Perceptual economy and the impression of visual depth

D. VICKERS

University of Adelaide, Adelaide, South Australia 5001

Three experiments were carried out to investigate the hypothesis that the impression of visual depth given by gradient patterns is due to the operation of a principle of "perceptual economy." The first experiment showed that the greater the number of values of the gradient variables in a pattern, the stronger was the tridimensional impression given by that pattern. The second showed that the greater the number of gradient variables, the less the amount of the pattern that had to be exposed before a tridimensional response was elicited. The third showed that judged slants increased as a function of the number of gradient variables in each pattern. The results suggest that the impression of visual depth given by a pattern represents a compromise between the economy gained from making a tridimensional specification of the elements in the pattern and the possible cost of making an error in coding the information in this way.

Gibson (1950a, 1966) has argued that an adequate basis for visual depth perception is provided by invariants in the visual array which is sampled by an O, either as he gazes fixedly at a particular scene or as he scans or moves around in his surroundings. Although the higher order variables examined by Gibson may in most cases provide effective stimuli for the perception of depth, however, there are several instances in which it seems difficult to apply the Gibsonian theory of ecological optics. In some cases, such as the "balloon demonstration" described by Ittelson and Kilpatrick (1951), the perceptual system appears to respond to a higher order variable, which, though present, is misleading. The result is that two stationary balloons are perceived as moving backwards and forwards, though it is only their brightness and size which covary. In other cases the perceptual system appears to respond inappropriately to a higher order variable, as when a rotating trapezoidal window is seen as an oscillating rectangle (Ittelson & Kilpatrick, 1951), or when a luminous, two-dimensional Necker cube, viewed in darkness, is seen as three-dimensional, even to the extent that the faces appear of different sizes (Gregory, 1966, p. 156). In yet other cases the "interpretation" of higher order variables by the perceptual system remains the same, though the stimulus situation is quite different. For example, Clark, Smith, and Rabe (1955) found that an unslanted trapezoid, viewed monocularly, was seen as a slanted rectangle, while a rectangle at a slant was seen veridically, although Os were warned of the possibility of being presented with a variety of combinations of shape and slant. As Gibson (1966, p. 287) points out, since his theory is primarily a theory of correct

perception, it must explain these and other examples of "incorrect" perception by supplementary assumptions, e.g., that the O makes a probable guess (p. 304). However, the examples of incorrect perception on which Gibson concentrates are those elicited by equivocal or contradictory information rather than those representing an actual distortion of the information in the visual array. Moreover, he does not offer a detailed account of what happens when the perceptual system is confronted with two or more hypotheses, each of comparable plausibility, i.e., under what conditions the perceptual system might be expected to reject one hypothesis, alternate between the two, or settle for a compromise.

An alternative approach which might serve both as a theory of perception and of "misperception" is similar to that first explored by Hochberg and McAlister (1953) and Attneave (1954), and more recently by Garner (1962, 1966, 1970). This suggests that stimulus information is perceived as invariant because this is a more economical way of encoding the sensory data. For example, it is possible to specify Pattern A in Fig. 1 as four parallel lines of varying thickness and interline interval. It is also possible, however, to specify the array tridimensionally so that thickness and interline interval become invariant. This form of coding minimizes the variation hypothetically admissible in the pattern, thereby reducing the redundancy in its description.

It may be objected that the concepts used by Hochberg and McAlister, Attneave, and Garner are very different, and indeed are not always used consistently by their respective authors (Vickers, 1967; Evans, 1967). One solution to the many contradictions has been

proposed by Vickers (1967) and independently (and much more systematically) by Evans (1967). Instead of the usual two, it is possible to distinguish between three information measures: (1) the average information associated with any combination of the values of each of a given number of variables when no constraints are imposed on their selection; (2) the average information associated with any combination when some constraints are used in its selection; and (3) the average information associated with any of the combinations actually presented to or sampled by an O. The difference between the first and second measures determines the amount of "constraint" redundancy, i.e., the amount of constraint, patterning, or similarity in the patterns actually shown or sampled. The difference between the second and third measures determines the amount of "discrimination" redundancy, or the extent to which the observed patterns differ among themselves (Evans, 1967). On this view, an economy hypothesis would imply that in the usual "free" perceptual situation (i.e., where the emphasis is on organizing the stimulus information) the perceptual system seeks to arrive at a form of coding which will minimize discrimination redundancy (or maximize constraint redundancy). That is, the O tries to devise a general description which will include the sampled patterns as a subset but which will also restrict the number of hypothetically admissible combinations of the values of the variables observed in the sampled patterns. One way of doing this, as in the case of Pattern A in Fig. 1, is to specify the information in such a way that a variable with a number of values is coded as an invariant.

Since a tridimensional specification of Pattern A allows all the correlated values of each variable to be predicted from the angle of orientation of the array, then, the greater the number of correlated values in such an array, the greater will be the relative economy of such a specification. That is, if angle of orientation is regarded as the only effective variable for a tridimensional specification, then adding more appropriate lines will not increase the number of angles required for a complete description of the array, whereas if thickness and interline interval are variables, as in a bidimensional specification, then the number of combinations of the values of these variables would increase with each additional line. Hence we should expect that a tridimensional response to Pattern A should become more probable if more lines were added to

Fig. 1. The four patterns used in Experiment 1.

it, provided the thickness and interline interval could remain invariant in a tridimensional specification.

According to Gibson's theory, on the other hand, it would seem that in the same situation no change in perception should occur from Pattern A to Pattern B, since "the same stimulus array coming to the eye will always afford the same perceptual experience insofar as it carries the same variables of structural information [Gibson, 1966, p. 248]."

EXPERIMENT 1

Stimuli

The four patterns used are shown in Fig. 1. Each pattern corresponds to the projection of a horizontal grid of equidistant straight lines onto a plane inclined at 75 deg to the grid, the number of lines increasing from 4 in A up to 19 in D.

Procedure

Two patterns were displayed side by

side before O and at right angles to his line of sight. Os were instructed to indicate which looked "more three-dimensional." The six pairs of patterns were presented to each of six Os in a 6 by 6 Latin square design in which the two possible positions of each pattern were randomized from one O to the next.

Results

Judgments were recorded for each pair of patterns, and a coefficient of rank order agreement for paired comparison data was calculated (Kendall, 1948). For six Os and four patterns $U = 0.89$, so that agreement between Os is highly significant, with $p < .001$. The results support the hypothesis in showing that the greater the number of values of the two variables which can be reduced by choosing a tridimensional rather than a bidimensional response, the greater is the tendency to choose the former.

EXPERIMENT 2

The four patterns used are shown in Fig. 2. When specified bidimensionally, Pattern I has one variable, interline interval, with several different values; II has an additional variable, namely, line thickness; while III has three and IV has four variables, each with several values. If each pattern is specified tridimensionally, however, the number of different values for each variable can be reduced to one. The economy hypothesis implies that the greater the number of variables that can be made invariant by adopting a tridimensional specification, the greater should be the probability of making such a response. Hence, we should expect a greater tendency for a tridimensional response to IV than to III, to III than to II, and to II than to I. In this case, prediction from Gibson's theory is more difficult. As Gibson (1950b) points out, the assertion that a gradient in texture is a stimulus for perceived slant would assume, in its simplest form, that the perception was a response to a mathematical property of the texture. In his later statement of the theory of ecological optics, this seems to be his own preferred interpretation (Gibson, 1966). Since the gradient variables in Patterns I-IV all specify a surface at the same slant, it would seem to follow that the depth impression from each should be the same. To paraphrase Gibson's own (1950b) analogy: the melody is the same, even though it is played on several different instruments in unison.

Stimuli

Pattern I corresponds to the projection of a horizontal grid of equidistant parallel lines onto a plane at 75 deg to the grid. The lines, however, are relatively thin and do not decrease in thickness from bottom to top in their projection. In Pattern II the lines are of appreciable thickness and decrease in thickness at the same rate as the interline intervals. Pattern III is similar to I, except that it is composed of broken lines. The lengths of the lines and the distances between them decrease at the same rate as the interline intervals. Finally, Pattern IV is similar to Pattern III, except that the broken lines are of the same thickness as those in Pattern II and decrease similarly in thickness from bottom to top. The patterns were constructed by photographically reducing 43 x 36 cm graphical projections, and mounting 15 x 10 cm matt prints of each reduction in black cardboard frames.

Procedure

Each O was given one pattern at a time just before him on a table. After

Fig. 2. The four patterns used in Experiment 2.

Os had given some spontaneous indication that they obtained a tridimensional effect from the complete pattern, the pattern was covered by an opaque slide. Os were then asked to uncover the patterns slowly, line by line, until they received an impression of depth. Immediately afterwards, the pattern was completely exposed, and Os were asked to bring the slide down from the top until the impression of depth disappeared. Each of the 24 student Os took about 7 min to complete all eight adjustments, and all 24 permutations of the four cards were used.

Results

The mean number of lines (or rows of broken lines) at which Os reported a change were calculated and are shown in Table 1, together with the results of one-tailed sign tests of the significance of the differences between the means for I and II, II and III, and III and IV. The results confirm the hypothesis in that the greater the

number of variables which can be made invariant by choosing a tridimensional rather than a bidimensional response, the greater is the tendency to choose the former.

EXPERIMENT 3

In the course of Experiment 2 several Os reported that the more vividly tridimensional patterns appeared to slope away at a more extreme angle than the others, so that the top of IV, for example, seemed "further away" than that of III. This would seem to throw doubt on a common assumption that distance

estimation should become more consistent as increased "cues," "information," or "redundancy" are added to a visual display. The alternative view suggested by this observation is that Os become progressively "bolder" in their estimations of slant as redundancy is increased, i.e., that their estimates become greater rather than simply less variable. The third experiment was conducted to test this observation, and to gain some measure of the possible range of quantitative differences in response.

Stimuli

The four patterns used were the same as those of Experiment 2, shown in Fig. 2.

Procedure

The apparatus is shown in Fig. 3. Each O was shown one of Patterns I, II, III, or IV, viewed monocularly through a window in a screen at a distance of about 2 m, and a fifth, finely-textured pattern (X), viewed binocularly through another window in the same screen. At the beginning of each trial, Pattern X was set to a roughly perpendicular or horizontal position. Os were asked to increase or decrease the slant of X to correspond as closely as possible with the apparent slant of the gradient pattern. Os performed one type of adjustment for all four patterns before performing the opposite adjustment. The order of the "increase" and "decrease" adjustments was balanced, and all 24 permutations of the order of the four patterns were used once. Each of the 12 student Os took about 15 min to complete all eight adjustments.

Results

The mean angle of all settings for each pattern was calculated for each O, the overall mean settings in degrees from the vertical being 48.67, 50.67, 64.50, and 66.67 for I, II, III, and IV, respectively. A test of average rank correlation (Jonckheere, 1954) with the number of gradient variables in each pattern gave a value of $\tau = .639$, which is highly significant ($p < .001$, one-tailed). Since the previous experiment has shown that the amount of a pattern which has to be exposed before a tridimensional

Table 1
Number of Gradient Variables in Each Pattern in Experiment 2, and the Number of Lines That Have to Be Exposed Before a Tridimensional Response is Obtained

Pattern	Number of Gradient Variables	Mean Number of Lines Exposed for 3-D Response	Significance of Differences	
			Patterns	Level
I	1	10.7	I-II	$p < .001$
II	2	8.4	II-III	$p < .001$
III	3	5.1	III-IV	$p < .001$
IV	4	3.0

Fig. 3. The apparatus used in Experiment 3.

response is elicited is an inverse function of the number of gradient variables in the pattern, we may interpret this trend as showing that the more readily the gradient pattern is seen as tridimensional, the greater the angle from O at which it appears to slope away.

DISCUSSION

It must be admitted that Experiment 1 has a certain unnerving simplicity and that a sceptic might argue that Os were judging the greater number of lines rather than tridimensionality. Although it would require a separate experiment to refute this point, an examination of the patterns in Fig. 1 would suggest that it is subjectively very implausible. In addition, the present interpretation agrees well with informal observation of a child's early attempts to master perspective drawing; when he is dissatisfied with the depth impression given by the railway tracks he has drawn as receding into the distance, a favorite strategy employed by a child is to double the number of sleepers or the number of fenceposts beside the track (an endeavor often frustrated by his inability to carry this out while preserving the same gradient). Analogous effects of textural grain on distance estimation have recently been reported by Corsini and Pick (1969).

Again, it may be argued, in connection with Experiment 2, that Patterns I and III contain contradictory gradients, since the thicknesses of the lines do not decrease. If the elements are seen as overlapping, however, it is possible to perceive the patterns in such a way that this contradiction is minimized. For example, Pattern I may be seen as the side of a clinker-built boat, or Pattern III as a tiled roof. In both

cases the shaded parts would constitute the lower edges of the elements, and these do not automatically decrease in projective size as a function of distance from the O. In any case, this objection would not account for the significant differences in response between I and III and between II and IV.

Experiments 1 and 2 were intended to test the hypothesis that perception is a process of encoding sensory data in as economical a way as possible. The experiments illustrate two techniques of studying the framing of perceptual hypotheses by O: the first consists of holding the number of variables constant, while increasing the number of values on each variable; the second consists of holding the number of values constant and increasing the number of variables. Similar experimental situations for studying the framing of perceptual hypotheses readily suggest themselves. For example, certain configurations of converging circles, drawn on a disk in a Michotte apparatus (Michotte, 1963) and seen through the viewing slot, appear at slow speeds to expand and contract in a bidimensional way but, at high speeds, to rotate about an axis, i.e., to display movement in three dimensions. Experiment 1 suggests that the greater the number of elements in a pattern which can be made invariant in a tridimensional response, the more readily will such a response be elicited. In a similar way, we may expect that the greater the number of appropriate elements in such moving patterns, the more readily will a tridimensional response be elicited, e.g., the slower the speed at which the disk must spin before the pattern is seen to rotate. Vickers (1967) has obtained results in a situation of this kind which show that

the speed at which the disks have to rotate before a tridimensional response is elicited is an inverse function of the number of converging circles drawn on the disk, and which lend support to the view that redundancy in the form of spatial repetition of the elements in a pattern may in some way be analogous to the repetition of the pattern in time. Similar results, with regard to the speed of rotation, have also been reported by Langdon (1951).

The results from Experiment 3 show that the order of magnitude of the corresponding slant setting follows the order of tendency to evoke a tridimensional effect, as measured in Experiment 2. In other words, the fewer the number of variables, or values of a given variable, that have to be exposed before a given pattern looks tridimensional, the greater the slant will appear to be when the whole pattern is exposed, even though the slant setting necessary to achieve the maximum invariance remains the same.

It may be that the phenomenon is peculiar to the particular set of patterns used, although some preliminary investigation with actual plane surfaces suggested this was not so; indeed, this result resembles the well-known continuum of judgments of phenomenal regression in relation to the amount of information about distance (Holway & Boring, 1941; Thouless, 1931). Similar results have frequently been reported in the literature (e.g., Clark, Smith, & Rabe, 1956; Attneave, 1969). Usually, these are interpreted as showing that perception of depth or slant becomes "more accurate" as increased "cues" are added to any array (Dember, 1965, p. 180; Forgas, 1966, p. 213). The ambiguity in the use of the term "accurate," however, disguises the point that the most direct way of applying the Gibsonian hypothesis to this situation would be to suppose that as more gradient data become available to O, the less variable his estimates of slant based on the data should become. In a similar way, a simple economy hypothesis would postulate that Os adopt a perceptual response that will make all gradient variables invariant. However, there is strictly only one angle at which the figures could be specified so as to make the values of the gradient variables equal, and all other slant responses would seem on this view to be more redundant.

The present results, on the other hand, suggest that O attempts to have his cake and eat it at the same time, i.e., he perceives the variables in the patterns as invariants, but adjusts the slope at which the hypothetical plane

containing them must lie, according to the amount of confirming evidence, even though this means distorting this evidence. Such an explanation could also account for the perception of a rotating trapezoidal window as an oscillating rectangle and would predict that simpler structures would be less likely to evoke the impression of movement in depth of a rigid body (e.g., Wallach & O'Connell, 1953). It seems worthwhile, therefore, to consider the possibility that the perceptual system does not simply tend to reduce redundancy in the coding of visual data to a minimum, but seeks to balance the advantage gained by this against the possible cost of making an error, or introducing some distortion into the data, by more efficient recoding. A moderate reduction in redundancy might warrant a moderate distance or slant estimate, while a high reduction might justify a bolder or increased estimate.

These effects must be studied in more detail before proposing theories of greater precision. It would seem useful, however, to try to develop the economy hypothesis, modified along the lines suggested above. It appears to be a hypothesis of possibly greater generality than the Gibsonian approach, while at the same time depending upon the same rich variety of stimulus information to which the theory of ecological optics draws attention. Besides suggesting an account of incorrect as well as correct perception, it could deal with the perception of two-dimensional stimulus patterns or "noninformative" stimuli (Gibson, 1966, p. 245) in the same way as that of arrays specifying real objects in depth. Finally, its application to perceptual learning would follow very similar lines to those proposed by Gibson (1966) and E. J. Gibson (1969). As Bower (1966)

suggests, infants may begin by perceiving invariants in stimuli and sequences of stimuli rather than learning how to perceive by progressing from simple percepts to complex ones, in the way we learn elementary geometry at school. On the present view, the perception of invariants would be an economy necessitated by the enormous information potential of the visual array and the limited cognitive capacity of the infant. Perceptual learning would accordingly take the form of developing perceptual hypotheses that combined economy with the discrimination redundancy necessary for particular situations. Variations in the achievement of such a compromise would seem to correspond naturally with differences in perceptual or cognitive style such as those distinguished by Barron (1952).

REFERENCES

- ATTNEAVE, F. Some informational aspects of visual perception. *Psychological Review*, 1954, 61, 183-193.
- ATTNEAVE, F. The determination of perceived tridimensional orientation by minimum criteria. *Perception & Psychophysics*, 1969, 6, 391-396.
- BARRON, F. Personality style and perceptual choice. *Journal of Personality*, 1952, 20, 385-401.
- BOWER, T. G. R. The visual world of infants. *Scientific American*, 1966, 215, 80-92.
- CLARK, W. C., SMITH, A. H., & RABE, A. Retinal gradient of outline as a stimulus for slant. *Canadian Journal of Psychology*, 1955, 9, 247-253.
- CLARK, W. C., SMITH, A. H., & RABE, A. Retinal gradients of outline distortion and binocular disparity as stimuli for slant. *Canadian Journal of Psychology*, 1956, 10, 77-81.
- CORSINI, D. A., & PICK, H. L. The effect of texture on tactually perceived length. *Perception & Psychophysics*, 1969, 5, 353-356.
- DEMBER, W. N. *Psychology of perception*. New York: Holt, Rinehart, & Winston, 1965.
- EVANS, S. H. Redundancy as a variable in pattern perception. *Psychological Bulletin*, 1967, 75, 18-26.
- FORGUS, R. H. *Perception*. New York: McGraw-Hill, 1966.
- GARNER, W. R. *Uncertainty and structure as psychological concepts*. New York: Wiley, 1962.
- GARNER, W. R. To perceive is to know. *American Psychologist*, 1966, 21, 11-19.
- GARNER, W. R. Good patterns have few alternatives. *American Scientist*, 1970, 58, 34-42.
- GIBSON, J. J. *The perception of the visual world*. Boston: Houghton-Mifflin, 1950a.
- GIBSON, J. J. The perception of visual surfaces. *American Journal of Psychology*, 1950b, 63, 367-384.
- GIBSON, J. J. *The senses considered as perceptual systems*. London: George Allen & Unwin, 1966.
- GIBSON, E. J. *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts, 1969.
- GREGORY, R. L. *Eye and brain*. London: World University Library, 1966.
- HOCHBERG, J. E., & McALISTER, E. A quantitative approach to figural "goodness." *Journal of Experimental Psychology*, 1953, 46, 361-364.
- HOLWAY, A. H., & BORING, E. G. Determinants of apparent visual size with distance variant. *American Journal of Psychology*, 1941, 54, 21-37.
- ITTELSON, W. H., & KILPATRICK, F. P. Experiments in perception. *Scientific American*, 1951, 185, 50-55.
- JONCKHEERE, A. R. A test of significance for the relation between m rankings and k ranked categories. *British Journal of Statistical Psychology*, 1954, 7, 93-100.
- KENDALL, M. G. *Rank correlation methods*. London: Griffin, 1948.
- LANGDON, J. The perception of a changing shape. *Quarterly Journal of Experimental Psychology*, 1951, 3, 157-165.
- MICHOTTE, A. *The perception of causality*. London: Methuen, 1963.
- THOULESS, R. H. Phenomenal aggression to the real object. II. *British Journal of Psychology*, 1931, 22, 1-30.
- VICKERS, D. Theories and experiments on visual discrimination and the perception of visual depth. Unpublished dissertation submitted for the degree of PhD in the University of Cambridge, 1967.
- WALLACH, H., & O'CONNELL, D. N. The kinetic depth effect. *Journal of Experimental Psychology*, 1953, 45, 205-217.

(Accepted for publication December 12, 1970.)