Visual information for object identity in apparent movement

WILLIAM H. WARREN
Hampshire College, Amherst, Massachusetts 01002

Apparent movement is used to examine the nature of the visual information which specifies object identity. Constructive feature-comparison theories rely on static formal information and predict that two phases of an apparent movement display must be featurally similar in order to appear as a single object in motion. An opposing Gibsonian model is based on abstract geometrical information and predicts that the two phases may differ radically in shape but must be ecologically transformable in order to be seen as a single object. The predictions are tested by presenting subjects with displays which are similar but not transformable, and transformable but dissimilar. Results show that the transformability of the display determines its perceived object identity, while featural similarity has no effect. An event theory of object identity is offered which claims that the *affordance structure* of an event, determined by geometric invariants specifies object identity.   ma-

The problem of phenomenal object identity or constancy in event perception can be stated as follows: How do different patterns of stimulation, occurring over a period of time, result in the perception of a unified event with continuous objects? What information specifies the phenomenal identity of an object undergoing change? The apparent movement phenomenon can be utilized to explore the information for object identity because it provides a situation where perceptually valid and invalid changes of shape can be presented.

Apparent movement occurs when two separated stationary lights (or phases) are flashed successively with an interstimulus interval (ISI) of about 60 msec. The result is "optimal movement" of a *single object* appearing to translate smoothly through space. Conventionally, the two phases of apparent movement displays have been the same shape, although several attempts to vary shape have been made. Orlansky (1940) found that low similarity seemed to inhibit movement. Squires (1959) reported findings which indicate that similar objects (e.g., three-cornered figures, rounded figures) tended to support good movement while dissimilar objects did not. Both these studies left the concept of similarity formally undefined and did not obtain reports of perceived object identity. Foster (1972a, b, 1973a, b) defined the phase relations with mathematical rigor, but experimented with rotations in the picture plane and not with changes of shape. Kolers (1972; Kolers

& Pomerantz, 1971) found that disparate pairs, such as a square and a triangle, could be resolved into optimal movement, appearing to "change into" one another when presented optimally or superimposed. He concluded that motion is the primary event and figure, or contour, secondary, and that the illusion of motion and identity in optimal movement is based solely on spatio-temporal conditions. However, Kolers used outline drawings instead of solid object-like shapes, presented them in the retinal periphery (1.7° and 5.4° from a right-hand fixation point), used very well-practiced subjects, and counted reports of "smooth continuous motion" instead of strict object identity.

The dominant theory of object identity in perceptual events has been called "visual synthesis" by Neisser (1967) and involves the integration of successive visual "snapshots" taken of a changing object over time (see also Stroud, 1955). Such constructive theories rely on static formal features to characterize each snapshot, as developed in recent work on pattern recognition (Dodwell, 1970; Neisser, 1967; Selfridge, 1959; Uhr, 1963). Successive snapshots are stored in memory and compared for featural similarity to determine their identity or disparity, similar ones being attributed to a common object and integrated into a unified percept. Two shapes are generally considered "similar" if they contain the same number and relative positioning of the same features. Other concepts of similarity have been proposed, such as ratio of overlap in prototypic forms (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). The common featural definition was tested in the present study.

J. J. Gibson's (1966a, b) theory of ecological optics suggests an opposing account of event percep-

tion with two central theses: (1) that the world contains sufficient information in the form of abstract invariants to specify its contents and relations, and (2) that perception of the world is veridical and direct because an exploring organism can detect such information. A theory of object identity based on Gibson's proposal states that people detect abstract information for the transformability of two shapes (Shaw & Pittenger, Note 1), specifically that geometric properties which remain invariant under ecologically valid transformations of shape directly specify object identity. Thus, a change of shape is not a problem to be explained but is itself a rich source of information about an object. This concept of transformability is consonant with Foster's (1973a) formalization of structure-preserving diffeomorphisms but constrains the permissible transformations with the notion of ecological validity. E. J. Gibson (Note 2) has catalogued such ecologically valid shape transformations, which include translation, looming, perspective shift, and such biomorphic deformations as locomotion and growth.

On this basis, Shaw has proposed (Pittenger & Shaw, 1975; Shaw & Wilson, in press) that every event is defined by its *affordance structure*, composed of a "structural invariant," the constants of the object or layout, and a "transformational invariant," the type of transformation or change the object undergoes. This suggests a framework in which to discuss object identity in apparent movement. Illusions such as apparent movement are successful when they simulate the information which specifies a given event. If the requirements of a natural affordance structure are not met, i.e., if there is no unified object specified by the shapes of the two phases and no ecological transformation which relates them, then the illusion of object identity in apparent movement should break down.

The two models yield conflicting predictions about the types of shape changes which should result in perceived object identity. The feature comparison hypothesis predicts that the two phases of an apparent movement display must be featurally similar, if not identical, to be attributed to a common object and appear as a single object in motion. In contrast, the transformability hypothesis predicts that the two phases may have different featural shapes as long as they are transformable, i.e., are related by a transformation defined within a mathematical group of geometric transformations (e.g., a projective group) and thereby specify an ecological transformation (e.g., a perspective shift). Group invariants are not features. They are not static formal attributes which must be compared inferentially but are abstract relations defined over a change of shape, independent of the specific shape being changed. The following experiment tests the feature and trans-

formability predictions with basic shapes by presenting displays which are featurally similar but not transformable, and others which are transformable but dissimilar.

METHODS

Subjects
Eight Hampshire College undergraduates, five women and three men, served as subjects. All were naive as to the nature and purpose of the experiment. Two other subjects were removed from the sample, one because he failed to perceive movement with the practice displays and one due to apparatus breakdown during testing.

Apparatus
The apparent-movement phenomenon was created on a three-channel tachistoscope (Scientific Prototype Model GB) with mercury-argon gas discharge lamps. Four timers were used to form a continuous presentation cycle, linking the "blank" channel and the two phases of the display, which could be switched on and off by the experimenter. Display backgrounds were equalized for intensity. Subjects viewed the display in a dim room through a binocular eyepiece at a distance of 165 cm.

Displays
The three practice and eight test displays are illustrated in Figure 1. The shapes were cut from flat black exposed photographic paper and mounted on off-white matboard. The prototypic form was the square, which had an area of 1 cm² and subtended 0.5° of visual angle. The triangle, two hexagons, rectangle, and practice lines were also 1 cm²; the four polar projections in Displays d, e, h, and i (see Gibson, 1957) obviously could not be controlled for area. All displays were presented so that the centers of gravity of the two phases were aligned on the same horizontal axis and separated by 1.5° of visual angle.

The displays were chosen along two dimensions: ecologically transformable vs. nontransformable, and featurally similar vs. dissimilar. A display was classified as similar if its two phases
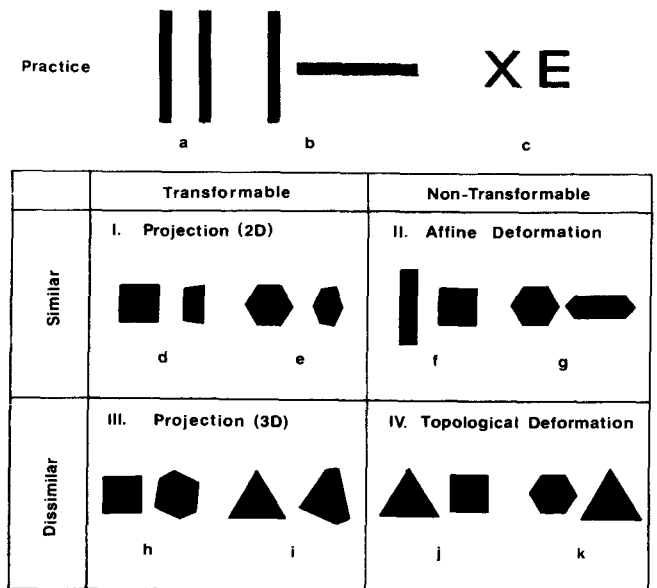

Figure 1. Two-phase displays used in the experiment.

had the same number of sides and corners, with changes of length and angle permitted.

The interaction of these two dimensions set up a four-condition design, each condition of which could be redescribed mathematically at a given level of transformation geometry (Figure 1). *Condition I*: similar and transformable (projective geometry, mapping a two-dimensional figure onto a two-dimensional plane). *Condition II*: similar but nontransformable (affine geometry—strain transformations). *Condition III*: dissimilar but transformable (projective geometry, mapping a three-dimensional object onto a two-dimensional plane). *Condition IV*: dissimilar and nontransformable (topology).

Procedure

Displays were always presented in the following cycle, consistent with the literature on optimal movement and pilot work: left phase for 60 msec, blank background for 60 msec (ISI), and right phase for 60 msec. A recycle interval (RCI) of 500 msec was used to create the effect of "single-shot" presentations (as in Squires, 1959; and Kolers & Pomerantz, 1971), rather than continuous back-and-forth movement. The blank remained on during the RCI and whenever the presentation cycle itself was switched off.

A 45-min experimental session began with the three practice displays, with the dual purpose of familiarizing subjects with the perceptual effect of one moving vs. two separate objects, and to eliminate any subjects not susceptible to the apparent-movement illusion. The subjects were asked to describe what they saw, and if the illusion was not obvious, were encouraged to see the display as one object in motion. Display a was the standard illusion; Display b illustrated a transformation (translation) which did not appear in the test series; and Display c illustrated the appearance of two separate objects to help clarify the subjects' criteria and break any set for seeing one object.

The test series then began. Subjects were instructed to fixate the spot where the figure appeared on the left, and the presentation cycle was switched on. After 20 presentations of the display, subjects were asked whether they saw one object or two, and were asked to describe any visible movement. The presentation cycle was left on during this time, and the responses recorded. Subjects had brief rest periods while display cards were changed.

Each subject was shown all eight displays in a different randomized sequence, which was repeated three times for that subject. The display sequences were counterbalanced across subjects for position and order effects, using a Latin square design.

RESULTS

Perceived Object Identity

Trials were scored with a "1" if the subject reported seeing one object, and with a "0" if she/he reported two separate objects. Since the variance of total subject scores was negligible ($s^2 = 0.5$), the data were homogeneous and trials (three per subject per display) were treated individually for statistical purposes. The total number of trials in which perceived object identity was reported, together with its percentage of total trials in each condition, is given in Table 1.

With transformable displays, the subjects reported object identity on 95.8% of the trials (96 trials per dimension), while this was true on a scant 6.3% of the trials with nontransformable displays. A test for significance of difference between two proportions (z statistic) showed this difference to be highly significant ($z = 12.40$, $p < .001$).[1] On the other hand, the dimension of similarity had no statistical effect; the subjects perceived one object on 53.1%

Table 1
Number and Percentage of Trials in Which Perceived Object Identity Was Reported

Transformable			Nontransformable			Totals
			Similar			
I. d	e	Total	II. f	g	Total	Total
23	24	47 (97.9%)	4	0	4 (8.3%)	51 (53.1%)
			Dissimilar			
III. h	i	Total	IV. j	k	Total	Total
22	23	45 (93.8%)	2	0	2 (4.2%)	47 (48.9%)
Totals		92 (95.8%)			6 (6.3%)	

Note—Number of trials per display = 24. Roman numerals and letters refer to displays in Figure 1.

of the similar display trials and on 48.9% of the dissimilar display trials ($z = 0.58$, n.s.).

This highly significant difference for the transformability dimension was maintained across pairs of conditions. Specifically, it occurred between Condition I and Condition II, with similar and transformable displays at 97.9% perceived object identity (48 trials per condition) and similar but nontransformable displays at only 8.3% ($z = 8.80$, $p < .001$), and between Condition IV and Condition III, with dissimilar and nontransformable displays at 4.2% and dissimilar but transformable displays at 93.8% ($z = 8.37$, $p < .001$). It is clear from Table 1 that any differences between Conditions I and III, or between Conditions II and IV, are not significant.

Verbal Reports

Nearly all reports on transformable displays were consistent with the perception of a single object—"a door swinging out" (Dislay d), "a cube rotating" (Display h)—and on nontransformable displays with the perception of two separate objects—"the rectangle moves right and the square blinks on and off" (Display f). In the few cases where nontransformable displays were seen as one object, the verbal reports were consistent with that perception as well. For instance, Display f was seen as one object on only four trials, with verbal reports of "the end of a book, turning to the cover," or "a rectangle shrinking and getting fatter."

Six of the eight subjects spontaneously reported seeing crucial Display h as a three-dimensional cube. The remaining two did not mention dimensionality, and in such cases after the last presentation of the display the subject was asked, "Does it look like it's flat or in three dimensions?" These two subjects then immediately saw the display in three dimensions and indicated that the perception was very compelling. Only three of the eight subjects spontaneously reported seeing Display i as a three-dimensional object; after the dimensionality question, two more reported a three-dimensional object, making a total of five.

The subjects who did not report Displays h and i as three-dimensional yet saw them as single objects

usually reported very ambiguous types of transformations, such as a two-dimensional square or triangle being "distorted," "crushed ... amorphous," "somehow transformed," or "getting its corner chopped off."

DISCUSSION

The results confirm the predictions of the transformability hypothesis. It appears that the two phases of a display are not compared and judged for identity on the basis of their featural similarity, for some similar displays are seen as disjunct objects (Displays f, g), while some dissimilar displays are seen as phenomenally identical (Displays h, i). Even when stimuli are not continuous through time but simulate mental snapshots, perception seems to be based on the transformation relation specified across the two phases of the display.

The three subjects who failed to see Display i in three dimensions yet reported it as a single object present a difficulty. They did not detect the depth information for a rigid three-dimensional object, but described a very ambiguous transformation. Some unspecified information for object identity must have been available in the display. An argument could be made for the influence of some Gestalt-like property of global similarity, independent specific local features. However, if this were the case, it would be expected that there would be fewer reports of object identity in Condition III than in Condition I, and in Condition IV than in Condition II. There were no such differences in the results. Alternatively, information for a two-dimensional biomorphic change of shape could have been available in Display i. The verbal reports suggest such a transformation, but further experiments on information for biomorphic and other transformations are needed.

The results differ from those of Kolers (1972; Kolers & Pomerantz, 1971), who found that disparate shapes can be made to "change into" one another under special circumstances. However, in the present study, under constant spatio-temporal conditions, the shape relations alone determined whether the display was seen as the optimal movement of one object or as two separate, often stationary, objects. Kolers' hypothesis cannot explain such an influence of "contour properties" on perceptual identity and movement because it considers them to be independent of and secondary to the spatio-temporal properties which determine identity and motion.

As noted in the introduction, the affordance structure of an event is composed of a component that remains constant and a component that changes (Pittenger & Shaw, 1975). The transformable displays in this experiment contain geometric information which specifies a single object undergoing a certain change—the affordance structure of a certain unified event. How may these transformability relations be described (see also Foster, 1973a)?

The transformable displays (d, e, h, i) are perspective shifts defined in projective geometry, which holds the geometric property of cross-ratios *constant* under transformation (Jeger, 1966). This property specifies the ecological transformability of the displays, hence their object identity, and constitutes the "structural invariant" of optimal movement. The constant cross-ratios, together with *changes* in other properties of length, angle, length ratios, and vector properties, jointly specify a perspective shift. This constitutes the "transformational invariant" of the event, which describes the specific transformation the object is undergoing. Subjects' descriptions of the resulting unified events were specific and consistent. On the other hand, nontransformable Displays j and k can only be transformed in topology, which does not maintain cross-ratios or any other metric contour properties.

However, the simple formulation that cross-ratios specify object identity cannot be established as the minimal necessary information. Displays f and g are defined in affine geometry, which holds cross-ratios *and* vector properties constant while length ratios, angle, and length may vary, yet they are not ecologically transformable and appear to subjects as two separate objects. Thus, this particular affordance structure specifies an ecologically invalid transformation. The unecological nature of affine transformations can be explained by the fact that they are defined as projections of figures with a projection point at infinity, so that a tilted rectangle is shortened but exhibits no line convergence. We do not naturally see objects tilting at infinite distances and thus affine transformations specify not tilting but squashing, which requires physical contact by an external agent (Shaw, McIntyre, & Mace, 1974).

The geometric invariants which specify the identity of rigid objects are apparently found in projective geometry and not in topology, for metric properties seem to differentiate transformable and nontransformable shapes (e.g., Displays j, k). On the other hand, biomorphic transformations of nonrigid objects (such as growth, flexion, and locomotion) may require excursions into topology to locate their defining invariants (see Pittenger & Shaw, 1975).

NOTE

1. The cutoff point for $p < .001$ is $z = 3.30$.

REFERENCE NOTES

1. Shaw, R. E., & Pittenger, J. *On perceiving change.* Paper presented at the Conference on Modes of Perception, Spring Hill, Minnesota, June 1975.

2. Gibson, E. J. *A classification of events for the study of event perception.* Paper presented at the meeting of the American Psychological Association, Chicago, September 1975.

REFERENCES

DODWELL, P. C. *Visual pattern recognition.* New York: Holt, Rinehart, & Winston, 1970.

FOSTER, D. H. A method for the investigation of those transformations under which the visual recognition of a given object is invariant. I. The theory. *Kybernetik,* 1972, **11**, 217-222. (a)

FOSTER, D. H. A method for the investigation of those transformations under which the visual recognition of a given object is invariant, II. An example experiment: The group of rotations SO(2) acting on a Landolt ring. *Kybernetik,* 1972, **11**, 223-229. (b)

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

FOSTER, D. H. An experimental examination of a hypothesis connecting visual pattern recognition and apparent motion. *Kybernetik,* 1973, **14**, 63-70. (b)

GIBSON, J. J. Optical motions and transformations as stimuli for visual perception. *Psychological Review,* 1957, **64**, 288-295.

GIBSON, J. J. *The senses considered as perceptual systems.* Boston: Houghton Mifflin, 1966. (a)

GIBSON, J. J. The problem of temporal order in stimulation and perception. *Journal of Psychology,* 1966, **62**, 141-149. (b)

JEGER, M. *Transformation geometry.* (A. Deicke & A. Howson, Trans.). London: George Allen and Unwin, 1966.

KOLERS, P. A. *Aspects of motion perception.* Oxford: Permagon, 1972.

KOLERS, P. A., & POMERANTZ, J. R. Figural change in apparent motion. *Journal of Experimental Psychology,* 1971, **87**, 99-108.

NEISSER, U. *Cognitive psychology.* New York: Appleton-Century-Crofts, 1967.

ORLANSKY, J. The effect of similarity and difference in form on apparent visual movement. *Archives of Psychology,* 1940, **35** (Whole No. 246).

PITTENGER, J., & SHAW, R. E. Aging faces as viscal-elastic events: Implications for a theory of nonrigid shape perception. *Journal of Experimental Psychology: Perception and Performance,* 1975, **1**, 374-382.

ROSCH, E., MERVIS, C. B., GRAY, W. D., JOHNSON, D. M., & BOYES-BRAEM, P. Basic objects in natural categories. *Cognitive Psychology,* 1976, **8**, 382-439.

SELFRIDGE, O. G. Pandemonium: A paradigm for learning. In D. U. Blake & A. M. Uttely (Eds.), *Proceedings of the symposium on the mechanization of thought processes.* London: Her Majesty's Stationary Office, 1959.

SHAW, R. E., McINTYRE, M., & MACE, W. The role of symmetry in event perception. In H. Pick (Ed.), *Perception: Essays in honor of J. J. Gibson.* Ithaca: Cornell Press, 1974.

SHAW, R. E., & WILSON, B. E. Generative conceptual knowledge: How we know what we know. In D. Klahr (Ed.), *Carnegie-Mellon symposium on information processing: Cognition and instruction. Vail, Colorado, June 1974.* Hillsdale, N.J: Erlbaum, in press.

SQUIRES, P. C. Topological aspects of apparent visual motion. *Psychologie Forschung,* 1959, **26**, 1-12.

STROUD, J. M. The fine structure of psychological time. In H. Quastler (Ed.), *Information theory in psychology.* Glencoe: Free Press, 1955.

UHR, L. "Pattern recognition" computers as models for form perception. *Psychological Bulletin,* 1963, **60**, 40-73.