

# The recognition of geometrically transformed text

PAUL A. KOLERS

RESEARCH LABORATORY OF ELECTRONICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY

*In two experiments it was found that the difficulty of recognizing geometrically transformed text varied with the transformation imposed upon it. In one experiment the text was in English and the Ss were native speakers of English; in the second, the text was in Hebrew and the Ss were native speakers of Hebrew. The two languages are read in opposite directions. Nevertheless, the order of difficulty of the transformations was similar with the two texts. The results are interpreted to show that the visual systems contains "transformation operators" as well as the more widely publicized "detectors."*

The term "object constancy" refers to the psychological process by which an object is perceived in a standard way despite variations in its image on the eye. In the case of size constancy, for example, the apparent or perceived size of a familiar object is found to change little even though its retinal image almost doubles when the distance between the perceiver and the object is halved. Past usage has emphasized the maintenance of a perceptual identity of the standard and transformed arrays, but the term "constancy" need not be restricted in this way, especially since we often speak of more-or-less constancy, of overconstancy and underconstancy. One can regard constancy as having dimensions, and as a special case of recognition. It is the recognition of transformed arrays that will be considered here.

The mechanism by which the constancies are achieved have not been explored much. The general opinion is consistent with Helmholtz' thesis that the constancies are the result of cognitive variables involving judgment and estimation. Conversely, some constancies may reflect the action of specific visual mechanisms. The recent discovery of visual mechanisms that react selectively to objects of particular configuration or in particular orientations (Lettvin et al, 1959; Hubel & Wiesel, 1965), and the demonstration that the orientation of one object can affect the perceived orientation of another (Howard & Templeton, 1966; Sekuler, 1965; Gilinsky, 1967), suggest that the orientation of an object may—like its brightness or contour—be a primitive characteristic used in the construction of its perceptual representation.

In the cases cited, the selective response to an object in one orientation, and the interaction between contours in different orientations, attest the existence of geometry-specific "detectors"; that is,

parts of the visual system tuned to respond to specific orientations. The achievement of constancy could not depend only on such detectors, however, for they would be incapable of supplying the scale transformations or contour transformations whose existence the occurrence of constancy implies. Some "higher-order" mechanism is required that intervenes between the response of the geometry-specific detectors and constancy. Two experiments reported in this paper show that the recognition of transformed objects is mediated in part by the operation of just such higher-order perceptual mechanisms.

## READING TRANSFORMED ENGLISH

Most studies of adaptation to a transformation have been performed with lenses or prisms (Held & Freedman, 1963; Harris, 1965; Rock, 1966). In these, the physical objects retain their identity, but their appearance is altered by the optics, and S's recognition of objects or adjustment to the alteration is measured. The approach followed here uses no optics. Rather, part of the physical world is transformed, and S's recognition of the transformed objects is measured. The objects transformed are passages of connected discourse, and the recognition is measured by the time S takes to read them aloud.

## Materials

Passages approximately 310 words long were prepared in each of eight geometric transformations. In one subset of four, the transformations were rotations in three-dimensional space: normal text (N), inversion (I), reflection (M), and rotation in the plane of the page (R). In the second subset, letter reversal (r) was applied to the preceding four, to make rN, rI, rM, and rR. Each passage of 310 words was prepared in only one transformation, and eight passages were prepared in each transformation. The passages were typed 26 lines to a page in IBM Courier font. All 64 passages were taken from a single source (Miller, 1962) in order to minimize differences in the text attributable to context, style, vocabulary, and the like. Examples of actual text appear in Fig. 1, where asterisks have been added to indicate the beginning of a line.<sup>1</sup>

## Procedure

An 8 by 8 Latin square was constructed in such a way that on each day S read one page of each of the eight transformations, but in a different order

N \*Expectations can also mislead us; the unexpected is always hard to  
perceive clearly. Sometimes we fail to recognize an object because we

R Emerson once said that every man is as lazy as he dares to be. It is  
kind of mistake a New England Puritan might be expected to make. It is

I formed a habit of reading a few pages of a book every day. It is  
of course a habit of reading a few pages of a book every day. It is

M \*Several years ago a professor who teaches psychology at a large  
university had to ask his assistant, a young man of great intelligence

r N \*On his first day in school he was thoroughly disoriented.  
He went to school for the first time; when he saw the school he felt

r R a very large number of people of different ages and  
different heights and different colors and different shapes.

r I to be determined as to whether the text was  
to be read in the same way as the text was

r M \*elap a no elcric der thgirb a swohs enO .serutcip tnereffid owt enigamI\*  
.dnuorgkcab yarg a no elcric neery thgirb a rehto eht ,dnuorgkcab wolley

Fig. 1. Examples of transformed text. Each pair of lines is a single transformation, coded at the left. The asterisks indicate where to begin each pair.

from other days. Across the eight days of testing, each kind of transformation preceded and followed every other kind exactly once. Cycling such a Latin square yields a "Latin cube," which is a design that insures that every S reads all eight pages of each transformation, but in a way that staggers the pages across days and changes the sequence of transformations within days.<sup>2</sup> This complexity of design was introduced to control statistically for the fact that the various passages could not be regarded a priori as of equal difficulty. In order to complete the balance of the cube, on each day S read half a page he had read the preceding day in the transformation he was to be tested with first.

Clear plastic envelopes were marked with an arrow to indicate the direction in which to read the enclosed page. The envelopes were placed singly in a holder in front of S who, following printed instructions, read aloud as rapidly and accurately as he could. In preliminary trials we found that Ss use different strategies in this kind of task which vary the latency of their beginning to read. Consequently, reading time was measured with a stopwatch from the uttering of the last word on the first line of text; thus, the data to be presented are based on 25 lines of text per page.<sup>3</sup>

The eight test days were preceded by two practice days. The Ss were scheduled so that weekends

occurred at different stages in the 10-day sequence, and were tested individually.

### Subjects

Thirty-two Ss, lower classmen from MIT and Harvard University, were tested in four replications of the Latin cube. All of them reported that they had normal or corrected vision. After the testing, the variables of handedness and reading knowledge of Semitic languages were found to have influenced performance. In this experiment, however, the only selection criterion used was that S be able to read a page of normal English text aloud in less than 2 min. Three potential Ss were excluded for failing to meet this criterion.

### Results

The amount of time taken to read pages aloud varies sharply with the transformation and decreases with practice, in the manner of the familiar "learning curve." The geometric mean times taken by the 32 Ss are shown in Fig. 2 on a double logarithmic scale. The curves shown there can be represented as power functions of the form  $\tau = kD^{-\alpha}$ , where  $\tau$  is time, D days (i.e., pages read),  $\alpha$  the slope of the line, and k a constant indicating the y-intercept. The least-squares value of  $\alpha$  for the individual transformations, which measures the rate

of improvement, varies from -0.03 for N to -0.30 for I. The near-zero value for N is to be expected from skilled readers; I is the transformation on which maximum improvement occurs.

The arithmetic means and variances of the time data are closely correlated. Consequently, the analysis of variance was performed on their logarithms (the geometric means are shown in Fig. 2). It yielded significant F-ratios for days, transformations, and subjects ( $p$  well below 0.01). The  $p$ -value for order, however, was greater than 0.10, indicating that the counterbalancing in the Latin cube succeeded in preventing practice on any one transformation from affecting performance on any other in a significant way. A critical difference was computed from the error mean square of the analysis of variance. A difference between any two points on the left ordinate in Fig. 2 greater than 0.026 is significant at the .05 level, and a difference greater than 0.034 is significant at the .01 level.

The order of difficulty of the transformations is of theoretical interest. Four transformations—N,

rN, R, and rR—occupy ranks 1, 2, 3, and 8 for 90% of the Ss on a daily basis. The filling of ranks 4-7 is equivocal, however, because of the differences in the rate at which the various transformations are mastered. Kendall's (1948) coefficient of concordance was used to evaluate the order of difficulty of the transformations; that is, the time taken to read eight pages of each of the eight transformations was correlated across the 32 Ss. The result,  $W = 0.814$ , shows that the Ss "agree" markedly, but not perfectly, on the order of difficulty of the transformations.<sup>4</sup> The opposite correlation, the rank-ordering of Ss by transformations, yielded  $W = 0.735$ , showing some tendency for Ss who read one transformation rapidly to read others rapidly as well.

### Discussion

It is not a new finding that mathematically equivalent transformations such as R, I, and M affect the perceiver differentially (Smith & Smith, 1962; Howard & Templeton, 1966). Here, the differences are revealed by the order of difficulty inferred from

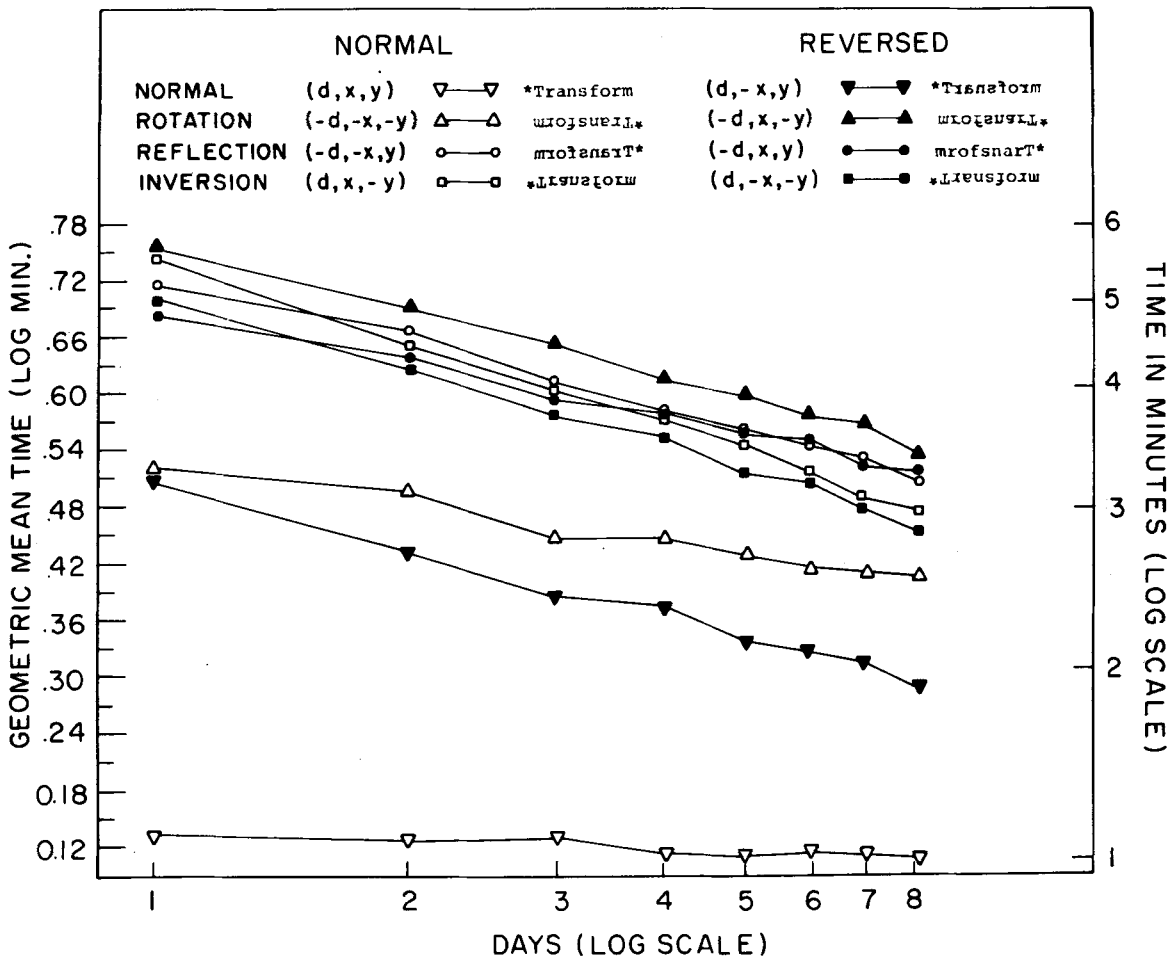


Fig. 2. Reading time for transformed English. The legend on the figure gives the orientation of the individual characters in terms of the horizontal (x) and vertical (y) cartesian coordinates, and d indicates the direction of reading.

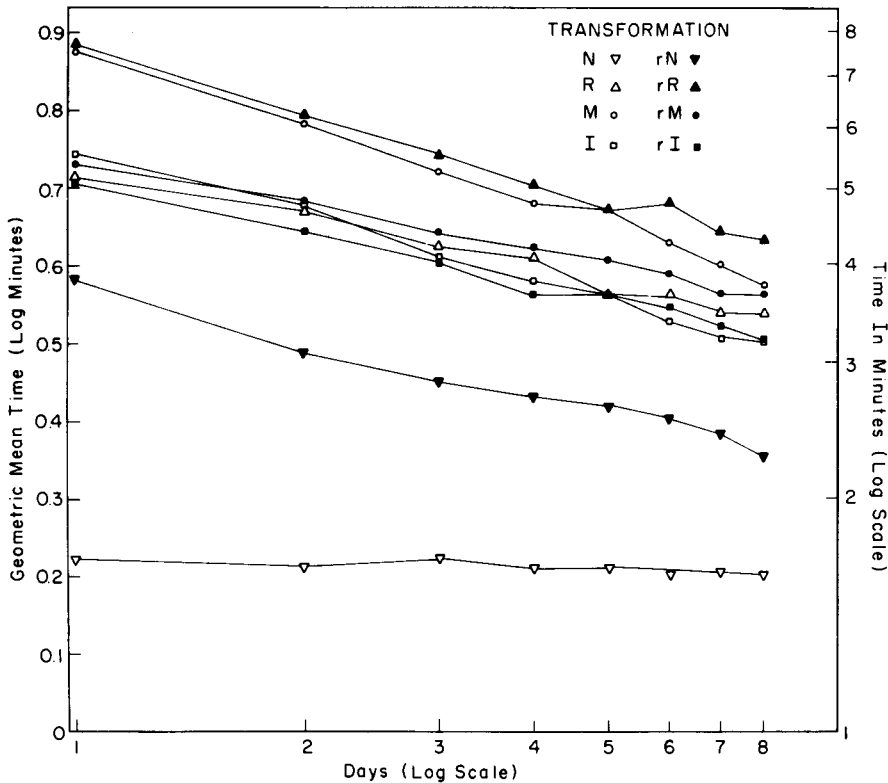


Fig. 3. Time taken by Israeli Ss to read transformed English.

the time required to read them. In addition, however, the data of Fig. 2 show that the transformations do not combine their effect on the reader in any mathematically simple way, such as by addition or multiplication. For example, the two transformations that are easiest to read (R and rN) produce in combination the transformation that is most difficult (rR). In some cases, the double transformation (rotation plus letter reversal) is easier to read than the simple rotation alone, as when rI is easier than I and rM is easier than M. It has proved to be impossible to solve the general equation  $T_i = T_1 * T_2$ , where  $T_i$  is the amount of time taken for any transformation,  $T_1$  and  $T_2$  are any two other transformations that might compose  $T_i$ , and  $*$  is any arithmetic operation. It has proved to be impossible also to solve the equation of the form  $T_i = O_1 * O_2 * O_3$ , where the Os are any binary operators such as letter reversal, rotation, reflection, or the like. Mathematically, the eight transformations may be represented as an Abelian group generated by three binary operators. The data obtained with them, however, cannot be represented in a similar way. The reader operating on the transformations of Fig. 1 clearly does not decode them into geometric constituents and operate on the constituents in a serial way. The implication of this result is that the "features" the perceptual system uses to decode the texts are different from the geometrical features we use to describe them. Hence, analysis of the

texts in terms of geometrically distinctive features has little relevance for understanding their perception. A proof that the interaction of three or four variables cannot represent the data will be presented in another place.

The failure of arithmetic interaction to account for the data involves the variable of letter reversal (r) mainly. The order of difficulty of the tridimensional rotations, R, I, and M, is remarkably stable across Ss, and it is this finding that is the main concern of the present paper. The order, however, appears to be sensitive to particular kinds of past experience, three of which will be described.

Semitic languages, such as Hebrew and Arabic, are written and read right to left. After the experiment was completed, it was found that ability to read a Semitic language had affected performance on the transformations of English. Many Ss with the ability took less time to read M than I, while Ss without the ability usually took less time for I than M. Since fewer than one-third of the Ss had the ability to read Semitic languages, and not all to the same degree, the order of difficulty R, I, M holds for the greater number of Ss. However, 23% of the Ss with no reading knowledge of left-going languages also took less time for M than I. Thus, here the effect of training is masked by other factors.

In another experiment, it was found that native speakers of English who were concentrating in Middle Eastern studies and had acquired a good

\* פעם אחת ישבנו אני ורבי חיים בביה המדרש ולא היה אדם  
 שם חוץ ממנו. ראיתי שמרתיע וקם וחוזר ויושב, וחוזר וקם N

\* אגזזזז טעהז \* לזזזז אגזזזז אגזזזז אגזזזז אגזזזז אגזזזז  
 אגזזזז אגזזזז אגזזזז אגזזזז אגזזזז אגזזזז אגזזזז אגזזזז R

\* סזא חזח אלז שרזחח חזחח סזזח זרזז זרזז זרזז זרזז זרזז זרזז זרזז זרזז  
 סזזז זרזזז, אשזזזז זרזזזז סזזז זרזזזזז זרזזזז זרזזזז זרזזזז זרזזזז M

\* גז טעזא טעקכז טדלגז כדלגז \* גקטז גז טדלגז טזא נקגז  
 זקגז \* טקטז טקז טעזז טקטז טקטז טקטז טקטז טקטז טקטז טקטז I

\* סיאלמ חונליאהו חונליא סיאלמ סיסדרפה. וזננינקל רוזחנז  
 סילס קוהל סיפטוקו תורוהבו סירוהב סהל סידמוע. תורזפ rM

\* נזננג אמזנז קטטז טזנכז טזנכז טזנכז טזנכז טזנכז טזנכז טזנכז טזנכז  
 טזנכז טקטז טזנכז טזנכז טזנכז טזנכז טזנכז טזנכז טזנכז R

Fig. 4. Examples of transformed Hebrew. The asterisks indicate the beginning of each pair.

reading knowledge of Semitic languages read M in far less time than other native speakers of English (Kolers, in press a). This is evidence that skill in reading in a particular direction can transfer both across languages and across alphabets. However, the particular role of experience remains unclear.

In an effort to clarify the issue, native speakers of Hebrew, Israeli graduate students at MIT, were tested for their ability to read transformed English. Eight Ss were tested through one cube of the statistical design. The findings with these Ss were that I was read in less time than R, and both were read in less time than M (Fig. 3). The relative ease of M for native speakers of English who read Semitic languages is confusing at this point, but the reversal of I and R among Israeli Ss raises an interesting and testable question that is examined below.

**READING TRANSFORMED HEBREW**

The question of interest is whether the order of difficulty of the transformations described above is the result of specific orientation-sensitive mechanisms in the visual system, or whether it is the result of higher-order mechanisms concerned with transformations of visual displays. The question was studied by requiring readers whose skill was restricted to left-going languages to read transformations of those languages analogous to the transformations of Fig. 1. If the biases revealed in the data of Fig. 2 were the result of the action of orientation-specific mechanisms, the order of difficulty of reading

transformed English would be different from the order of difficulty for transformed Hebrew, for example. The difference would show up specifically in the comparison of inverted text (I) and text rotated in the plane of the page (R). The reason is as follows.

Rotation in the plane of a page of a right-going language produces characters that face to the left, are inverted top to bottom, and read leftward, as illustrated by R in Fig. 1. Inversions of a right-going text produce characters that face rightward, run top to bottom, and read rightward, as in I. A rotation in the plane of a page of left-going text, such as Hebrew, produces characters that face to the right, are inverted, and read rightward, as in R of Fig. 4. An inversion of Hebrew produces characters that face to the left, are inverted, and read leftward, as in I of Fig. 4. Thus, a rotation in the plane of a page of English produces a geometry formally identical to an inversion of Hebrew, and an inversion of English produces a geometry formally identical to rotation in the plane of the page of Hebrew. If the data of Fig. 2 were the result of orientation-detecting mechanisms, inverted Hebrew should be easier to read than rotated Hebrew; however, if the results are due to the action of transformation-operators in the visual system, rotated Hebrew should be easier to read than inverted Hebrew, as is the case with English. The distinction being drawn is between orientation-specific detectors and transformation-specific operators.

**Method**

Hebrew passages 100 words long were taken from a single novel by S.Y. Agnon and were prepared in all of the transformations illustrated in Fig. 1. This experiment was designed to complement the former one with English, but for reasons beyond control had to be curtailed. Only six transformations were used, and the testing was performed with six Ss. Testing followed the design of a 6 by 6 balanced Latin cube, but the Ss were tested for only five days.

**Subjects**

Ss who read Hebrew or Hebrew and Arabic well and had only a rudimentary skill with European languages (less than the second-grade level) were located in a religious vocational school in Israel.<sup>5</sup> Only six Ss were found who met the criteria of reading skill for Semitic languages and age range (16-20).<sup>6</sup>

**Results**

The data shown in Fig. 5 indicate that the order of difficulty of the transformations of Hebrew is similar to the order for English. Normally-oriented Hebrew required the least time, and of the transformations on it, R is easiest and is followed by I and M. Thus, the order for the simple rotations is identical in the two languages. The one discrepancy in the data for the two languages occurs with the transformations involving letter reversal (r): Readers of English usually require more time for rR than for rM, but the reverse is true for the data

of Fig. 5. The significance and reliability of this reversal are unknown. The main finding is the similarity of the order of difficulty of transformations imposed on right-going and left-going languages. Thus, the same transformations applied to different bases yield the same results. It is the transformation imposed and not the specific geometry of the characters that determines the order of difficulty.

**Qualitative Observation**

A particular kind of error occurred in the reading of both the transformed Hebrew and the transformed English. It is an instance of a directional bias rather than an orientational bias. It is tangential to the demonstrations above, but has significance for a theory of reading and perceptual recognition, and so will be described briefly. To understand the import of the observation, it is necessary to know that seven letters of the Hebrew alphabet have distinctive terminal forms. The difference between normal and terminal form is expressed by the rule valid for six of the pairs that an open horizontal at the bottom of a normal character is replaced by a vertical extending below the line in a terminal character. The seventh letter, mem, violates this rule, but is different in its two forms in other ways.<sup>7</sup> As illustrations, consider the following pairs of words. The name "Moshe (h)" is written with three Hebrew characters, mem, shin, heh, and the word "Hashem" is written heh, shin, terminal mem. The word "maon" is written mem, ayin, vav, terminal nun, and "noam" is written nun, vav, ayin, terminal mem. Finally,

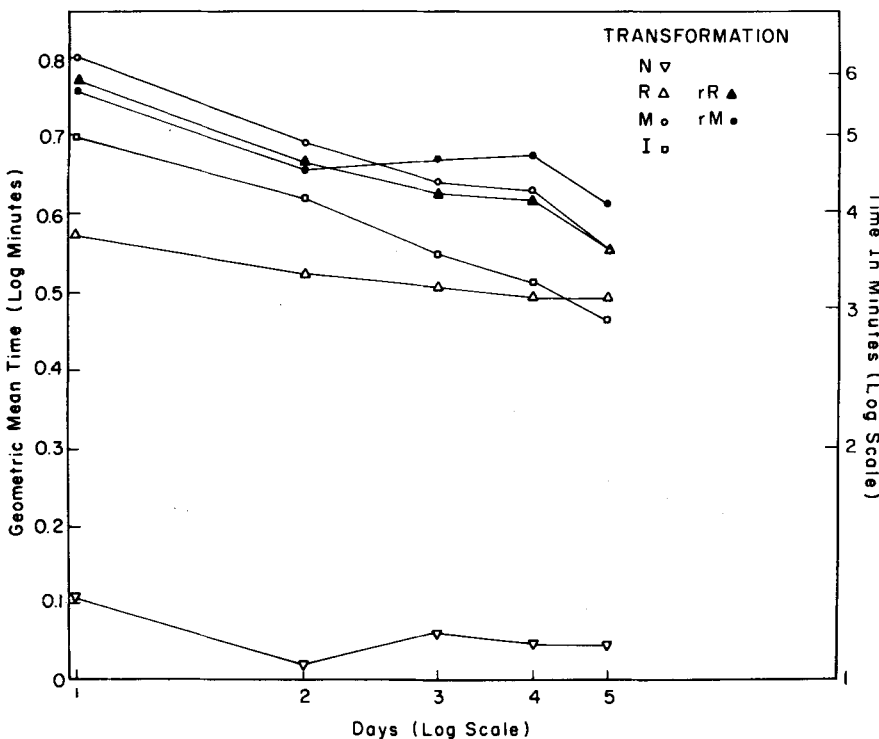


Fig. 5. Reading time for transformed Hebrew.

"alats" is written ayin, lamed, terminal tsadi, and "tsala" is written tsadi, lamed, ayin. The appearance of mem, nun, and tsadi are quite different in the normal and terminal forms and, of course, different from each other. Despite these differences in appearance, and despite the rule that a terminal letter never legally can be used to begin a Hebrew word (any more than a capital letter can end an English word), the Ss sometimes read "Hashem" as "Moshe," "maon" as "noam," "tsala" as "alats," and so on.

This kind of error is not restricted to Hebrew; a similar error occurred with the reading of English in transformation rM when the Ss read "was" as "saw," for example, or "ton" as "not." Errors of this kind occurred only in transformation rM, a transformation in which individual characters appear in their normal orientation but the order of reading them is reversed. It has been shown that skilled readers of English can read rM aloud more rapidly from left to right as nonsense than from right to left as sense (Kolers, in press b). In the present context this directional bias is shown to be strong enough that it induces the reader of Hebrew to violate a rule of the language (that words cannot begin with terminal letters). The error raises an interesting problem regarding the relation between recognition of objects and motions of the eye.

### Discussion

Biases of form perception, representing a spatial polarization of the visual system, have been documented for many lower animals (e.g., Sutherland, undated). Polarizations have been documented for humans as well, most notably in the right-left and up-down confusions of children (Ghent, 1961; Howard & Templeton, 1966). Gilinsky (1967) has recently shown that some of these polarizations can extend over time sufficiently that prior inspection of a grating in one orientation affects the perceptibility of a second grating, the direction and magnitude of the effect varying with the relative orientation of the two. There is therefore little question that the orientation of objects strongly affects their recognition. Such biases are implicated also in the findings that equivalent mathematical transformations of text require different amounts of time to be decoded (Figs. 2 and 5). The identity in the order of difficulty of transformations across languages that normally go in opposite directions, however, implies a mechanism of a higher-order than orientation-sensitive detectors. Such a higher-order mechanism appears to be concerned with the recognition of visually transformed objects, yet is object-independent, in contrast to object-dependent detectors. That is to say, the typical "feature-detector" is selectively sensitive to objects in specific orientations, making it geometry-specific. The mechanisms revealed in the present experiment, by contrast, are

transformation-specific rather than object-specific: The mechanisms appear to be indifferent to the specific geometry of elements they operate on but perform identical operations upon them. Transformation operators of this kind would seem to be significant in object constancy and in the adaptation to optically transformed arrays.

There are other behavioral implications to the directional biases reading imposes that may only be mentioned. For example, European and American card catalogs usually begin at the upper left and proceed to the lower right. Israeli card catalogs, in contrast, go from upper right to lower left. Zippers on American luggage usually open from left to right; on Israeli luggage they open from right to left. While these examples are trivial taken by themselves, their existence implies that spatial orientation may be measurably different in different cultures.<sup>8</sup>

### References

- Cochran, W. G., & Cox, G. M. *Experimental designs*. 2nd edition. New York: Wiley, 1957.
- Ghent, L. Form and its orientation. *Amer. J. Psychol.*, 1961, 74, 177-190.
- Gilinsky, A. S. Masking of contour-detectors in the human visual system. *Psychon. Sci.*, 1967, 8, 395-396.
- Harris, C. S. Perceptual adaptation to inverted, reversed, and displaced vision. *Psychol. Rev.*, 1965, 72, 419-444.
- Held, R., & Freedman, S. J. Plasticity in human sensorimotor control. *Science*, 1963, 142, 455-462.
- Howard, I. P., & Templeton, W. B. *Human spatial orientation*. New York: Wiley, 1966.
- Hubel, D. H., & Wiesel, T. N. Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *J. Neurophysiol.*, 1965, 28, 229-289.
- Kendall, M. G. *Rank correlation methods*. London: Griffin, 1948.
- Kolers, P. A. Reading temporally and spatially transformed text. In K. Goodman (Ed.), *The psycholinguistic nature of the reading process*. Detroit: Wayne University Press, to appear (a).
- Kolers, P. A. Direction of scan and the nature of the task. In M. Kinsbourne (Ed.), *Hemispheric asymmetries of function*. London: Tavistock, to appear (b).
- Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., & Pitts, W. H. What the frog's eye tells the frog's brain. *Proceed. Inst. Radio Engngs., New York*, 1959, 47, 1940-1951.
- Miller, G. A. *Psychology, the science of mental life*. New York: Harper & Row, 1962.
- Rock, I. *The nature of perceptual adaptation*. New York: Basic Books, 1966.
- Sekuler, R. W. Spatial and temporal determinants of visual backward masking. *J. exp. Psychol.*, 1965, 70, 401-406.
- Smith, K. U., & Smith, W. K. *Perception and motion*. Philadelphia: Saunders, 1962.
- Sutherland, N. S. The methods and findings of experiments on the visual discrimination of shape by animals. *Exp. Psychol. Soc. Monogr. No. 1*, undated.

### Notes

- Two pieces of carbon paper inserted back-to-back between two pieces of typing paper will yield all of the transformations; but words have to be typed backwards for some of them.
- The statistics were derived from formulas in Cochran and Cox (1957). I thank Professors R. A. Carlson and W. G. Cochran for their help in formulating the design.
- Thanks to Mesdames Anne Boyer and Kathryn Rosenthal who collected the data.
- The time spent correcting errors is not the basis of the rank order of the transformations. This was established by finding the

time taken to read lines on which Ss made no errors and normalizing these times to a page of 25 lines for each S individually. The rank order of the normalized data was identical to the original rank order.

5. I thank Rabbi Shalom Rosenblit of Kiriat Hanoar, Boys Town Jerusalem, through whose great kindness the testing was made possible. Its conduct was arranged through the Israel Institute of Applied Social Research, and I thank the Director and staff, especially I. M. Schlesinger, for their cooperation, and Tamar Tolodano for assistance in preparing and scoring the texts.

6. The optimal data for comparison would be obtained from Ss who are skilled readers of left-going languages and lack familiarity with others. There are few such Ss available. Most skilled readers

of Semitic languages, for example, also have a considerable knowledge of some European language. Skilled readers of only left-going languages exist in some religious seminaries in Israel and (perhaps) in other countries, but efforts to obtain Ss in the appropriate age range from this source were not successful.

7. The Hebrew alphabet is illustrated in most desk-size dictionaries of English.

8. This work was supported principally by the National Institutes of Health (Grant IPOIGM-14940-01), and in part by the Joint Services Electronics Program (Contract DA 28-043-AMC-02536(E)).

*(Accepted for publication September 12, 1967.)*