An Affine Group Model and the Perception of Orthographically Projected Planar Random Polygons

JOHAN WAGEMANS, PIET VANDEN BOSSCHE, NORBERT SEGERS,
AND GÉRY D'YDEWALLE

University of Leuven

Starting with Shepard and Cooper's seminal work on the perception of rotated and/or reflected 2D and 3D patterns, a series of chronometric studies have been published about the matching of visual patterns with respect to shape, regardless of differences in size or orientation. These studies are restricted to the group of Euclidean geometry. Using random polygons, we extended this research to the group of affine transformations. From the total group defined by four parameters, subgroups were made by different parameterizations of the basic matrix (e.g., holding some parameters constant or giving them 0 and 1 values). A subgroup on level l is defined as the smallest subgroup enclosing two subgroups of a lower level $l-1$. In this manner, a hierarchical model with five levels is constructed with the total group at the top and the identity at the bottom. The detectability of the different transformations, as measured with reaction times in a *Same/Different* paradigm, was investigated to evaluate the psychological plausibility of this mathematical model. More specifically, we were interested in finding out which of all mathematically equivalent paths described best human visual perception of orthographically projected planar random polygons. Going from the top of the model to the bottom, significant differences were obtained by (1) removing the skewing in the Y -direction, (2) removing the scaling in the Y -direction, and (3) removing the skewing in the X -direction, respectively. © 1994 Academic Press, Inc.

1. INTRODUCTION

One of the central problems in the study of shape perception is the phenomenon of shape constancy; that is, the fact that humans tend to see figures as having the same shape if they differ only in position, orientation, size, and mirror image reflection.

We thank P. Locher, R. Shepard, J. Townsend, K. Verfaillie, and an anonymous reviewer for critically reading earlier drafts of the text. This interdisciplinary collaboration is made possible by grants from the University of Leuven (OT/89/11) and from the Belgian Fund for Joint Basic Research (FKFO). The equipment used is partially sponsored by the Belgian Ministry of Science (AI-Impulse Program) and the University of Leuven (GOA). The revision of the manuscript profited from discussions in the context of the VIVA-project sponsored by the EC (BRA-6448). The first author has a grant from the National Fund for Scientific Research (N.F.W.O. Belgium). Correspondence should be sent to Johan Wagemans, Laboratory of Experimental Psychology, University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium (e-mail: fpaas10@blekul11).

Research about the perception of form undergoing orientational change has been dominated by the concept of mental rotation. There is behavioral evidence suggesting that when processing geometrically transformed shapes, human observers are mentally transforming the internal representations of the shapes in order to make certain decisions about them. For example, Shepard and Metzler (1971) found that the time taken by subjects to determine the identity of two objects was proportional to the angle between them. That is, subjects behaved as though they mentally rotated one of the objects at a constant rate until both were at the same orientation. Likewise, Bundesen and Larsen (1975) reported that the reaction time needed to notice the identity of shapes in the same orientation but of different sizes increased linearly as a function of their size ratio. Again, they suggested that a process of mental size transformation could explain this finding.

Although the literature inspired by these results and theoretical notions has grown rapidly (see Shepard & Cooper, 1982, for a review), it is almost systematically restricted to the set of Euclidean transformations. Nevertheless, a much larger set of transformations seems relevant to the human perception of shape equivalence. Recent experimental work and theoretical analyses have shown that affine information plays an important role in several perceptual tasks such as viewing pictures and cinema at a slant (Busey *et al.*, 1990; Cutting, 1987, 1988) and recovering structure from motion (Koenderink & van Doorn, 1991; Todd & Bressan, 1990; Todd & Norman, 1991). More specific evidence about the use of regularities which are preserved under affine transformations has been provided as well (Wagemans, 1992, 1993).

Geometries can be distinguished on a systematic basis. A geometry is defined by a group of transformations¹ and its invariants. The group of Euclidean transformations of the plane is generated by rotations, translations, and reflections, which all preserve distance. An image of an object becomes smaller with increasing distance between the object and the image plane. Despite the fact that distances are not invariant under this transformation (hence, it is not Euclidean), some relations are preserved: Transformed distances are constant multiples or fractions of original distances. This case can be described by the group of similarity transformations; that is, the group generated by the Euclidean transformation together with the dilations. Besides manipulation of the distance between the object and the image plane, there are other transformations that do not involve parallelism between the

¹ A group $G, *$ of transformations is a set of transformations equipped with an operation with the following properties:

(1) Closure of $*$: The combination of any pair of transformations leads to another transformation that also belongs to the group. $\forall x \in G, \forall y \in G: (x * y) \in G$.

(2) Associativity of $*$: The transformation can be bracketed as follows, $\forall x, y, z \in G: x * (y * z) = (x * y) * z$.

(3) Identity element: There is a transformation e called the identity element which has no effect. $\exists e \in G, \forall x \in G: x * e = x = e * x$.

(4) Inverse for $*$ in G : For each transformation, there is an inverse transformation that reverses the effect of the first. $\forall x \in G, \exists x^{-1} \in G: x * x^{-1} = e = x^{-1} * x$.

object plane and the image plane. When we rotate a planar figure around any axis and then project it orthogonally onto a fixed plane, the image is called an affine transformed version of the original figure. It is this group of affine transformations that will be explored further in this article. Transformations that destroy even more properties are, for example, projective (including perspective) and topological.

As indicated above, most classic research about perceptual processing of geometrically transformed shapes has been restricted to the set of Euclidean transformations, although some work has focused on size transformation as well (hence, the similarity group). In addition, a major shortcoming of most previous research is that the mathematics of group theory has not been taken very seriously. Recently, two new developments have attempted to remedy these shortcomings. However, each of them has solved only one problem at the expense of sacrificing the other requirement. Our goal in this paper is to formulate a mathematically elegant model for a fairly complex set of transformations. To better appreciate our contribution, we provide some essential background about both of the new developments that have appeared since the publication of the overview of the classic work on shape equivalence (Shepard & Cooper, 1982).

First, Chen and Chen (1982, 1987) proposed a mathematically elegant group model for similarity transformations. The rationale followed by Chen and Chen (1987) is that, within a group of transformations, the smaller a subgroup,² the more invariants there are under its elements, and the more invariants, the easier it should be to recognize the images generated by the specific subgroup. The results confirmed this reasoning in so far that shape equivalence was easier to determine in a *Same/Different* task when sense was preserved than when size was preserved. In other words, the reversal of the sense of a form created more instability in an image than changes in size. The subgroup that preserves sense and verticality maximizes the symmetry and produced the shortest reaction time. Despite the mathematical elegance of this group model and its empirical support, its application is quite restricted. Similarity transformations are but a small subgroup of the more realistic affine transformations.

The second new development is offered by Leyton's work (1987a, 1987b). According to his observations, an arbitrary parallelogram is perceived as a deviation of a square under some affine transformation. Furthermore, the removal of the transformation T is not done in a single step but involves successive stages of prototypification, undoing three transformations from T in a systematic order, namely (1) rotation, (2) shear, and (3) stretch. The affine transformations of the plane considered by Leyton do not include reflection or translation, hence, his

² $H, *$ is a subgroup of the group $G, *$ if and only if:

- (1) H is a subset of G ,
- (2) $H, *$ is a group. $H, *$ is the group generated by an element x of G (respectively, a subset D of G) if and only if $H, *$ is the smallest subgroup of $G, *$ enclosing x (respectively, D).

work is restricted to the group of all real two by two matrices with positive determinant. Each matrix of this group can be described as a product of the following form:

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

This decomposition is known as the Iwasawa decomposition (Leyton, 1987b). The notation implies that rotation $\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$ is applied first, then shear $\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$, and finally stretch $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$. The transformation is written at the left side of the element on which it acts: $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}$. The order is very important because multiplication of matrices is not commutative; for example,

$$\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} a & bc \\ 0 & b \end{pmatrix} \neq \begin{pmatrix} a & ac \\ 0 & b \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}.$$

Leyton (1987a, 1987b) claimed that the successive steps between the rotated parallelogram and the square are indicative of the psychological relevance of the Iwasawa decomposition. To explain the order in which the components are removed, Leyton (1987a, p. 29) introduced the cognitive stability principle: "The cognitive system removes the properties of a stimulus in sequence starting from those properties ranked as most unstable and progressing successively to those properties ranked as most stable." Subjects rated rotation to be the most unstable property of the initial percept, shear as the next most unstable, and stretch as the most stable. Hence, the order of removal is in agreement with the principle.

Although Leyton's work is more general than Chen and Chen's (1982, 1987), because it concerns affine instead of similarity transformations, it is also mathematically less elegant, as we attempt to make clear next. Implicit in this line of reasoning is the assumption that all different orders of removing factors of the Iwasawa decomposition (i.e., a path from top to bottom in Fig. 1) are mathematically equivalent. However, mathematically speaking, removing factors in different orders makes no sense because the multiplication of matrices is not commutative. Applying first a rotation and then a stretch to a square, for example, does not yield a rotated rectangle but a rotated parallelogram. In other words, Leyton would have to permit different (say, shear) transformations to occur when a different order of unravelling is taken. His program only works because it supposes the very same (say shear) transformation that was imposed on the initial object, for which he has, of course, no independent evidence. In conclusion, one cannot factorize the Iwasawa decomposition the way Leyton does. This does not mean, of course, that the Iwasawa decomposition cannot be psychologically relevant, but it is not mathematically grounded.

FIG. 1. All possible factorizations of the Iwasawa decomposition. Taken from "A Limitation Theorem for the Differentiable Prototypification of Shape" by M. Leyton, 1987, *Journal of Mathematical Psychology*, 31, p. 314.

In summary, Chen and Chen (1987) deal with the group of similarity transformations, which is smaller than the group of affine transformations implied by Leyton's work. On the other hand, Chen and Chen proposed a group model: It is a mathematically sound model with subgroups of the total group. In contrast, the Iwasawa decomposition is a decomposition of a particular affine transformation, not of the group of affine transformations. Our research is aimed at combining "the best of both worlds": We want to extend Chen and Chen's (1987) work to the group of affine transformations in a mathematically justified way. We constructed a model of subgroups of the affine group and designed an experiment using polygons transformed by elements of the different subgroups in order to investigate the psychological salience of certain subgroups of transformations. Contrary to Leyton's "model" the different paths from top to bottom in our model can be interpreted as mathematically equally plausible form prototypifications. As Leyton was, we are interested in finding out which path is psychologically the most relevant one.

Before describing the details of our experiment let us start with our affine group model. We consider all affine transformations (excluding translations) that can be represented by the general linear group of two by two matrices:³

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{R}; ad - bc \neq 0 \right\}.$$

³ Some people (including an anonymous reviewer) would doubt whether these matrices constitute real subgroups, because they are defined locally instead of globally. According to that reasoning, the rotations with different centers on the Cartesian plane do not form a group, but the rotations with different centers and translations with different directions and distances do form a group. In contrast, we only considered affine transformations with a (common) fixed point. We excluded translations (as did Chen & Chen and Leyton) because perception of two shapes already involves a translation when a saccadic eye movement is made to bring both subsequently into foveal vision. However, all affine transformations with a fixed point can be represented by the General Linear group of real two by two matrices, $GL(2, \mathbb{R})$, and all matrices in Fig. 2 definitely describe the general form of a subgroup of $GL(2, \mathbb{R})$.

In order to construct a hierarchical model, we start with four subgroups enclosing only one type of transformation, at the lowest level:

$$\left\{ \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \middle| a \in \mathbb{R}_0 \right\}$$

$$\left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \middle| b \in \mathbb{R} \right\}$$

$$\left\{ \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \middle| c \in \mathbb{R} \right\}$$

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix} \middle| d \in \mathbb{R}_0 \right\}.$$

These four subgroups represent, respectively, scaling in the X -direction, skewing in the X -direction, skewing in the Y -direction, and scaling in the Y -direction.⁴ At a higher level, the subgroups are generated by two subgroups of a lower level. A subgroup generated by a set is the smallest subgroup enclosing it. In this way, a group structure is formed with five levels; the total group at the top and the identity at the bottom. Figure 2 shows our hierarchical model with a matrix representing a specific subgroup and the way a square is transformed by it at each node.⁵ Table 1 indicates how any transformation of each subgroup can be written as a product of matrices out of the subgroups that generate the considered subgroup. (Note that these decompositions are not unique.)

EXPERIMENT

In our experiment, the parameters a , b , c , and d of the matrix could take eight different values (0, 1, and three positive and three negative different from 0 and 1). The combinations with zero determinant were excluded, because in that case a dimension would be lost, so that the transformed polygon would yield a line. The

⁴ From a purely abstract, mathematical standpoint of affine transformations in general, there is no reason why rotations, reflections, shearings, stretchings, or translations along or about any particular axes should have a privileged status. However, there is good evidence that the psychological world is not isotropic (e.g., pervasiveness of the oblique effect, cf. Appelle, 1972). Shepard (1984; Shepard & Hurwitz, 1984) has suggested that the special status of the vertical (Y) axis might be based on the uniform gravitational field of the world we have evolved in, whereas the horizontal (X) axis might be (literally!) based on the ground surface on which we move.

⁵ The reader should not be misled by this way of representing the group structure. In essence, it is a graph of chains of subgroups and, therefore, very similar to the technique of relating group theory and perception originally proposed by Chen and Chen (1982, 1987). The only problem is that it is not always clear what the invariant properties are for each subgroup. For example, the skewings in the X -direction have only one eigenspace, the X -axis, with eigenvalue 1: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$. This means that X -skewings do not change anything in the X -direction; they preserve some kind of length in the X -direction. More exactly, they preserve the distance between points with the same Y -coordinate (see Fig. 2). We admit that these invariants and their perceptual relevance are not always intuitively clear.

FIG. 2. Our hierarchical group model with at each node the matrix representing a specific subgroup and a square transformed by an element of it.

TABLE 1

Possible Decomposition of the Matrices that Are Used in the Group Structure

Matrix	Decomposition
$\begin{pmatrix} a & 0 \\ c & 1 \end{pmatrix}$	$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$
$\begin{pmatrix} 1 & 0 \\ c & d \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix}$
$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$
$\begin{pmatrix} 1 & b \\ 0 & d \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$
$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}$	$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix}$
$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$
$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$	$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & d - bc/a \end{pmatrix} \begin{pmatrix} 1 & b/a \\ 0 & 1 \end{pmatrix}$

determinant could be positive or negative, which implies that reflections were allowed. Note that reflection was not allowed in the experiments of Leyton (1987a, 1987b).

At the lowest level of the group structure (excluding the identity), each matrix could take six different values for, respectively, a , b , c , or d which resulted in 4×6^4 transformations for that level. At the next level, we can have 6^2 possible combinations for each matrix. At the third level there are only two matrices with each 6^3 possible transformations, so there were 432 possible transformations at this level. At the level of the total group, the factors a , b , c , and d could take only four values (two positive and two negative), because otherwise the number of trials would be too large. As a result, at the level of the total group there were 4^4 transformations possible. This yielded a total of 601 transformations in the case of the subgroups and 256 for the total group.

Method

Subjects. As a consequence of the large number of trials, only a restricted number of well-motivated subjects could be run. In addition to the first two authors, two other research assistants, involved in totally different research topics,

FIG. 3. Five examples of the random polygons used in the experiment.

volunteered. They were naive about the purpose of the experiment. All observers had normal or corrected-to-normal vision.

Stimuli. The planar figures used for the transformations were ten random polygons created by connecting 24 coordinates generated by a method described in Wagemans *et al.* (1991, 1992). Figure 3 shows five random polygons used in the experiment. Before transforming the polygons, they were randomly given one of 35 possible orientations between 0° and 360° to avoid possible bias effects caused by possibly intrinsic orientations of the polygons (as a consequence of the procedure used to connect the coordinates). Due to software restrictions, the random polygons were divided into two sets of five. Combination resulted in a total of $5 \times 35 = 175$ possible stimuli per set. Each subject was randomly assigned one of the two stimulus sets in each block, and at each trial one of the possible 175 patterns was randomly selected to be transformed using one of the possible parameterizations of the matrices formulated above. Each transformation was repeated five times, with a new random polygon for each repetition.

Apparatus. The coordinates of all patterns were stored in separate files that could be accessed by a Turbo Pascal program steering the experiment on an IBM-AT-compatible with a Phoenix 80386 processor and a Hercules graphics board. They were presented on a monochrome screen with white phosphor against a uniform blackground. Subjects were tested individually. They were seated at a distance of 114 cm from the screen. The experimental room was completely darkened and indirect light reflecting from the screen was prevented by a cardboard with a circular opening attached to the screen. All this was done to avoid visual reference frames. The individual patterns, and the whole stimulus area (both patterns plus the separating space) subtended respectively 2° and 6.5° of visual angle. A headrest was used to prevent head rotation.

Procedure. The experiment was divided into two major parts: one for the subgroups and one concerning the total group. In each trial, two stimuli were simultaneously projected side by side on two halves of a computer screen. One of the stimuli was an untransformed random polygon and the other was a transformed version of the same or another polygon. The subject's task was to decide whether the two patterns were the same or not, regardless of an affine transformation between them. Note that this task differs from the one used in the classic Shepard and Cooper experiments, where subjects had to discriminate between a rotation in the picture plane and a reflection of the same polygon. The total number of trials

for the subgroups is $601 \text{ transformations} \times 5 \text{ patterns} \times 2 \text{ (Same/Different)} = 6010$ trials, and for the group $256 \times 5 \times 2 = 2560$ trials. The trials were divided into 23 sessions of 261 or 262 trials for the subgroups, and 10 sessions of 256 trials for the group. Each session took approximately 25 minutes. The order of the sessions was randomized for each subject. One author and one naive subject first did the sessions concerning the group, followed by the sessions concerning the subgroup, while the other two subjects received the reverse order.

Before starting the experiment, the task was clarified to the naive subjects with a square and a random quadrangle, undergoing different exemplars of the set of transformations. In addition, subjects were given an exercise session with the random polygons used in the experiment. Each trial was preceded by a fixation cross that was presented for 750 ms in the middle of the screen. Then, after a 500-ms interval, the two stimuli were presented. Subjects had to make Same/Different responses by pushing one of two response buttons on a panel connected with the PC configuration. When the answer was correct a high-frequency tone (660 Hz) was given, in the case of a wrong answer a low-frequency tone (440 Hz) was given. The only reason why this feedback was provided was to keep the motivation and arousal of the subjects at an optimal level. Stimuli remained on the screen until an answer was given. Both reaction times and accuracy were registered. In case of a wrong answer, the trial was retaken at a random position in the rest of the session.

Data analysis. Because we were interested in the effect of the kind of affine transformation, represented by one of the ten subgroups and the total group, on the reaction time needed to consider the affine equivalence of the two shapes, we tested a restricted number of a priori contrasts using the error term yielded by an analysis of variance (ANOVA) of the response time data in a randomized block design (df_e is 17125). First of all, we looked for significant differences between average reaction times at the different levels in our hierarchical model. Second, we compared the reaction times for subgroups that were connected by a path in the group structure.

RESULTS AND DISCUSSION

The mean reaction times associated with the transformations at each level in our hierarchical model differed from each other. (Level 1 refers to the four basic matrices that were used to construct our group model; Level 4 refers to the the matrix that represents the total group). A priori comparisons showed that the average response time for Level 1 ($\bar{X} = 1178$ ms, $SD = 590$, $N = 480$) was smaller than that for Level 2 ($\bar{X} = 1285$ ms, $SD = 733$, $N = 2880$), which was smaller than that for Level 3 ($\bar{X} = 1347$ ms, $SD = 780$, $N = 8640$), which did not differ reliably from that for Level 4 ($\bar{X} = 1352$ ms, $SD = 757$, $N = 5120$): $F = 8.67$, 15.79 , and 0.17 , respectively, with $p < 0.003$, 0.0001 , and 0.68 .

In summary, a pattern generated by a subgroup representing a basic transformation of our hierarchical model is recognized faster as a transformed version of the target stimulus than is a figure that is generated by a subgroup representing a combination of two of these basic matrices. Likewise, a figure generated by a subgroup representing the combination of two of the basic matrices yields faster responses than a stimulus generated by a subgroup representing a combination of three of the basic matrices. Recognizing shape equivalence of a planar random polygon transformed by an element of the total group is not reliably more difficult. Figure 4 represents comparisons of reaction times for stimuli generated by subgroups that are connected by a path in our hierarchical group model.

FIG. 4. Comparison of mean reaction times for the same responses between transformed stimuli and target stimuli, for the stimuli that were generated by subgroups that are connected by a path in our hierarchical group model. $\bar{X}$, mean reaction time; SD, standard deviation; N, number of observations; n.s., not significant.

These results show that—in order to judge equivalence of shapes related by the most general affine transformation (i.e., generated by the total group)—undoing skewing in the *Y*-direction gives faster reaction times ($F = 16.97$, $p < 0.0001$), while undoing skewing in the *X*-direction did not make it easier. The reaction times for the latter are even longer ($F = 11.72$, $p < 0.0006$). When we undo the subgroup that gives a significant improvement in reaction time compared with the total group, from scaling in the *Y*-direction, we also found an effect ($F = 4.61$, $p < 0.03$), but not when we undo it from scaling in the *X*-direction ($F = 1.22$, $p < 0.3$). The next (marginally) significant difference in reaction time is found when we undo the subgroup that represents scaling in the *X*-direction and skewing in the *X*-direction, from skewing in the *X*-direction ($F = 3.12$, $p < 0.08$), but not when we undo it from scaling in the *X*-direction ($F < 1$).

In summary, going from the top of the group structure to the bottom, the significant manipulations are, respectively, (1) removing the skewing in the *Y*-direction, (2) removing the scaling in the *Y*-direction, and (3) removing the skewing in the *X*-direction. Following Leyton's rationale, we could conclude that, although each path in our hierarchical group model is mathematically equivalent, only one route seems relevant for human perception.

Reformulating our results in Chen and Chen's terms, we found that within the group model, the smaller the subgroup, the easier it is to recognize the images generated by it. More specifically, the images generated by the subgroup representing scaling in the *X*-direction are recognized faster than the images generated by the subgroup representing the combination of scaling and skewing in the *X*-direction. The images generated by the latter are recognized faster than those generated by the subgroup representing a combination of scaling and skewing in the *X*-direction and a scaling in the *Y*-direction, which in turn yield smaller average response times than images generated by the total group. So, Chen and Chen's (1987) conclusion, based on the group of similarity transformations, that the smaller the subgroup, the more invariants there are under its transformations (see Note 4), and the more invariants, the easier it is to recognize the images generated by the specific subgroup, seems to hold for the larger set of affine transformations investigated here also.

FINAL REMARKS

Of course, one might wonder why it would be adaptive for our perceptual system to work in the way discovered here in the more natural setting in which we have evolved. As Shepard (1981, 1984) has forcefully argued, shape constancy is important because it enables us to recognize a rigid object in the 3D world as the same despite variations in the relative orientation and distance between the object and the viewer. Considered from this standpoint (as Shepard did in reviewing our paper), the patterns (planar polygons), transformations (affine, corresponding with parallel projection), and task (Same/Different judgments) used in this experiment constitute only a very special case.

However, there is good evidence that our perceptual system often ignores the difference between orthographic (or parallel) projection used here and perspective (or central) projection occurring in real vision (e.g., Stevens, 1980; Todd, 1984; Ullman, 1979). With respect to the patterns themselves, we are currently exploring arbitrarily curved, closed contours instead of polygons, and the basic results seem pretty similar. Finally, the task could be adapted to capturing more basic processes (i.e., those which seem more deeply engrained in our perceptual system). For example, we could design experiments with Shepard's apparent motion paradigm (1981; Farrell & Shepard, 1981), testing whether the time intervals between two subsequently presented affine equivalent shapes could be shorter when they are related by a simpler transformation according to our group structure. Even finer estimates could be obtained by spatio-temporal probing (Robins & Shepard, 1977) or the window technique (McBeath & Shepard, 1989). Applying and extending these paradigms to more realistic (nonplanar) shapes seems very promising as a means of evaluating the perceptual relevance of group theory. However, it goes without saying that this constitutes a much more ambitious research project which will not be realized on the short term. Our main purpose was to show that it seems a fruitful route to follow. As such the results presented here only add another source of evidence for a transformational approach to visual perception, which is consonant with Shepard's own long-term research achievements (e.g., Carlton & Shepard, 1990a, 1990b).

REFERENCES

- APPELLE, S. (1972). Perception and discrimination as a function of stimulus orientation: The "oblique effect" in man and animals. *Psychological Bulletin*, **78**, 266-278.
- BUNDESEN, C., & LARSEN, A. (1975). Visual transformation of size. *Journal of Experimental Psychology: Human Perception and Performance*, **1**, 214-220.
- BUSEY, T. A., BRADY, N. P., & CUTTING, J. E. (1990). Compensation is unnecessary for the perception of faces in slanted pictures. *Perception & Psychophysics*, **48**, 1-11.
- CARLTON, E. H., & SHEPARD, R. N. (1990a). Psychologically simple motions as geodesic paths. I. Asymmetric objects. *Journal of Mathematical Psychology*, **34**, 127-188.
- CARLTON, E. H., & SHEPARD, R. N. (1990b). Psychologically simple motions as geodesic paths. II. Symmetric objects. *Journal of Mathematical Psychology*, **34**, 189-228.
- CHEN, M., & CHEN, K. C. (1982). A transformation analysis of form recognition under plane isometries. *Journal of Mathematical Psychology*, **26**, 237-251.
- CHEN, M., & CHEN, K. C. (1987). A group model of form recognition under plane similarity transformations. *Journal of Mathematical Psychology*, **31**, 321-337.
- CUTTING, J. E. (1987). Rigidity in cinema seen from the front row, side aisle. *Journal of Experimental Psychology: Human Perception and Performance*, **13**, 323-334.
- CUTTING, J. E. (1988). Affine distortions of pictorial space: Some predictions for Goldstein (1987) that La Gournerie (1859) might have made. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 305-311.
- FARRELL, J. E., & SHEPARD, R. N. (1981). Shape, orientation, and apparent rotational motion. *Journal of Experimental Psychology: Human Perception and Performance*, **7**, 477-486.
- KOENDERINK, J. J., & VAN DOORN, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America A*, **8**, 377-385.

- LEYTON, M. (1987a). Nested structures of control: An intuitive view. *Computer Vision, Graphics, and Image Processing*, **37**, 20–53.
- LEYTON, M. (1987b). A limitation theorem for the differentiable prototypification of shape. *Journal of Mathematical Psychology*, **31**, 307–320.
- MCBEATH, M. K., & SHEPARD, R. N. (1989). Apparent motion between shapes differing in location and orientation: A window technique for estimating path curvature. *Perception & Psychophysics*, **46**, 333–337.
- ROBINS, C., & SHEPARD, R. N. (1977). Spatio-temporal probing of apparent rotational movement. *Perception & Psychophysics*, **22**, 12–18.
- SHEPARD, R. N. (1981). Psychophysical complementarity. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 279–341). Hillsdale, NJ: Erlbaum.
- SHEPARD, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, **91**, 417–447.
- SHEPARD, R. N., & COOPER, L. A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.
- 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*, **171**, 701–703.
- STEVENS, K. A. (1980). *Surface perception from local analysis of texture and contour* (Technical Report 512). Cambridge, MA: MIT AI Lab.
- TODD, J. T. (1984). The perception of three-dimensional structure from rigid and nonrigid motion. *Perception & Psychophysics*, **36**, 97–103.
- TODD, J. T., & BRESSAN, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics*, **48**, 419–430.
- TODD, J. T., & NORMAN, J. F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception & Psychophysics*, **50**, 509–523.
- ULLMAN, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- WAGEMANS, J. (1992). Perceptual use of nonaccidental properties. *Canadian Journal of Psychology*, **46**, 236–279.
- WAGEMANS, J. (1993). Skewed symmetry: A nonaccidental property used to perceive visual forms. *Journal of Experimental Psychology: Human Perception and Performance*, **19**, 364–380.
- WAGEMANS, J., VAN GOOL, L., & D'YDEWALLE, G. (1991). Detection of symmetry in tachistoscopically presented dot patterns: Effects of multiple axes and skewing. *Perception & Psychophysics*, **50**, 413–427.
- WAGEMANS, J., VAN GOOL, L., & D'YDEWALLE, G. (1992). Orientational effects and component processes in symmetry detection. *Quarterly Journal of Experimental Psychology*, **44A**, 475–508.

RECEIVED: January 7, 1991