

Factors affecting the magnitude of the Ponzo perspective illusion among the Baganda

PHILIP L. KILBRIDE

Bryn Mawr College, Bryn Mawr, Pennsylvania 19010

and

HERSCHEL W. LEIBOWITZ

The Pennsylvania State University, University Park, Pennsylvania 16802

The magnitude of the Ponzo perspective illusion, utilizing stimuli with varying amounts of context, was determined for a population of Ugandan villagers. The subjects were classified as two-dimensional, three-dimensional, or mixed perceivers based upon their verbal responses to photographs portraying symbolic depth cues. The illusion magnitude among those subjects classified as three-dimensional was similar to that of college observers who demonstrate an increase in illusion magnitude with increasing background context. The two-dimensional observers' responses were similar to those of villagers tested previously, showing no significant increase for the same conditions. The role of two-dimensional flatness cues was eliminated on the basis of a control experiment. The data are interpreted as reflecting the operation of a cognitive factor determining responsiveness to symbolic depth cues in two-dimensional reproductions.

Leibowitz and Pick (1972) have recently reported that the Ponzo illusion is essentially nonexistent among a group of rural Baganda in Uganda. The illusion was evaluated by means of a series of stimuli (Figure 1) with increasing amounts of context, including a plain background, a background consisting of two converging lines (classical Ponzo figure), a photograph of a plowed field, and a photograph of railroad tracks. Whereas college students in the United States (Leibowitz, Brislin, Perlmutter, & Hennessy, 1969) as well as in Uganda (Leibowitz & Pick, 1972) demonstrated a progressively larger illusion effect as the richness of the background depth cues increased, the rural Ugandans demonstrated little or no illusion under any of the background conditions.

There are several interpretations to account for the lack of or diminution of illusion magnitude among

The field research and analysis for this study was supported by Grants MH22538 and MH08061 from the National Institute of Mental Health. The senior author would like to express his gratitude to the Makerere Institute of Social Research, Makerere University College, Kampala, Uganda, for their assistance during his tenure there as a senior research associate from June 1972 to May 1973. Professors Clive Davis and Michael C. Robbins offered invaluable assistance during various phases of the field research and Mr. Joseph Busuulwa served as our Muganda colleague throughout various phases of the field work. We would also like to gratefully acknowledge the advice of Professor Howard Hoffman regarding statistical procedures and Mr. Christopher L. Stitt, who assisted in statistical analysis. Professors Matthew Yarczower and Robert J. Miller provided critical evaluations for which we are thankful.

adult observers. In line with the ecological hypothesis (Segall, Campbell, & Herskovitz, 1966), the influence of these depth cues is assumed to depend upon the previous history of exposure to the cues in question. (For a comprehensive literature review, see Miller, 1973.) Utilizing the same stimuli, Leibowitz et al. (1969) have shown that the illusion magnitude was less for Guamanian college students, for whom the depth cues in question are less familiar, than for Pennsylvania or Ugandan college students (Leibowitz & Pick, 1972). However, this interpretation must be rejected in the case of the rural Ugandans, since their environment is rich in the specific cues, particularly perspective, represented in these stimuli. Alternatively, Leibowitz and Pick (1972) have suggested that the educated subjects may have learned to disregard the "flatness cues" (Hardy & Perrin, 1932; Schlosberg, 1941), inherent in two-dimensional line drawings and photographs. Such cues, (e.g., absence of binocular disparity, surface reflections, presence of the border of the photograph, etc.), might have dominated the responses for the uneducated rural villagers whose exposure to two-dimensional reproductions is limited. On the other hand, for the educated groups, with a history of exposure to two-dimensional reproductions of various kinds, the flatness cues could be ignored, so that the depth cues would become effective. From this point of view, the strength of the learned depth cues would be equivalent for these groups while observing in a natural environment, but in the experimental situation, while viewing two-dimensional representa-

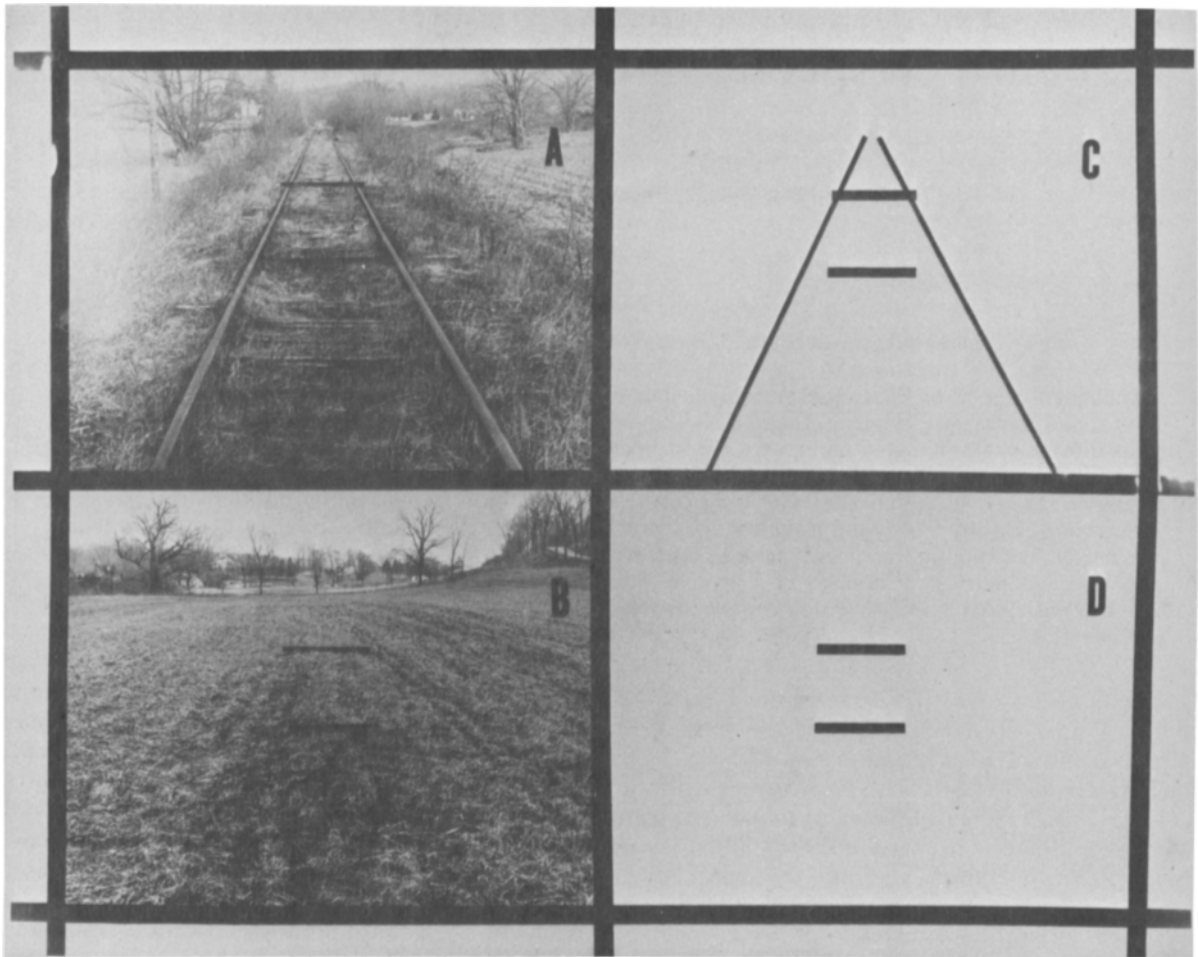


Figure 1. The stimuli used in the main experiment. The extent of the horizontal lines in these examples is the same (after Leibowitz & Pick, 1972).

tions, the same depth cues would be effective only for those observers who had learned, as a concomitant of the educational process, to ignore flatness cues.

A third interpretation, which might be called cognitive or perceptual "style," is suggested by the finding that many Baganda are not cognizant of cues used to symbolically indicate depth in photographic portrayals (Kilbride & Robbins, 1969). For example, traditional Baganda tend overall to perceive pictorial content two-dimensionally in contrast with the more educated Baganda, who, through their exposure to formal education and modern mass media, are cognizant of pictorial cues such as perspective, object size, and superimposition. Since response to the Ponzo pictures is activated at least in part by recognition of symbolic cues to distance, it seems reasonable to expect that the traditional rural Baganda would be relatively unresponsive to such cues. Conversely, Uganda university students,

cognizant of pictorial depth cues, should therefore demonstrate a higher responsiveness to the Ponzo stimuli.

The purpose of this study is to compare alternative interpretations for the previously observed absence of illusion magnitude in two-dimensional reproductions among the rural Baganda. This study utilizes the same stimuli and procedures as in the previous experiments, and was carried out in the same geographical area of Uganda.

EXPERIMENT I

Method

The original sample consisted of 105 Baganda. For reasons to be mentioned later, only 82 Baganda, 53 males and 29 females, are included in the present analysis. These subjects ranged in age from 16 to 80 years, with a mean of 31 years. Formal education varied from none to 12 years, with a mean of 5.5 years. The subjects were selected from the same rural Ugandan population as in the previous



Figure 2. Photograph of a rural road in Uganda. This photograph, along with Figure 1A, were utilized in the preliminary experiment to classify the subjects as "two"- or "three"-dimensional perceivers.

study. However, an effort was made to include subjects who varied in the amount of their formal education as well as two-dimensional reproductions in their homes. Most of these subjects were full or part-time cultivators, with some working in clerical or commercial occupations. Therefore, the sample varied in both education and in terms of their exposure to mass media. Many homes, for example, contain photographs, printed materials, and in a few cases television sets. Their visual environment typically consists of small hills of no more than 500 ft in height, separated by valleys. Human settlement is usually confined to the higher elevations, with the homes dispersed and embedded among plantain gardens. Numerous roads and paths transverse the area, offering vistas with perspective, sometimes for several miles. The Baganda tend to reside on or near roads which are important in their daily commerce and communication.

Before administration of the Ponzo figures, each subject was shown the track stimulus, Figure 1D, and a photograph of a local road, Figure 2, and asked "What do you see?" While asking the question, the experimenter traced his finger over the railroad tracks or the borders of the road. The purpose of this procedure was to classify subjects as "two"- or "three"-dimensional perceivers, following the distinction made by Kilbride and Robbins (1969). The subjects were classified as three-dimensional if responses to both photographs corresponded to three-dimensional objects, i.e., road, track, river, or bridge. Two-dimensional perceivers were so classified if their responses corresponded to two-dimensional objects in both cases, i.e., house, ladder (most frequent), hill, roof, hill with ladder. Those subjects who gave a two-dimensional response to one stimulus and a three-dimensional one to the other were classified as "mixed." A description of the three groups by sex, age, and education is given in Table 1. Table 2 shows the range of object identifications considered to be "two-dimensional."

Administration and scoring of the Ponzo test was identical to that followed in the previous studies.¹ Twenty-three subjects were eliminated because they failed to produce correct judgments for both extreme conditions in one or more series of stimuli.

Results

The principal datum, the magnitude of the Ponzo illusion expressed as the percentage overestimation of the upper line for the four backgrounds, is presented for the various background conditions in Figure 3,

with the data for the subjects separated, depending upon whether they were classified in the "two," "three," or "mixed" dimensional categories. For comparison, the data from the previous study for the villagers and for the college students are also reproduced. It will be noted that the data for the three-dimensional perceivers are similar to those of the college students in the previous Uganda (as well as Pennsylvania) study. The general trend of the data is the same for the mixed and for the two-dimensional perceivers, but the absolute values are systematically lower.

A two-way analysis of variance with repeated measures was performed for the illusion scores for the "2-D," "mixed," and "3-D" groups (Winer, 1962). This analysis indicated that there was a significant effect of groups ($F = 11.05$, $df = 2,79$, $p < .01$), a significant series effect ($F = 43.64$, $df = 3,237$, $p < .01$), and a significant interaction ($F = 4.64$, $df = 6,237$, $p < .05$).

An analysis of the simple main effects for the series factor was performed. This analysis showed that for both the "3-D" group and the "mixed" group there was a significant series effect ($F = 47.60$, $df = 3,237$, $p < .01$, and $F = 47.60$, $df = 3,237$, $p < .05$, respectively), whereas the "2-D" group showed no significant effect of series ($F = 2.31$, $df = 3,237$, $p < .05$). These findings are interpreted to support the expectation that Baganda who perceive depth (e.g., perspective) in pictures are responsive to the Ponzo illusion, while those who perceive no depth are not responsive.²

An analysis of the simple main effects for the group factor was also performed. This analysis showed that for both the control and geometric viewing conditions there was no significant group effect ($F = .41$, $df =$

Table 1
Sex, Age, and Educational Data for Each Subgroup

	Male	Female	Mean Age (Years)	Mean Educational Level (Years)
2-D	11	8	36.2	4.4
Mixed	8	3	39.2	4.1
3-D	34	18	27.4	6.2

Table 2
Range of Two-Dimensional Object Identifications

Road Stimulus	N	Railtrack Stimulus	N
House	14	Ladder	15
Hill	3	House	4
Roof	1	Roof	2
		Hill	2
		Hill w/Ladder	1
		Poles	1

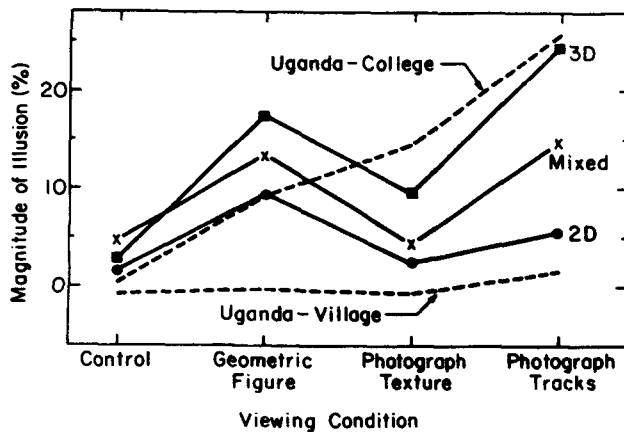


Figure 3. The magnitude of the Ponzo illusion for the principal stimuli reproduced in Figure 1. The subjects have been separated into groups based on their responses to the preliminary test. For comparison, the data from the previous study are indicated by the dotted lines.

2,316, $p > .05$, and $F = .90$, $df = 2,316$, $p > .05$, respectively). For both the phototexture and the phototrack series, there was a significant group effect ($F = 3.94$, $df = 2,316$, $p < .05$, and $F = 21.92$, $df = 2,316$, $p < .01$, respectively). These data indicate that there is a significant difference among "2-D," "mixed," and "3-D" groups in the photographic viewing contexts only.

In effect, the classification of the subjects, based on the results of the preliminary tests, reveals that the results for the rural Ugandans classified as three-dimensional perceivers is strikingly similar to that for the college-educated groups previously tested. The two-dimensional perceivers show no illusion magnitude for any of the background conditions as reported previously for a group of rural Ugandans.

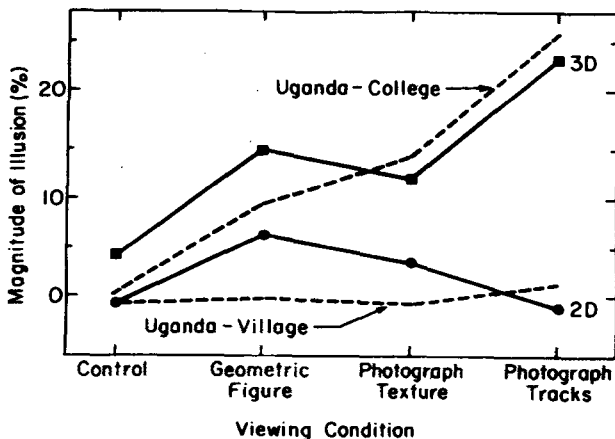


Figure 4. The magnitude of the Ponzo illusion under presentation conditions designed to eliminate or minimize cues to flatness. Classification as 2D or 3D is based on the preliminary test. For comparison, the data from the previous study are indicated by the dotted lines.

The mixed group produced intermediate illusion values.

EXPERIMENT II

Experiment II was designed as a direct test of the hypothesis that flatness cues may have been responsible for the lack of illusion magnitude in the previous study as well as the similar results for the two-dimensional perceivers in Experiment I of the present investigation. To this end, the same stimuli were presented in a viewing box specially designed to minimize flatness cues. The stimuli were mounted vertically at a distance of 63 cm and viewed against an opaque background. This viewing distance was chosen because it represented the "ortho" viewing position, i.e., the distance at which the absolute angular dimensions of the stimuli are identical to those in the original scene from the camera position. This distance has been shown to maximize the "reality" or "plastic" depth in photographs (Hardy & Perrin, 1932; Schlosberg, 1941). This arrangement also minimizes cues to flatness such as the familiar background against which the stimuli are ordinarily viewed, borders, and reflections from the surface.

Twenty-six additional rural Baganda subjects, selected and tested in the same manner as those in Experiment I, were utilized. The data, separated into two-dimensional and three-dimensional perceivers (there were no mixed responses among this group) are plotted in Figure 4. The three-dimensional group is similar to the previous three-dimensional perceivers as well as the college students, while the two-dimensional group, as was true for the previous subjects, demonstrated no increase in illusion magnitude with increasing background complexity.

A two-way analysis of variance with repeated measures on one factor for the illusion magnitudes for the "two-dimensional" and "three-dimensional" groups indicated that there was a significant effect of groups ($F = 28.80$, $df = 1,24$, $p < .01$), a significant series effect ($F = 4.10$, $df = 3,72$, $p < .01$), and a significant interaction ($F = 5.10$, $df = 3,72$, $p < .01$).

An analysis of the simple main effects for the series factor showed that for the "three-dimensional" group there was a significant series effect ($F = 6.90$, $df = 3,72$, $p < .01$), whereas the "two-dimensional" group showed no significant effect of series ($F = 2.47$, $df = 3,72$, $p > .05$). These findings are interpreted to support the cognitive style interpretation, since Baganda who perceive no depth (e.g., perspective) in pictures are not responsive to Ponzo stimuli in a viewing situation where flatness cues have been markedly reduced, whereas "three-dimensional" perceivers are responsive.³

An analysis of the simple main effects for the group factor was also performed. This analysis showed that for both the control and geometric viewing conditions

there was no significant group effect ($F = 1.69$, $df = 1.24$, $p > .05$, and $F = 4.01$, $df = 1.24$, $p > .05$, respectively). For both the phototexture and the phototrack series, there was a significant group effect ($F = 4.27$, $df = 1.24$, $p < .05$, and $F = 31.66$, $df = 1.24$, $p < .01$, respectively). These data indicate that there is a significant difference between "2-D" and "3-D" groups in photographic viewing contexts only.

DISCUSSION

The present study provides empirical support for the hypothesis that a cognitive factor is responsible for the previously observed differences in illusion magnitude between populations of college students as compared with rural Ugandan villagers. Contrariwise, the hypothesis that the group differences are a result of the ability to disregard flatness cues must be rejected. The critical factor is whether the subjects are sensitive to symbolic depth cues as evidenced by their responses to the pretest photographs. While it is easy to describe the operational definition of two- vs. three-dimensionality based on such a test, it is more difficult to hypothesize or specify the underlying mechanisms. However, it is apparent that the tendency to respond three-dimensionally to photographs is related to response magnitude to the Ponzo illusion, particularly in the photographic reproductions.

A logical question in relation to the previous study (Leibowitz & Pjck, 1972) is that none of the subjects in that experiment showed any illusion response whatsoever to any of the stimuli, thus implying that they were all two-dimensional. In the present study, a deliberate attempt was made to identify rural Ugandan villagers who might be expected to fall into the three-dimensional category. In effect, the present population represents a biased sample in relation to the previous group. Although we have no way of knowing what percentage of the rural Ugandan population falls into the 2-D, 3-D, or mixed perceiver categories, the fact that the previous study, in which the sampling was random, produced consistently low illusion effects suggests that most of this population falls into the 2-D category. Clearly, more detailed specification of the categories and their possible relationship to response to photographs and two-dimensional stimuli is needed. At the moment, we can conclude with certainty only that we have identified a factor which allows us to predict the responsiveness of subjects to the illusion stimuli.

Most importantly, the present data have methodological implications for cross-cultural research. Previous research has demonstrated that many populations, in addition to the Baganda, might also be considered as two-dimensional perceivers (see

Miller, 1973, for a review). Therefore, since pictorial depth perception is not a cultural universal, it is necessary to take this fact into account in research involving two-dimensional stimuli. Unless attention is paid to this precaution, there is a danger of falsely interpreting a lack of responsiveness as due to a reduced strength of the space cues in question. As an alternative, it might be feasible to consider the possibility of utilizing three-dimensional stimuli, thus completely avoiding the question of whether the subjects are responsive to cues presented in the two-dimensional mode.

REFERENCES

- HARDY, A. C., & PERRIN, F. H. *The principles of optics*. New York: McGraw-Hill, 1932. Pp. 465-469.
- KILBRIDE, P. L., & ROBBINS, M. C. Pictorial depth perception and acculturation among the Baganda. *American Anthropologist*, 1969, 71, 293-301.
- LEIBOWITZ, H., BRISLIN, R., PERLMUTTER, L., & HENNESSY, R. Ponzo perspective illusion as a manifestation of space perception. *Science*, 1969, 166, 1174-1176.
- LEIBOWITZ, H., & Pjck, H. Cross-cultural and educational aspects of the Ponzo illusion. *Perception & Psychophysics*, 1972, 12, 403-432.
- MILLER, R. J. Cross-cultural research in the perception of pictorial materials. *Psychological Bulletin*, 1973, 80, 135-150.
- SCHLOSBERG, H. Stereoscopic depth from single pictures. *American Journal of Psychology*, 1941, 54, 601-605.
- SEGALL, M. H., CAMPBELL, D. T., & HERSKOVITS, M. J. *The influence of culture on visual perception*. Indianapolis: Bobbs-Merrill, 1966.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

NOTES

1. Test administration was in Luganda. Instructions were translated into Luganda through the technique of back translation. Each subject was asked (for 10 cards on each of the four series): "On each of these pictures there are two horizontal lines." "I want you to tell me which of the two lines is longer than the other. If it is the top, say 'top.' If it is the bottom, say 'bottom'" ("Ku buli bifaananyi bino, kuliko ennyiriri biri ez'obukiika. Njagala ombulire kunyiriri ezo ebiri luluwa olusinga lunalwo obuwanvu. Bweruba nga lwawaggulu 'gamba nti lwawagullu' bweruba nga lwawansi 'gamba nti lwawansi'"). In all stimulus cards, the upper line was constant in length while the length of the lower lines were presented in random order. For all observation conditions, the equality value was determined by interpolation as the midpoint of the region at which the subject's responses changed. The magnitude of the overestimation, or illusion, is the percentage overestimation of the upper line.

2. One-way analyses of variance and, where appropriate, Newman-Keuls tests were also performed for the illusion scores for the "3-D," "mixed," and "2-D" groups (Winer, 1962). The results showed a significant effect of treatment for the "3-D" and "mixed" groups ($F = 33.93$, $df = 3,153$, $p < .05$, and $F = 3.15$, $df = 3,30$, $p < .05$, respectively). All possible comparisons were significant

($p < .05$) for the "3-D" group. For the "mixed" group, the control condition was significantly different ($p < .05$) from the photograph track condition, with all other individual comparisons insignificant. The "2-D" group showed no significant effect of treatment ($F = 2.75$, $df = 3,52$, $p < .05$). Overall, these further analyses indicate that for the "3-D" group the observed dip in magnitude of the illusion between the geometric figure condition and the photograph texture condition is a significant difference. Therefore, our present assumption that phototexture contains more depth (perspective) information than a line-drawn railway track may be questionable.

3. One-way analyses of variance and, where appropriate, Newman-Keuls tests were also performed for the illusion scores for

the "3-D" and "2-D" groups. The results showed a significant effect of treatment for the "3-D" group ($F = 5.55$, $df = 3,27$, $p < .05$), whereas the "2-D" group showed no significant effect of treatment. For the "3-D" group, the control condition was significantly different ($p < .05$) from the photographic track condition, with all other individual comparisons insignificant.

(Received for publication September 9, 1974;
revision received February 3, 1975.)