for this task.

Second, subjects could align the objects with respect to the viewer-centered frame, so that they see identical views of each object (when the objects are identical in shape). When a pair of identical objects are aligned with respect to this frame, corresponding parts of the two objects point in directions 90 degrees apart (relative to the table top).

Third, subjects may align the objects so that there is an identical spatial relationship of each object's parts to the scene (e.g., to salient directions of the table and room). Because the objects are widely separated, different views (and retinal images) of the two objects are produced when the objects are aligned with respect to this scene-based frame.

Figure 2. *Line-drawing illustrations of the two experimental stimuli.*

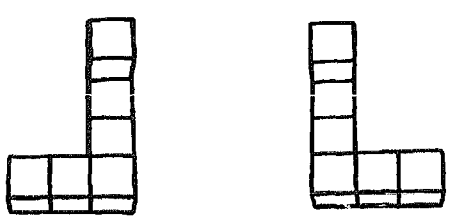
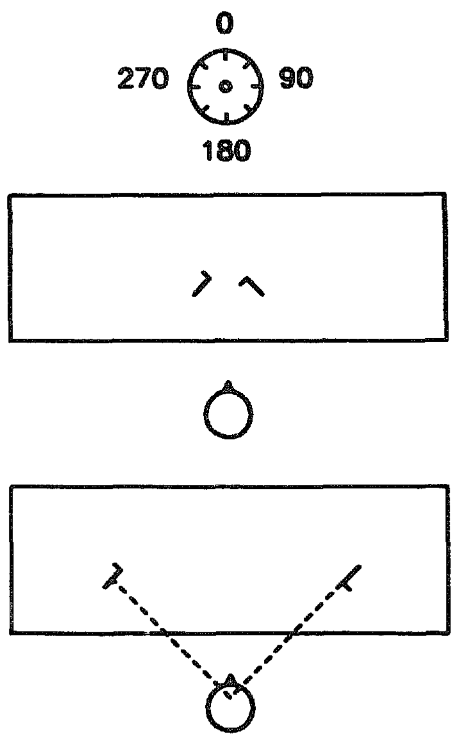


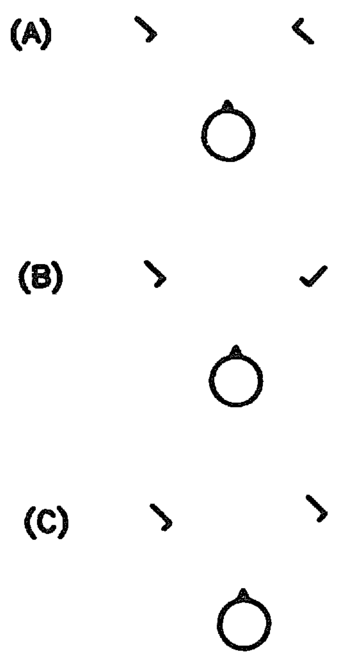
Figure 3. *View from above of the experimental setting in Experiments 1, 2, 3, and 5, and the 90-Degree Separation task in Experiment 4. At the top of the page are the compass directions used to orient the stimuli. The direction in which the front of the subject's body faced was 0 degrees. The middle and bottom panels show the relationship of the subject to the stimuli when the objects were either close together or far apart. An object's two end segments are depicted as a long line (for the segment in contact with the table top) and a short line (for the top segment in Figure 2).*



Fourth, subjects could use an alignment that compromises between the viewer-centered and scene-based frames. Such compromise effects have been found in studies of perception in other paradigms by Attneave (1972), Corballis, Nagourney, Shetzer, and Stefanatos (1978), Gilinsky (1955), Rock (1973), and Uhlarik, Pringle, Jordan, and Misceo (1980).

In Experiment 1, the standard object was at either a 225 or 315 degree orientation (Figure 3), and the comparison object was initially either 180 or 90 degrees counterclockwise away from the orientation of the standard. If observers preferred to use scene-based representations of the object in order to compare them, then presumably the comparison object would be set to 225 and 315 degree orientations respectively. If their preference was to use viewer-centered representations, then the comparison object would be set to

Figure 4. *An illustration of viewer-centered and scene-based alignments when observer's lines of sight to the objects are 90 degrees apart. (a) The initial presentation of stimuli on a trial (standard on left at 315 degrees, comparison on right at 135 degrees); (b) viewer-centered alignment; and (c) scene-based alignment. (See Figure 3 and the text.)*



an orientation 90 degrees clockwise from the orientation of the standard, i.e., at the 315 or 405 degree orientations respectively. Responses between 225 and 315 degrees when the standard was at 225 degrees, or between 315 and 405 when the standard was at 315 degrees, would be "compromises" between the scene-based and viewer-centered alignments. So, for example, if the standard was pointing "southwest", then scene-based alignment of the comparison object was "southwest", viewer-centered alignment was "northwest", and "west" was the exact compromise between the scene-based and viewer-centered alignments.

Method

Subjects

Twelve University of California at San Diego (U.C.S.D.) undergraduates, who had not been in any related experiments, participated for $5.50 an hour.

Stimuli

The stimuli were two identical wooden replicas of each of the two shapes in Figure 2. They subtended 8 to 9.5 degrees of the subject's visual angle; the height of the major axis of each cube-figure was 6.8 cm. The objects were shown with their central segment vertically aligned, and their bottom segment resting on the table top in front of the subject. The stimuli were presented at different orientations about the vertical axis. Figure 3 illustrates the compass directions by which an object's orientation was designated (for the experimenter only). The 0, 90, 180, and 270 degree compass directions were aligned with the significant directions in the room in which the subject performed the experiment. On two of the four test trials with an identical pair, the bottom segment of the standard object pointed toward 225 degrees, and on the other two such trials, the bottom segment pointed toward 315 degrees. Identical pairs of objects differed in orientation by 90 or 180 degrees. (The orientation difference between objects with mirror image shapes is undefined because there is no orientation where all corresponding parts of the shapes are aligned.) On all other trials, the bottom segment of an object pointed toward either 45, 135, 225, or 315 degrees.

Design

All subjects performed the same set of 12 practice and 8 test trials in the same random order. For each response and stimulus condition, there was an equal number of trials with identical and mirror image pairs, and with each of the two shapes in Figure 2. The "standard" object was leftmost for the subject, and the "comparison" object was rightmost. The two objects were either close together directly in front of the subject or far apart (one to the subject's left and the other to the right, see Figure 3). Each of the two shapes was equally often the standard and comparison object. On half of the trials with an identical pair, the difference in orientation of the two objects was 90 degrees; on the other half, the difference in orientation was 180 degrees. The comparison object was never presented at an orientation where it would be perfectly aligned with the standard object with respect to either the viewer-centered or scene-based frames.

Procedure

The task was to judge whether the two objects presented on a trial were identical or were mirror images of one another. In two sets of trials, subjects were instructed to make their judgment by imagining the comparison object at the orientation of the standard, without physical manipulation of the objects. In two other sets of trials, they were instructed to make their judgment by physically rotating the comparison object to an orientation that allowed them to make their judgments. In the latter cases, they were instructed to physically rotate the comparison object as soon as possible without reasoning about the objects' shapes, and to turn the object about its vertical axis without lifting it off the table. After either mentally or physically rotating the comparison object, subjects indicated their judgments by a verbal response (i.e., "same" or "mirror"). Subjects were instructed to view the objects by rotating their head about their vertically-aligned neck. Trials on which subjects moved their head in other ways were repeated later in a session until performed correctly.

To start the session, subjects performed eight practice trials with the standard and comparison objects directly in front of them (Figure 3). On the first four practice trials, subjects made their judgment by imagining one object at the orientation of the other. On the second four practice trials, they physically rotated the comparison object (the object to their right) about the vertical axis to an orientation that allowed them to judge whether the objects were identical or mirror images. In the last four practice trials, the standard and comparison objects were far apart (Figure 3), and subjects made their judgments by mentally rotating the objects. Finally, subjects performed eight test trials with the objects far apart and physically rotated the comparison object about its vertical axis.

Results

Subjects' discrimination of identical and mirror image pairs of objects was incorrect on less than .5 percent of the test trials. They rotated the comparison object to orientations (Table 1) that were reliably different from alignment with either the scene-based or viewer-centered frames (two-tailed t tests, $df = 47$, $p < .001$, $t = 6.81$ and $t = 11.14$, for 225 and 315 degree conditions). The preferred orientations were on average *15* degrees from alignment with the standard object with respect to the scene-based frame of reference, and *75* degrees from alignment with respect to the viewer-centered frame of reference. Every subject on every trial either turned the comparison object to near scene-based alignment or to an alignment no more than 40 degrees away from scene-based alignment toward viewer-centered alignment. The extent of compromise varied across individuals (overall compromise for each was: 38.25, 37.75, 26, 26, 22.5, 21.75, 11, 9.75, 9.25, 8.5, 8.25).

Table 1. *Experiment 1: Orientation to which subjects rotated the identical comparison object*

Direction of subjects' rotation	Orientation of standard					
	225 degrees			315 degrees		
	Response	*SD*	Frequency	Response	*SD*	Frequency
Clockwise	237.88	9.11	16	328.78	10.06	18
Counterclockwise	243.88	21.98	8	334.33	5.06	6

Note. Based on 12 subjects and a total of 48 observations. Degrees refer to compass directions in Figure 3.

Clockwise and counterclockwise rotations

It is possible that the orientation to which subjects rotated the comparison object influenced, or was influenced by, whether they turned it in a clockwise or counterclockwise direction. For example, the direction of their rotations may have been influenced by incidental features of their grasp of the objects or by constraints on motion of the joints of the hand and arm. However, when identical objects were 180 degrees apart, *86* percent of the time subjects rotated the comparison object *past* viewer-centered alignment and through the shortest angle to their final orientation (an orientation that nearly aligned the objects with respect to the scene-based frame). That subjects chose to pass the object through the shortest angle to its final orientation suggests that at the *start* of the rotation they knew (within at most 15 degrees) the final orientation.[3] This indicates that the direction of rotation was determined by the final orientation (and not vice versa).

Discussion

To compare the shapes of objects at different orientations and positions, subjects in Experiment 1 relied primarily on scene-based representations. However, they did not rotate the comparison object to exact scene-based

[3]Because the comparison object was never presented in perfect alignment with the viewer-centered frame, when identical objects differed in orientation by 90 degrees, a *clockwise* rotation was always the shortest angle to turn the object to achieve alignment in the scene-based frame. As expected, subjects chose the shortest angle of rotation 83 percent of the time in these cases.

alignment but rotated it to a *compromise* orientation between scene-based and viewer-centered alignment. Interestingly, they often rotated the comparison object past the viewer-centered alignment, toward scene-based alignment. This compromise alignment is not an artifact of averaging across subjects, absolute orientations, or trials; on every trial each subject turned the comparison object either to near scene-based alignment or to an alignment no more than 40 degrees away from scene-based alignment toward viewer-centered alignment. (However, there were individual differences in the extent of the compromise alignments.) It is not clear why subjects use these compromises. They may better allow the observer to decide that the objects have the same shape or are at the same orientation (see General Discussion). Each of four following experiments confirmed the results in Experiment 1 and investigated a different implication of the use of these compromise—but predominantly scene-based—alignments for comparing shape.

Experiment 2: Spatial relationship of an object to the significant directions in a scene

In Experiment 2, we examined how the spatial relationship between (a) the significant directions in a scene and (b) the principal axes (or surfaces) of the object influences the magnitude of the deviation from scene-based alignment. We varied the extent to which the object's principal surfaces were aligned with the significant directions in the scene (e.g., the edges of the table and the walls of the room, which were in alignment), and predicted that when the object's principal surfaces were aligned with these significant directions (i.e., when the spatial relations of the object and scene are "salient"), observers would place the comparison object at an orientation nearer to perfect scene-based alignment. In the Parallel condition, the segments of the standard object pointed in directions parallel to the scene's significant directions. In the Nearly-Parallel condition, the standard object's segments were 20 degrees away from alignment with one of the significant directions of the scene (e.g., a wall of the room). In the Skewed condition, the standard object's segments were at a 45 degree angle to the significant directions of the scene.

Method

Subjects

Eleven U.C.S.D. undergraduates, who had not been in any related experiments, participated for $5.50 an hour.

Stimuli

The stimuli were identical to those in Experiment 1. With respect to the compass directions in Figure 3, there were three sets of orientations of the standard object. In the Parallel condition, the bottom segment of the standard object pointed toward 0, 90, 180, or 270 degrees (Figure 3). In the Nearly-Parallel condition, the standard's bottom segment pointed toward 70, 160, 200, or 340 degrees. In the Skewed conditions, the bottom segment of the standard object pointed toward 45, 135, 225, and 315 degrees.

Design and procedure

All subjects performed the same set of 18 practice and 24 test trials in the same random order. They first completed 6 practice trials with both objects directly in front of them: on 3 trials they were instructed to rotate the shapes mentally, and on 3 trials they were instructed to physically rotate one of the objects. Next they performed 12 practice trials with the objects far apart (Figure 3), making their judgments by mentally rotating one of the objects. Finally, subjects completed 24 test trials by physically rotating the comparison object. All other aspects of the design, procedure, task, and instructions were identical to those in Experiment 1.

Results

In the Parallel condition, subjects rotated the comparison object to very nearly perfect alignment with the standard object with respect to the scene-based frame. However, as the standard object was less closely aligned with the significant directions of the subjects' local environment, the final orientation of the comparison object regressed slightly towards the viewer-centered alignment (Table 2). The largest regression towards viewer-centered alignment, observed in the Skewed condition, was 26 percent of the total angle between scene-based and viewer-centered alignments.

Analyses of variance (ANOVAs) were conducted of the angle between the orientation to which subjects rotated the comparison object and the orientation of the standard object. Separate ANOVAs for each condition showed no effect of absolute angle of the standard object, and another ANOVA, ignoring that factor, showed a reliable difference between the Parallel, Nearly-Parallel, and Skewed conditions ($F(2,20) = 33.59$, $p < .001$). As in Experiment 1, within these main trends there was some variation across individuals in the Nearly-Parallel and Skewed conditions. (With 0 degrees indicat-

Table 2. *Experiment 2: Mean and standard deviation of orientation to which subjects rotated identical comparison object*

					Mean angular difference
Parallel	0	90	180	270	
	1.0 (2.52)	91.7 (2.19)	179.9 (4.05)	269.6 (2.90)	.75 (2.13)
Nearly-Parallel	70	160	200	340	
	80. (7.32)	175.6 (8.16)	215.6 (23.43)	357.4 (10.45)	14.95 (11.63)
Skewed	45	135	225	315	
	67.91 (26.31)	159.73 (7.71)	248.09 (19.71)	336.46 (16.03)	23.05 (14.59)

Note. Based on 11 subjects. Degrees refer to the compass directions in Figure 3.

ing scene-based alignment and 90 degrees indicating viewer-centered alignment, the overall means for each individual were: for Nearly-Parallel: 9, 6, 12, 5.5, 16.25, 23.5, 7.75, 25.75, 12, 13, 28.25; for Skewed: 29, 8.25, -2.25, 13.25, 26.75, 30.25, 15.75, 19.5, 54, 27.75, 31.25.) The error rate for the discrimination of mirror image and identical pairs of objects was less than .5 percent.

Clockwise and counterclockwise rotations

Subjects often rotated the comparison object beyond the orientation aligning it with the standard object with respect to the viewer-centered frame and toward the orientation that nearly aligned it with the standard object with respect to the scene-based frame. However, the tendency was less pronounced here than in Experiment 1, occurring only on the six trials when the orientation difference between identical objects was 180 degrees. It occurred about a third of the time for trials in the Parallel and Nearly-Parallel condition, and about a half of the time for trials in the Skewed condition.

Discussion

In comparing shapes of objects at different orientations and positions, observers in Experiment 2 relied for the most part on scene-based rather than viewer-centered representations. However, they used a partial compromise (of up to 26 percent) between the scene-based and viewer-centered frames as the principal surfaces of the object were less aligned with the significant directions in a scene. This result may be related to the possibility that the spatial relations between the scene and the objects' parts become less "salient" as the object is less aligned with the significant directions in the scene. The increasing compromise could have provided views of the objects that were increasingly informative for comparing shape or in deciding that the two objects were at the same orientation. These views presented corresponding aspects of the objects, though not from exactly the same perspective.

This shows the effect of the principal directions of large surfaces in a scene, and is reminiscent of various other environmental or contextual effects on the perception of spatial relations and shape (e.g., Bauermeister, 1964; Day & Wade, 1969; Howard, 1986; Robinson, 1972; Rock, 1973; Sedgewick, 1986). Possibly analogous compromise alignments have been observed in the assignment of up-down directions under some conditions. These compromises were among gravitational upright, retinal upright, and conspicuous directions defined by parallelism, pointing, and bilateral symmetry (e.g., Attneave, 1968; Attneave & Reid, 1968; Attneave & Olson, 1967; Palmer & Bucher, 1981; Rock, 1973). For example, Corballis and his coworkers have used head tilt to investigate the relative influence of retinal and environmental (i.e., gravitational) frames of reference on discrimination of identical or mirror-reversed letters, numbers, and dot patterns. The results often vary for different tasks and stimuli (Corballis et al., 1978). Consistent with our results, RT-orientation functions show that when discriminating correct and reversed letters and numbers, subjects use a representation compromising between retinal and environmental/gravitational frames, but close to the environmental/gravitational frame (Corballis, Zbrodoff, & Roldan, 1976). Other perceptual tasks that seem to be performed with respect to an environmental/gravitational frame are speeded identification of line slopes (Attneave & Olson, 1967), tachistoscopic identification of letters (Corballis, Anuza, & Blake, 1978), and memory for novel figures (Rock & Heimer, 1957). However, perceptual grouping (Olson & Attneave, 1970) and the judgment of the symmetry of dot patterns (Corballis & Roldan, 1975; Corballis et al., 1976) appear to be performed in a retinal frame. It is not clear why subjects rely to varying extents on different frames of reference for the different kind of tasks.

Experiment 3: Subjective and objective scene-based alignment

Results in Experiment 3 confirmed our belief that observers in Experiments 1 and 2 were relying more on scene-based than viewer-centered representations. We assessed the extent to which subjective perception of perfect alignment of the objects with respect to the scene-based frame, varied systematically from objective perfect alignment. In the "Shape Comparison" task, subjects performed a replication of Experiment 2. In the "Scene-Based Alignment" task, they performed trials in which they were instructed to turn one object so that it was at exactly the same orientation *with respect to the local environment* as another object (with the same shape).

Method

Subjects

Nine MIT undergraduates, who had not been in any related experiments, participated for $5.50 an hour.

Stimuli, design, and procedure

The stimuli were identical to those in Experiments 1 and 2, and were presented exactly as in Experiment 2. First, in the Shape-Comparison task, subjects performed a replication of Experiment 2; then, in the Scene-Based Alignment task, they performed only the trials in Experiment 2 that used identical pairs of objects. In Scene-Based Alignment task, subjects were instructed "to turn the object on your right so that it is at exactly the same orientation with respect to the local environment (that is, with respect to the room and table, etc.) as the object on your left." All other features of the design and procedure in the Scene-Based Alignment task were identical to those in Experiment 2.

Results

In the Scene-Based Alignment task, as the standard object was less closely aligned with the significant directions of the subjects' local environment, the perceived orientation of perfect scene-based alignment regressed slightly towards the viewer-centered alignment (Table 4). The regression of perceived scene-based alignment towards the viewer-centered alignment was about 13 degrees whereas the regression of the final orientation in subjects' replication of Experiment 2 (the Shape-Comparison task) was 26 degrees.

Data in the Shape-Comparison task (Table 3) were highly correlated with those in Experiment 2 ($r = .998$, $F(1,10) = 2248.72$, $p < .0001$). The error

Table 3. *Experiment 3: Shape-comparison, replication of Experiment 2. Mean and standard deviation of orientation to which subjects rotated identical comparison object*

					Mean angular difference
Parallel	0	90	180	270	
	0.33 (2.40)	90.67 (2.11)	181.11 (2.33)	271.22 (2.82)	0.83 (2.31)
Nearly-Parallel	70	160	200	340	
	75.33 (7.72)	188.56 (25.24)	238.22 (11.25)	363.22 (20.13)	23.83 (21.88)
Skewed	45	135	225	315	
	74.22 (28.83)	167.56 (14.25)	245.89 (12.98)	352.67 (21.27)	30.09 (21.24)

Note. Based on 9 subjects. Degrees refer to the compass directions in Figure 3.

rate for the discrimination of mirror image and identical object pairs was less than .5 percent.

Separate ANOVAs conducted for the Shape-Comparison and Scene-Based Alignment tasks (ignoring absolute orientation within each condition) showed reliable differences between the Parallel, Nearly-Parallel, and Skewed conditions ($F(2,16) = 27.51, p < .001$, and $F(2,16) = 12.57, p < .001$). Separate ANOVAs on each condition in each task showed a reliable effect of absolute orientations only in the Nearly-Parallel condition in the Shape-Comparison task ($F(3,24) = 5.10, p < .01$). This latter effect was caused by responses to the 70 degree orientation that were relatively close to scene-based alignment and by responses to the 200 degree orientation that were relatively far from scene-based alignment (see Footnote 3). An ANOVA (ignoring absolute orientation) on data in the Skewed and Nearly-Parallel conditions in the Shape-Comparison and Scene-Based Alignment tasks, showed that the final orientations in the Skewed and Nearly-Parallel conditions were different for the two tasks ($F(1,8) = 14.09, p < .01$).

Discussion

The results of Experiment 3 show that the *perceived* orientation of perfect scene-based alignment of the objects in this task is often different from *objec-*

Table 4. *Experiment 3: Scene-based alignment, perceived perfect scene-based alignment. Mean and standard deviation of orientation to which subjects rotated identical comparison object*

					Mean angular difference
Parallel	**0**	**90**	**180**	**270**	
	0.11 (2.13)	**90.11 (2.18)**	**180.67 (2.31)**	**270.79 (2.53)**	**0.42 (.7)**
Nearly-Parallel	**70**	**160**	**200**	**340**	
	74.00 (8.98)	**169.11 (19.33)**	**223.00 (13.15)**	**344.44 (22.11)**	**10.18 (18.05)**
Skewed	**45**	**135**	**225**	**315**	
	58.11 (8.86)	**151.00 (13.34)**	**239.89 (12.66)**	**333.56 (8.47)**	**15.64 (11.07)**

Note. Based on 9 subjects. Degrees refer to the compass directions in Figure 3.

tive perfect alignment. This suggests that there is a perceptual illusion in this situation. As the alignment of the objects with significant directions in the scene became less obvious, perceived perfect alignment regressed slightly toward viewer-centered alignment. The latter regression was half that observed when subjects aligned the two objects to compare their shapes. Thus, the regression away from objective scene-based alignment found in Experiment 2 and in the Shape-Comparison task in Experiment 3 seems due equally to two effects: (i) a change in the perceived orientation of perfect scene-based alignment and (ii) a preference for using compromise alignments when comparing two objects' shapes. These results suggest that in comparing shape subjects rely on representations based predominantly on a scene-based frame, even when principal surfaces of the objects are not aligned with the significant directions in the scene.

Experiment 4: Scene-based and body-based representations of shape

The compromise alignments in the preceding experiments could have been produced if, contrary to instruction, subjects turned or spread their shoulders appropriately to use a strategy of referring the shape to a body-based frame of reference. Suppose subjects examined each object by first turning their shoulders slightly about the vertical so that the front of their bodies faced slightly more in the direction of the inspected object. If subjects wanted to compare shapes with respect to the frame of reference of their bodies, they would have to compensate for the difference in orientations of their bodies as each object was inspected. This requires setting the comparison object to an orientation that regresses away from scene-based alignment toward viewer-centered alignment by the sum of the angles that subjects' shoulders were turned toward each object.

This possibility was controlled for by having people perform this task when their shoulders and upper trunk were immobilized and their lines of sight to the objects formed a 150 rather than 90 degrees angle, as in Experiments 1–3. Subjects performed the task in Experiment 2 under two conditions: when their lines of sight to each object formed a 150 degree angle and when they formed a 90 degree angle.

We predicted that subjects would perform the 90-Degree and 150-Degree Separation tasks in the same way: by relying primarily on scene-based frames to compare the shape of the objects, but using alignments that compromised slightly away from scene-based alignment toward viewer-centered alignment, as the principal surfaces of the objects were less aligned with the scene's significant directions. For observers to compromise in the same way between viewer-centered and scene-based alignments in the two tasks they must compensate for the difference in object separation in the two tasks. For a given spatial relationship of the two objects to the significant directions in the scene, observers will see *different* aspects of the objects in the 90-Degree Separation task than in the 150-Degree Separation task. Of course, the orientation of the objects with respect to the scene in the Parallel, Nearly-Parallel, and Skewed conditions is not changed by whether the lines of sight to the objects are 90 or 150 degrees apart. We can make the following predictions assuming observers compromise in the same way between viewer-centered and scene-based alignments in the two tasks.

In the Nearly-Parallel condition of the 150-Degree Separation task, if the comparison object is turned 5 degrees away from scene-based alignment in the direction of viewer-centered alignment, it is at an orientation (with respect to the observer) similar to that in the Skewed condition of the 90-Degree Separation Task. To produce the same relationship of views of the objects as in the 90-Degree Separation task, the regression of the final orientation in the Nearly-Parallel condition of the 150-Degree Separation task would be about 31 degrees: the initial 5 degrees plus about 26 degrees, the average amount of regression found in the Skewed condition in Experiment 2 and the Shape-Comparison task in Experiment 3.

In the Skewed condition of the 150-Degree Separation task, the comparison object must be more than 30 degrees away from scene-based alignment toward viewer-centered alignment to begin to reveal surfaces like those visible in the standard object. However, when the comparison object is 30 degrees away from scene-based alignment, the observer's view of the object is like that in the Skewed condition in the 90-Degree Separation task. Therefore, we predicted that the regression of the final orientation in the Skewed condition in the 150-Degree Separation task would be about 56 degrees: the initial 30 degrees plus 26 degrees, the average amount of regression found in the Skew condition in Experiment 2 and the Shape-Comparison task in Experiment 3.

Method

Subjects

Six MIT undergraduates, who had not been in any related experiments, participated for $5.50 an hour.

Stimuli, design, and procedure

The stimuli were identical to those in Experiments 1 and 2, and were presented differently in the two tasks. In the 90-Degree Separation task, subjects performed a replication of Experiment 2 (i.e., the observer's line of sight to each object formed a 90 degree angle). In the 150-Degree Separation task, they performed another replication of Experiment 2 but with the line of sight forming a 150 degree angle. In addition, in the 150-Degree Separation task, subjects' shoulders and upper body were bound to an immobile high-backed chair. All other features of the design and procedure in the 150-Degree Separation task were identical to those in Experiment 2. Half the subjects performed the 90-Degree Separation task first, then the 150-Degree Separation task; the other half performed the 150-Degree Separation task, then the 90-Degree Separation task.

Results

There was no reliable difference between the performances of subjects doing the tasks in different orders (tested by ANOVAs both with and without the factor of absolute orientation within condition, analogous to those described later, $p > .05$). The order factor was ignored in other analyses.

Data in 90-Degree Separation task (Table 5) were highly correlated with

Table 5. *Experiment 4: 90-Degree Separation, observer's lines of sight to two objects form a 90 degree angle. Mean and standard deviation of orientation to which subjects rotated identical comparison object*

					Mean angular difference
Parallel	0	90	180	270	
	1.33 (1.97)	91.37 (.94)	181.0 (1.29)	271.67 (2.38)	1.33 (1.54)
Nearly-Parallel	70	160	200	340	
	88.33 (17.1)	181.18 (10.96)	229.17 (19.66)	356.5 (14.23)	21.29 (14.80)
Skewed	45	135	225	315	
	79.5 (18.70)	162.83 (16.77)	235.17 (16.00)	349.33 (8.84)	26.71 (14.95)

Note. Based on 6 subjects. Degrees refer to the compass directions in Figure 3.

those in Experiment 2 ($r = .998$, $F(1,10) = 2187.81$, $p < .0001$) and in the Shape-Comparison task in Experiment 3 ($r = .999$, $F(1,10) = 4208.58$, $p < .0001$). The error rate for discriminating mirror image and identical object pairs in each task was less than .5 percent. Separate ANOVAs conducted for 90-Degree and 150-Degree Separation tasks (ignoring absolute orientations within conditions) showed reliable differences among the Parallel, Nearly-Parallel, and Skewed conditions ($F(2,10) = 22.86$, $p < .0001$, and $F(2,10) = 29.17$, $p < .0001$). Separate ANOVAs on each condition in each task showed a reliable effect of absolute orientation only in the Skewed condition of the 90-Degree Separation task ($F(3,15) = 4.98$, $p < .01$). This latter effect was caused by responses to the 225 degree orientation that were relatively close to the scene-based alignment (see Footnote 3). Another ANOVA (ignoring absolute orientation) showed that the final orientations in both the Skewed and Nearly-Parallel conditions were different for 90-Degree and 150-Degree Separation tasks ($F(1,5) = 25.23$, $p < .01$).

When the angle between the lines of sight to each object was 150 degrees, the observed final orientations were not reliably different (two-tailed t test, $p > .05$) from the predictions described earlier. These predictions are based on the assumption that subjects used a scene-based frame in the Parallel condition, but compromised between scene-based and viewer-based alignments in the Nearly-Parallel and Skewed conditions, so that the relationship

Table 6. *Experiment 4: 150-Degree Separation, observer's lines of sight to two objects form a 150 degree angle. Mean and standard deviation of orientation to which subjects rotated comparison object*

					Mean angular difference
Parallel	0	90	180	270	
	1.5 (1.89)	92.17 (3.34)	181.5 (.96)	271.0 (1.83)	1.54 (1.75)
Nearly-Parallel	70	160	200	340	
	93.67 (21.5)	199.17 (19.36)	245.67 (16.4)	385.667 (12.66)	38.54 (17.83)
Skewed	45	135	225	315	
	94.00 (24.93)	188.0 (17.36)	272.5 (32.54)	366.67 (29.97)	50.29 (26.76)

Note. Based on 6 subjects. Degrees refer to the compass directions in Figure 3.

between their views of the two objects was like that obtained 90-Degree Separation task (and like that used by observers in Experiments 2 and 3).

Discussion

These results indicate that when the lines of sight to the objects form a 90 or 150 degree angle, the observer compares shapes by using representations that are often primarily scene-based. Because we obtain the same results with these quite different angles of view, it is likely that the tendency to use primarily scene-based representations holds for a very wide range of angles. When the objects are 150 degrees apart and observers' shoulders and trunk are immobilized, they set the comparison object close to the orientation predicted by assuming that they would perform like those in Experiment 2. Observers in Experiment 2 used perfect scene-based alignments in the Parallel condition and alignments compromising between scene-based and viewer-centered frames in the Nearly-Parallel and Skewed conditions. This indicates that the compromise alignments observed in the preceding experiments did not result because subjects turned their body or shoulders to perceive (or represent) each object with respect to a body-based frame of reference. The extent of compromise between scene-based and viewer-centered alignment depends considerably on the spatial relations of an object with significant directions in the scene. The regression away from scene-based alignment toward viewer-centered alignment can range up to a *third* of the angle between the scene-based and viewer-centered frames. Additional research is necessary to clarify more generally the role of body-based frames of reference in natural visual perception and cognition.

Experiment 5: Compromise alignment and topology of object surfaces

In the compromise alignments we observed in Experiments 1–4, corresponding cube faces on the two objects were simultaneously visible, though from different viewpoints. The additional correspondence in views of the two objects presumably aided in comparing the objects' shape or in deciding that the two objects were aligned. The amount of compromise between the scene-based and viewer-centered alignments that is required to reveal corresponding features of the two objects depends on the topology of the objects' surfaces (Arnold, 1984; Callahan & Weiss, 1985; Koenderink, 1984; Koenderink & van Doorn, 1976, 1979; McCrory, 1980). In Experiments 1–4, the objects had regular planar surfaces meeting at 0 or 90 degree angles.

In Experiment 5, we examined whether objects with other topological structures, such as those with smoothly varying surfaces, would produce compromise alignments different from those found for the objects in the preceding experiments. The results of this experiment showed the generality of the findings in Experiments 1–4. In the "Cubes-Figure" task, subjects replicated Experiment 2 with the Shepard-Metzler objects; in the "Helix" task, they replicated Experiment 2 with helical shapes formed by a 180 degree turn of wire in either a clockwise or counterclockwise direction. Both the local surface and global shape of the helical objects varied smoothly. Studies of the discrimination of mirror image and identical pairs of such helical forms have not been previously reported.

Method

Subjects

Eight MIT undergraduates, who had not been in any related experiments, participated for $5.50 an hour.

Stimuli, design, and procedure

The stimuli in the Cubes-Figure task were identical to those used in Experiment 2. The stimuli in the Helix task were composed of cotton-coated wire in a (perfect) helical shape formed by either a clockwise or counterclockwise turn of 180 degrees. Each wire shape was mounted on a thin cardboard disk so that the axis of the helix was vertical. The objects subtended the same visual angle in maximum linear extent as the objects used in Experiments 1–4. Half the subjects performed the Cubes-Figure task, then the Helix task; the other half performed the Helix task first, then the Cubes-Figure task. All other features of the design and procedure were identical to that in Experiment 2.

Results

There was no difference in the results for subjects performing the Cubes-Figure and Helix tasks in different orders (tested by ANOVAs both with and without absolute orientation within condition, see following analysis, $p > .05$). The order factor was ignored in other analyses. The data in the replication of Experiment 2 in Cubes-Figure task (Table 7) were highly correlated with those in Experiment 2, the Shape-Comparison task of Experiment 3, and the 90-Degree Separation task of Experiment 4 ($r = .999$, $F(1,10) =$

Table 7. *Experiment 5: Cubes-Figure task, Metzler & Shepard Objects. Mean and standard deviation of orientation to which subjects rotated comparison objects*

					Mean angular difference
Parallel	0	90	180	270	
	1.63 (2.60)	90.75 (2.59)	181.38 (2.99)	271.00 (1.80)	1.19 (2.56)
Nearly-Parallel	70	160	200	340	
	82.63 (6.87)	174.13 (27.61)	232.88 (12.92)	356.50 (13.35)	19.03 (20.65)
Skewed	45	135	225	315	
	69.13 (8.95)	159.00 (18.65)	247.75 (5.74)	343.38 (23.37)	24.81 (15.80)

Note. Based on 8 subjects. Degrees refer to the compass directions in Figure 3.

4683.13, $p < .0001$; $r = .999$, $F(1,10) = 4141.04$, $p < .0001$; $r = .999$, $F(1,10) = 3643.21$, $p < .0001$).

The data in the Helix task (Table 8) were very similar to those in the other replications of Experiment 2. Subjects often reported aligning the bottom of the comparison object with respect to the standard, and then comparing the direction of the turn of the top of the helices. This method is analogous to that reported by subjects in Experiment 1. (Rock, di Vita, & Barbeito, 1981, and Rock & di Vita, 1987, found that people are poor at recognizing smooth but irregular wire shapes viewed at different orientations; such inability may have influenced observers' perceptual strategy here.) Although the mean for the Nearly-Parallel condition in the Helix task was relatively low, there were no reliable differences between subjects' alignments for the two kinds of objects in Cubes-Figure and Helix tasks (in each case, $p > .05$ by a two-tailed t test).

Separate ANOVAs conducted for the Cubes-Figure and Helix tasks (ignoring absolute orientations within conditions) showed significant differences between the Parallel, Nearly-Parallel, and Skewed conditions ($F(2,14) = 16.24$, $p < .001$ and $F(2,14) = 17.83$, $p < .001$). Separate ANOVAs on each condition in the Cubes-Figure and Helix tasks showed a reliable effect of absolute orientation only for the Nearly-Parallel condition in the Helix task. This latter effect was caused by responses to the 200 degree orientation that were relatively far from the scene-based alignment, and by responses to the 70 and 340 degree orientations that were relatively close to the scene-based

Table 8. *Experiment 5: Helix task. Mean and standard deviation of orientation to which subjects rotated comparison objects*

					Mean angular difference
Parallel	0	90	180	270	
	2.00 (2.92)	91.13 (1.76)	181.00 (2.45)	272.50 (1.50)	1.66 (2.31)
Nearly-Parallel	70	160	200	340	
	75.75 (11.20)	175.88 (13.73)	231.13 (21.77)	336.00 (28.54)	12.19 (23.81)
Skewed	45	135	225	315	
	79.13 (18.09)	156.25 (17.89)	249.75 (9.86)	332.63 (21.52)	24.44 (17.28)

Note. Based on 8 subjects. Degrees refer to the compass directions in Figure 3.

alignment.[4] In each part, the error rate for the discrimination of mirror image and identical object pairs was less than .5 percent.

General discussion

Observers cannot compare the handedness of two objects without physically or mentally 'correcting' for any difference in their orientation. Our experiments examined whether observers preferred to correct for the difference in orientation of two objects by aligning them with respect to a viewer-centered frame or to a scene-based frame. Our findings suggest that to compare shape observers largely rely on representations that encode orientations with respect to the scene in which the objects lie. These findings appear to be fairly general because the same pattern of alignments is used for comparing objects with quite different surface topology, and for a range of angles between the lines of sight to each object. Use of such a scene-based frame has an important computational advantage over other frames, such as a viewer-centered frame: An observer can move about without having to update the represented dispositions of previously perceived objects. Additional studies of the sort presented in this report are necessary to clarify the role of different frames of reference in perceptual activity (see the recent varied research of Cheng & Gallistel, 1984; McDermott & Davis, 1984; Parsons & Shimojo, 1987, 1988; Shepard & Hurwitz, 1984).

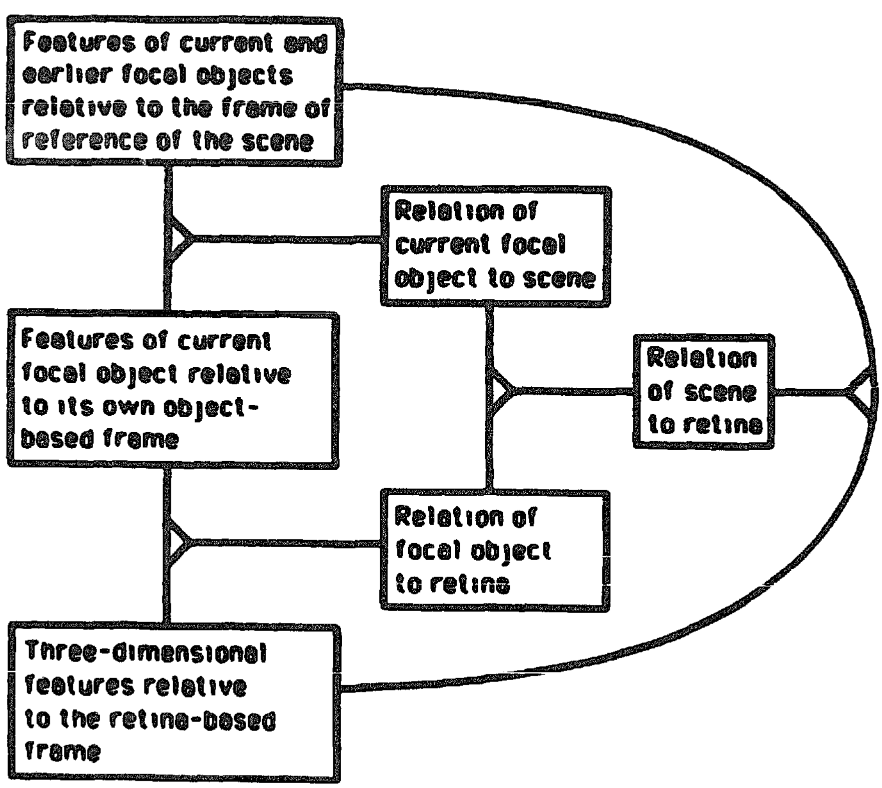
Analyses of the individual data show that the slight regression towards viewer-centered alignment is not a consequence of averaging data from many cases of scene-based alignment and a few cases of viewer-centered alignment. This leaves two main classes of explanation. The first is that subjects perform the task by using a frame of reference that is a compromise between scene-

[4]There are three observations of reliable effects of absolute orientation on the extent of compromise between scene-based and viewer-centered alignments in the Nearly-Parallel and Skewed conditions. In two cases, this effect occurred in the replication of the Shape-Comparison task from Experiment 2. An ANOVA of responses in each condition was conducted on the data from all performances of this task in Experiments 2–5. This analysis showed a reliable effect of absolute orientation only in the Nearly-Parallel condition ($F(3,66) = 5.86, p < .001$). The effect occurred because the responses in the 70 degree orientation were relatively near scene-based alignment and those in the 200 degree orientation were relatively far from scene-based alignment: the overall means were 10.85, 19.67, 27.81, and 18.57 for the 70, 160, 200, and 340 degree orientations. The only other effect of absolute orientations was in the Nearly-Parallel condition in the Helix task (in Experiment 5). The range of compromise observed there varied more widely than usual across all the absolute orientations. It remains a mystery why these variations in compromise alignments occurred for different absolute orientations. These effects could have been due to one of the factors not counterbalanced in our experimental design: for example, either the order of trials (all subjects saw the same sequence of trials) or the interaction between a particular-handed object shape and the absolute orientation at which it was presented (e.g., some preference for viewing the object from some angle). Further research is necessary to understand this finding.

based and viewer-centered. The second is that they use several frames of reference—viewer-centered, object-based, and scene-based—and that the final relationship between the object-based and scene-based frames is the result of a compromise between conflicting requirements. Although the data do not discriminate between these two types of explanation, we prefer the second because it is compatible with a more detailed theory of the frames of reference that are used in vision and spatial reasoning (Hinton, 1981b; Hinton & Parsons, 1981).

According to this theory, the initial, retina-based (viewer-centered) representation of an object is mapped into a canonical object-based representation which allows it to be recognized, and this object-based representation is

Figure 5. *A sketch of the groups of units and the interactions that could be used for object recognition and spatial reasoning. Each of the groups on the left represents spatial structures relative to a different reference frame (see Hinton, 1981b for further details). Within this apparatus, mental rotation could be performed by holding the object-based representation fixed and continuously altering the representation of the mapping between the object-based and scene-based units. The changing mapping would cause changes in the contents of the scene-based representation.*



then mapped into a scene-based representation that explicitly encodes the relationship of the object (and its major constituents) to the scene-based frame. Figure 5 illustrates the relationships among these three different representations of an object. Using these representations, the same-different judgement can be performed in two stages: First, the relationship between the moved object and the scene is altered (mentally or physically) until at least two, non-parallel limbs of the moved object are aligned (in the scene) with the corresponding limbs of the unmoved object. Then a third, non-coplanar limb of the moved object is compared with the corresponding limb of the unmoved object to see whether it points in the same or opposite directions in the scene.

If the judgements really are made in this way, there are two conflicting requirements on the orientation to which the moved object is rotated. The first requirement is that the orientation of the moved object with respect to the scene-based frame must be approximately the same as the orientation of the unmoved object in order to judge whether corresponding limbs of the two objects point in the same or opposite directions in the scene. But since this is a binary judgement it should be robust against small differences in the scene-based orientations of the two objects, so the first requirement does not entirely determine the final orientation of the moved object. Other studies of mental rotation (e.g., Hock & Tromley, 1978) have suggested that it is not necessary to rotate a letter or number into perfect vertical alignment in order to judge its handedness.

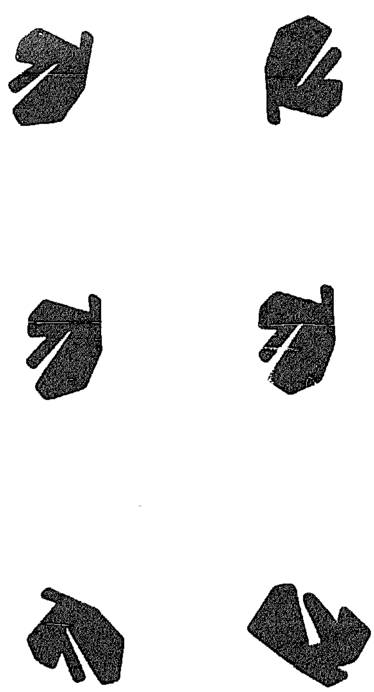
The second requirement is that the subject correctly identifies corresponding limbs in the two objects. It probably becomes easier to detect the correspondence between two limbs as the views of the two limbs become more similar—as two views become more similar it becomes less likely that surfaces which are visible in one limb are invisible in the other. So the regression towards viewer-centered alignment may have the function of making it easier to identify the correspondences. Further research is required to test this explanation.

Appendix 1: Why object-based frames are insufficient for the task

Hinton and Parsons (1981) suggest why object-based frames alone may be insufficient for the task of comparing shapes at different orientations. They hypothesize that when people impose an object-based frame they can use either a left-handed or a right-handed frame, and they have no explicit representation of the handedness of the frame they impose. They typically choose whichever handedness leads to a familiar object-based shape description. So when presented with a pair of mirror image objects, they will impose frames of opposite handedness and the two objects will receive *identical* object-based descriptions. The only difference lies in the handedness of the imposed frames which is not consciously available.

One problem with this explanation is that the assumption that we do not know the handedness of non-upright frames seems rather arbitrary. If the orientation of the frame is represented as a unitary entity such as a point in orientation space, it is surprising that a handedness cannot be associated with each such point. But this inability is much less surprising if the representation of the frame's orientation is componential. For example, the orientations and senses of the x, y, and z axes could be represented separately (subject to the constraint that they are mutually orthogonal). The handedness of the frame composed of these axes would then be a third-order property of the encoding, which could not be learned by any linear associator. (This can be proved using the group invariance theorem of Minsky and Papert, 1969.) So if the

Figure 6. *Illustration of how the similarity of a pair of shapes can influence the effect of their orientation difference on perceptual processing. It is more difficult to determine whether the shapes in panel (a) are identical than the shapes in panel (b) or (c). (The shapes in panel (b) are the same as those in (a) but are at the same orientation. See Attneave & Arnoult, 1956, and Cooper & Podgorny, 1976.)*



orientation of a frame has a componential representation it becomes reasonable to postulate that we only know the handedness for frames with the canonical, upright orientation (it is easy to associate handedness with one particular case).

There is a second problem for any theory that uses handedness considerations to explain why shape comparison tasks cannot be solved by just comparing object-based representations. This hypothesis does not explain why object-based frames cannot be used for discriminating differently-oriented, but similar, shapes that have the same handedness but differ slightly in some small feature. Yet mental (or physical) rotation is likely used for this task also. (The more similar two shapes are, the more likely is it that we prefer to view them from a comparable perspective to detect any differences in their shape. See Figure 6.) One possible explanation is that the object-based descriptions of the *whole* objects are too similar. To make the discrimination reliably it may be necessary to attend to parts of the two objects where they differ most. This may require many pairwise comparisons of corresponding parts of the two objects. The computations involved in finding corresponding parts are harder if the objects are in different orientations (see Parsons, 1988; Shepard & Farrell, 1985), so it may be worth the 'overhead' of performing an initial mental (or physical) rotation of one object to make these computations easier (also see Just & Carpenter, 1985, p. 166 ff.).

References

Andersen, R.A., Essick, G.K., & Siegel, R.M. (1984). The role of eye position on the visual response of neurons in area 7a. *Society for Neuroscience Abstracts, 10*, 934.

Andersen, R.A., Essick, G.K., & Siegel, R.M. (1985). The encoding of spatial location by posterior parietal neurons. *Science, 230*, 456–458.

Arnold, V.I. (1984). *Castrophe theory*. Berlin: Springer-Verlag.

Attneave, F. (1968). Triangles as ambiguous figures. *American Journal of Psychology, 81*, 447–453.

Attneave, F. (1972). Representation of physical space. In A.W. Melton & E.J. Martin (Eds.), *Coding processes in human memory*. Washington, DC: Winston.

Attneave, F., & Arnoult, M.D. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin, 53*, 452–471.

Attneave, F., & Farrar, P. (1977). The visual world behind the head. *American Journal of Psychology, 90*, 549–563.

Attneave, F., & Olson, R. (1967). Discriminability of stimuli varying in physical and retinal orientation. *Journal of Experimental Psychology, 74*, 149–157.

Attneave, F., & Pierce, C.R. (1978). Accuracy of extrapolating a pointer into perceived and imagined space. *American Journal of Psychology, 91*, 371–387.

Attneave, F., & Reid, K. (1968). Voluntary control of frame of reference and slope equivalence under head rotation. *Journal of Experimental Psychology, 78*, 153–159.

Ballard, D.H. (1986). Cortical connections and parallel processing: Structure and function. *The Behavioral and Brain Sciences, 9*, 67–120.

Bauermeister, M. (1964). Effect of body tilt on apparent verticality, apparent body position, and their relation. *Journal of Experimental Psychology, 67*, 142–147.

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. Pomerantz (Eds.), *Perceptual Organization*. Hillsdale, NJ: Erlbaum.

Bridgeman, B., Lewis, S., Heit, G., & Nagle, M. (1979). Relation between cognitive and motor-oriented systems of visual position perception. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 692–700.

Callahan, J., & Weiss, R. (1985). A model for describing surface shape. *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition* (pp. 240–245).

Carpenter, P.A., & Just, M.A. (1978). Eye fixations during mental rotation. In J.W. Senders, D.F. Fisher, & R.A. Monty (Eds.), *Eye movements and the higher psychological functions* (pp. 115–133). Hillsdale, NJ: Erlbaum.

Cattell, J. (1900). On the relations between time and space in vision. *Psychological Review, 7*, 325–343.

Cheng, K., & Gallistel, C.R. (1984). Testing the geometric power of an animal's spatial representation. In H.L. Roitblat, T.G. Bever, & H.S. Terrace (Eds.), *Animal Cognition* (pp. 409–423). Hillsdale, NJ: Erlbaum.

Cooper, L.A. (1975). Mental rotation of random two-dimensional shapes. *Cognitive Psychology, 7*, 20–43.

Cooper, L.A., & Podgorny, P. (1976). Mental transformations and visual comparison processes: Effects of complexity and similarity. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 503–514.

Cooper, L.A., & Shepard, R.N. (1973). Chronometric studies of the rotation of mental images. In W.G. Chase (Ed.), *Visual information processing* (pp. 75–176). New York: Academic Press.

Corballis, M.C., Anuza, T., & Blake, L. (1978). Tachistoscopic perception under head tilt. *Perception & Psychophysics, 24*, 274–284.

Corballis, M.C., Nagourney, B.A., Shetzer, L.I., & Stefanatos, G. (1978). Mental rotation under head tilt: Factors influencing the location of subjective reference frame. *Perception & Psychophysics, 24*, 263–273.

Corballis, M.C., & Roldan, C.E. (1976). Detection of symmetry as a function of angular orientation. *Journal of Experimental Psychology: Human Perception & Performance, 1*, 221–230.

Corballis, M.C., Zbrodoff, J., & Roldan, C.E. (1976). What's up in mental rotation? *Perception & Psychophysics, 19*, 525–530.

Day, R.H., & Wade, N.J. (1969). Mechanisms involved in visual orientation constancy. *Psychological Bulletin, 71*, 33–42.

Gilinsky, A. (1955). The effect of attitude on the perception of size. *American Journal of Psychology, 68*, 173–192.

Girgus, J.S., Gellman, I.H., & Hochberg, J. (1980). The effect of spatial order on piecemeal shape recognition: A developmental study. *Perception & Psychophysics, 28*, 133–138.

Gross, C.G., Rocha-Miranda, C.E., & Bender, D.B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology, 35*, 96–111.

Hinton, G.E. (1981a). A parallel computation that assigns canonical object-based frames of reference. *Proceedings of the 7th International Joint Conference on Artificial Intelligence*. Vancouver, BC.

Hinton, G.E. (1981b). Shape representation in parallel systems. *Proceedings of the 7th International Joint Conference on Artificial Intelligence*. Vancouver, BC.

Hinton, G.E., & Parsons, L.M. (1981). Frames of reference and mental imagery. In A.D. Baddeley and J. Long (Eds.) *Attention and Performance* (Vol. 9, pp. 261–277). Hillsdale, NJ: Erlbaum.

Hochberg, J. (1968). In the mind's eye. In R.H. Haber (Ed.), *Contemporary theory and research in visual perception*. New York: Holt, Rinehart, & Winston.

Hochberg, J. (1982). How big is a stimulus? In J. Beck (Ed.), *Organization and representation in perception* (pp. 191–217). Hillsdale, NJ: Erlbaum.

Hock, H., & Tromley, C.L. (1978). Mental rotation and perceptual uprightness. *Perception & Psychophysics*, *24*, 529–533.

Howard, I.P. (1986). The perception of posture, self motion, and the visual vertical. In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), *Handbook of perception and human performance (Vol. 1: Sensory processes and perception)*. New York: Wiley.

Hubel, D.H., & Wiesel, T.H. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology (London)*, *160*, 106–154.

Hubel, D.H., & Wiesel, T.H. (1974). Uniformity of monkey striate cortex: A parallel relationship between field size, scatter, and magnification factor. *Journal of Comparative Neurology*, *158*, 295–302.

Just, M.A., & Carpenter, P.A. (1985). Cognitive coordinate systems. *Psychological Review*, *92*, 137–171.

Kaushall, P., & Parsons, L.M. (1981). Optical information and practice in the discrimination of 3-D mirror-reflected objects. *Perception*, *10*, 545–562.

Keenan, J.M. & Moore, R.E. (1979). Memory for images of concealed objects: A re-examination of Neisser and Kerr. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 374–385.

Koenderink, J.J. (1984). The internal representation of solid shape and visual exploration. In L. Spillman & B.R. Wooten, *Sensory experience, adaptation, and perception* (pp. 132–142). Hillsdale, NJ: Erlbaum.

Koenderink, J.J., & Doorn, A.J. van (1976). Singularities of the visual mapping. *Biological Cybernetics*, *24*, 51–59.

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

Kosslyn, S.M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Lynch, J.C. (1980). The functional organization of posterior parietal association cortex. *The Behavioral and Brain Sciences*, *3*, 485–534.

Mach, E. (1897). *The analysis of sensations*. English translation (1959), New York: Dover.

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

McCrory, C. (1980). *Profiles of surfaces*. Coventry: University of Warwick.

McDermott, D., & Davis, E. (1984). Planning routes through uncertain territory. *Artificial Intelligence*, *22*, 107–156.

Metzler, J. (1973). *Cognitive analogues of the rotation of three-dimensional objects*. Unpublished doctoral dissertation, Stanford University.

Metzler, J., & Shepard, R.N. (1974). Transformational studies of the internal representation of three-dimensional objects. In R. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium* (pp. 174–201). Hillsdale, NJ: Erlbaum.

Mishkin, M., Ungerleider, L.G., & Macko, K.A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, *6*, 414–417.

Minsky, M.L., & Papert, S. (1969). *Perceptrons*, Cambridge, MA: MIT Press.

Morgan, M.J., Findlay, J.M., & Watt, R.J. (1982). Aperture viewing: A review and a synthesis. *Quarterly Journal of Experimental Psychology*, *34A*, 211–233.

Noda, H., & Warabi, T. (1986). Discharges of Purkinje cells in monkey's flocculus during pursuit eye movements and visual stimulus movements. *Experimental Neurology*, *93*, 390–403.

Olson, R.K., & Attneave, F. (1970). What variables produce similarity grouping? *American Journal of Psychology*, *83*, 1–21.

Palmer, S.E. (1975). Visual perception and world knowledge: Notes on a model of sensory cognitive interaction. In D.A. Norman & D.E. Rumelhart (Eds.), *Explorations in cognition*. San Francisco: W.H. Freeman.

Palmer, S.E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, *9*, 441–474.

Palmer, S.E., & Bucher, N. (1981). Configural effects in perceived pointing of ambiguous triangles. *Journal of Experimental Psychology: Human Perception & Performance*, *7*, 88–114.

Parsons, L.M. (1987a). Imagined spatial transformation of one's body. *Journal of Experimental Psychology: General, 116*, 172–191.
Parsons. L.M. (1987b). Imagined spatial transformation of one's hands and feet. *Cognitive Psychology, 19*, 178–241.
Parsons, L.M. (1987c). Visual discrimination of abstract mirror-reflected three-dimensional objects at many orientations. *Perception & Psychophysics, 42*, 49–59.
Parsons, L.M. (1988). Serial search and comparison of pairs of objects. *Memory & Cognition, 16*, 23–35.
Parsons, L.M., & Shimojo, S. (1987). Perceived spatial organization of cutaneous patterns on surfaces of the human body in various positions. *Journal of Experimental Psychology: Human Perception & Performance, 13*, 488–504.
Parsons, L.M., & Shimojo, S. (1988). *Spatial information processing in the cutaneous perception and motor production of patterns*. Manuscript submitted for publication.
Pinker, S. (1980). Mental imagery and the third dimension. *Journal of Experimental Psychology: General, 109*, 354–371.
Pinker, S. (1985). Visual cognition: An introduction. In S. Pinker (Ed.), *Visual cognition* (pp. 1–64). Cambridge, MA: MIT Press.
Pinker, S., & Finke, R.A. (1980). Emergent two-dimensional patterns in images rotated in depth. *Journal of Experimental Psychology: Human Perception & Performance, 6*, 69–84.
Robinson, J.O. (1972). *The psychology of visual illusion*. London: Hutchinson & Co.
Rock, I. (1973). *Orientation and form*. New York: Academic.
Rock, I. (1981). Anorthoscopic perception. *Scientific American, 244*, 103–111.
Rock, I., & di Vita, J. (1987). A case of viewer-centered object perception. *Cognitive Psychology, 19*, 280–293.
Rock, I., di Vita, J., & Barbeito, R. (1981). The effect on form perception of change of orientation in the third dimension. *Journal of Experimental Psychology: Human Perception & Performance, 7*, 719–732.
Rock, I., & Heimer, W. (1957). The effect of retinal and phenomenal orientation on the perception of form. *American Journal of Psychology, 70*, 493–511.
Sedgewick, H.A. (1986). Space perception. In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), *Handbook of perception and human performance (Vol. 1: Sensory processes and perception)*. New York: Wiley.
Shepard, R.N., & Cooper, L. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.
Shepard, R.N., & Farrell, J.E. (1985). Representations of the orientations of shapes. *Acta Psychologica, 59*, 103–121.
Shepard, R.N., & Hurwitz, S. (1984). Upward direction, mental rotation, and discrimination of left and right turns in maps. *Cognition, 18*, 161–193.
Shepard, R.N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 191*, 952–954.
Shimojo, S., & Richards, W. (1986). "Seeing" shapes that are almost totally occluded: A new look at Parks's camel. *Perception & Psychophysics, 39*, 418–426.
Thomson, J.A. (1980). How do we use visual information to control locomotion? *Trends in Neuroscience, 3*, 247–250.
Thomson, J.A. (1983). Is continuous visual monitoring necessary in visually guided locomotion? *Journal of Experimental Psychology: Human Perception and Performance, 9*, 427–443.
Ungerleider, L.G., & Mishkin, M. (1982). Two visual systems. In D.J. Ingle, M.A. Goodale, & R.J.W. Mansfield (Eds.), *Analysis of visual behavior*. Cambridge, MA: MIT Press.
Uhlarik, J., Pringle, R., Jordan, K., & Misceo, G. (1980). Size scaling in two-dimensional pictorial arrays. *Perception & Psychophysics, 27*, 60–70.
Van Essen, D.C. (1985). Functional organization of primate visual cortex. In A. Peters & E.G. Jones, *The cerebral cortex (Vol. 3)*. New York: Plenum Press.

Résumé

Cinq expériences ont examiné l'utilisation de deux types de représentations (en coordonnées fixes par rapport à la scène, ou en coordonnées égocentriques) dans la comparaison de formes sous différentes orientations. Deux objets, soit les dessins utilisés par Metzler et Shepard (1974), soit des hélices, étaient placés sur une table. Les lignes de visée de l'observateur vers ces objets étaient séparés par 90 ou 150 degrés. Chaque observateur devait décider si les formes des objets étaient identiques ou imagés par un miroir, et devait *tourner physiquement* l'un des objets vers une orientation qui lui permettait de prendre une décision. Dans cette tâche, les observateurs tournent souvent l'objet jusqu'à ce qu'il ait la même relation à la table que l'autre objet, utilisant donc un alignement par rapport à la scène extérieure bien que cela produise des images rétiniennes très différentes pour les deux objets. Les réponses régressent jusqu'à à un tiers vers un alignement en coordonnées égocentriques lorsque les surfaces principales d'un objet sont de moins en moins bien alignées avec des directions saillantes de la scène. Le même comportement d'alignement est observé avec des paires d'objets de topologie de surface très différentes. Quand on demande aux sujets de tourner un objet pour qu'il soit parfaitement aligné avec l'autre par rapport à la scène, les réponses reviennent à nouveau vers un alignement en coordonnées égocentriques. Ce retour atteint 15 degrés quand l'objet standard (qui n'est pas tourné) est moins bien aligné avec les directions saillantes de la scène. Globalement, ces résultats suggèrent que dans cette tâche de comparaison des formes, on utilise des représentations en coordonnées fixes par rapport au monde extérieur plutôt qu'en coordonnées égocentriques.