Perception, 1997, volume 26, pages 153–170

The perception of surface folding in static and animated displays

Manfredo Massironi
Istituto di Psicologia dell'Università di Verona, via San Francesco, 37129 Verona, Italy
Nicola Bruno
Dipartimento di Psicologia, Università di Trieste, via Università 7, 34123 Trieste, Italy
Received 3 April 1996, in revised form 15 November 1996

Abstract. How do we interpret outline drawings of surfaces? Although pictorial depictions are projectively ambiguous, observers demonstrate definite preferences of interpretation. Additionally, they commit typical errors. A study is reported of one specific arrangement of surfaces as it is represented in outline drawings, namely the arrangement that results when two arbitrary surfaces are joined at a common edge to form an angle in 3-D ('phenomenic folding'). With some of these arrangements, observers report that the angle formed by the two surfaces is zero (complete folding). With others, they report that the angles are greater than zero (incomplete folding). Both interpretations are actually valid. Several investigators have proposed that observer preferences such as these are due to a tendency to prefer a 3-D interpretation that will make the depicted 3-D shape regular.

Three experiments were performed to test this regularisation hypothesis. In the first, observers were shown pairs of four-sided polygons joined at one equal side. Their task was to imagine how the smaller polygon could be folded completely towards the larger, and, subsequently, to report on its position after the folding ('mental folding'). Reported positions were consistent with 3-D interpretations that caused figural regularisations. In the second and third experiments, observers were shown drawings of diamonds and parallelograms folded along a number of differently positioned and oriented segments ('folding edge'). Their task was to estimate verbally the extent of the dihedral angle formed by the two surfaces. Results indicated that the perception of incomplete folding is determined by 3-D interpretation of the orientation of the drawing with respect to the picture plane. In a fourth experiment, observers were asked whether projective equivalences might be disambiguated by animating two kinds of displays that yield the 'incomplete folding' effect but that should be distinguishable on the basis of the trajectories of the vertexes of the folding parts. Results demonstrated that even in these conditions observers are unable to interpret the foldings correctly. These results might be taken to indicate that projective, static information leading to a simpler and more regular interpretation of the display can prevail over explicit motion information that should force the system to achieve a nonregular solution.

1 Introduction

Some outline drawings are immediately interpreted as surfaces that have been folded over themselves. Two or more outlines are juxtaposed appropriately at some sides and the resulting interpretation on folding can be quite compelling (see figure 1—some compelling outlines).

Massironi (1988) investigated a number of such outlines under the rubric 'phenomenic folding'. He isolated four necessary factors and a number of facilitating conditions. Of the necessary factors, three pertained to the geometry of the outlines: (i) one side (the folding line) of the two outlines must coincide; (ii) the two outlines must be on the same side of the folding line; and (iii) arrow-junctions must be present at the ends of the folding line. These factors are suggestive of a role of mechanisms involved in the interpretation of occlusions. Accordingly, Massironi's fourth factor pertained to the perceptual set of observers: (iv) the outlines must be perceived as two overlapping surfaces. These results were interesting because they linked perceived folding with two phenomena that had been studied by Gestalt psychologists: amodal completion behind occlusions (Kanizsa and Gerbino 1981; Metelli 1960; Michotte et al 1962)

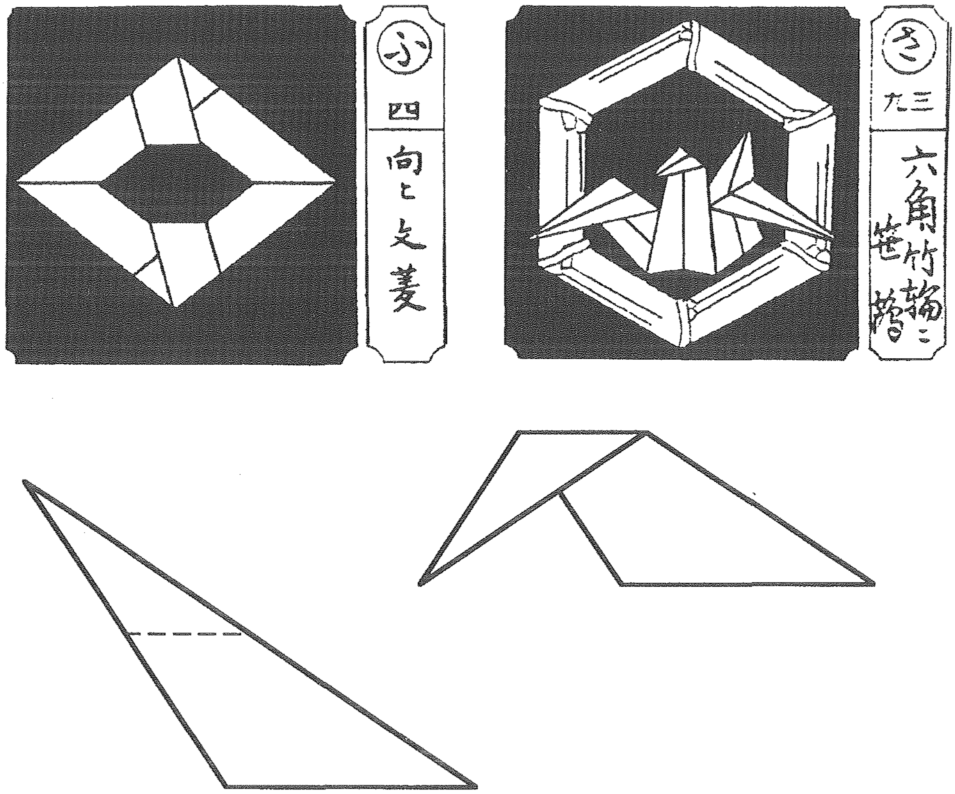

Figure 1. Outline drawings of folded surfaces. Top: traditional Japanese heraldic crests (*Japanese Design Motifs*, 1972, New York: Dover). Bottom: Triangle and its folding edge marked by a dotted line before and after a complete folding.

and perceived causality (Michotte 1946). With the appropriate geometry, observers perceive two overlapping parts, one in the foreground and one in the background. This perceptual interpretation is connected with a previous event whereby one part of the figure was bent towards the other which remained fixed. This rotation is experienced as the cause of the present state of affairs and is the basis for the double awareness that is typical of these outlines. One perceives the two overlapping parts but is, nonetheless, aware of their belonging to a single structure. In his study, Massironi (1988) carefully selected only outlines that were consistently perceived as being parallel to the picture plane. For reasons that will be clarified presently, when these figures are perceived as folded, it is more likely that the folding appears complete, that is that the folded part appears to have undergone a 180° rotation towards the fixed part. This is not true in general. Several types of outline are readily perceived as folded, but the folding appears incomplete, that is the folded part appears to have been rotated less than 180°, in some cases, significantly less (see figure 1, incomplete vs complete). In order to shed light on this point we must first present the geometrical rule for drawing the complete folding of a two-dimensional pattern. Consider the pentagon ABCDE in figure 2a, and the folding line (FL) around which the quadrilateral MCBN (folded part) must be rotated by 180°. The result of this rotation is obtained by drawing from points B and C two segments perpendicular to FL and extending them in such a way that $C(FL) = (FL)C'$ and $B(FL) = (FL)B'$, then joining B' with N and C' with M, (see continuous line pattern in figure 2b).

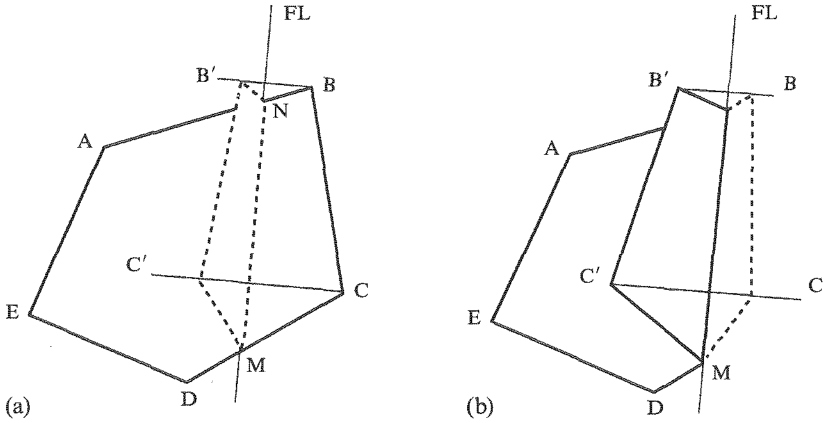

Figure 2. Geometric construction of a complete folding; explanation in the text.

What is puzzling in the context of our research is that in many cases the results of geometric complete folding are seen as incomplete. How and why does this happen? To answer this question the problem of the projective ambiguity of phenomenic foldings must be considered. Let us observe the dashed quadrilateral in figure 2a, which can be the representation of the folded part MCBN captured at the moment of its rotation around FL. In this case this can be an incomplete folding of MCBN or, alternatively, it can represent the complete folding of a different folded part such as the dashed quadrilateral in figure 2b. Like the folded part, MC'B'N in figure 2b could be either the result of a complete folding of the quadrilateral MBCN (figure 2a) or the incomplete folding of a larger different quadrilateral. In conclusion, from a projective point of view each phenomenic folding is, if we do not know the shape of the unfolded figure, ambiguous. Instead, from a perceptual point of view, each phenomenic folding is unambiguous in the sense that we see only complete or incomplete foldings according to factors that are independent from the geometric constraints previously described.

The search strategy of the following experiments was substantially explorative and was therefore focused on discovering these factors in order to try to understand how they work. Each experiment sheds some light on the structural nature of phenomenic folding, but none answers all, or even most, of the questions that might be asked about performance of the particular task studied.

Inspecting phenomenic folding outlines led us to suspect that these cases may be related to a process of regularisation. Some irregular outlines may become more regular if interpreted as being slanted in depth rather than parallel to the picture plane. The proposal that regularisation may be involved in the perceptual interpretation of pictorial displays has been advanced by several investigators (Attneave and Frost 1969; Deregowski and Parker 1992; Hochberg and Brooks 1960; Leeuwenberg 1971; Leeuwenberg and van der Helm 1991; Perkins 1972; Perkins and Cooper 1980), but it has rarely been linked to the perception of a previous causal event as in perceived foldings. The aim of experiment 1 was to investigate how potential regularisations may constrain the interpretation of foldings involving these types of outline figures. Observers were shown two outlines joined at a folding line and were requested to fold them mentally.

2 Experiment 1

A 'mental folding' task was used to study the relationship between events implicit in phenomenic folding and potential regularisations of involved outlines. In each phenomenic folding there were three fundamental components, namely, the fixed part, the folded part, and the folding line inclination. This experiment was aimed at verifying the role of these three components according to a regularisation hypothesis that foresaw that the mental folding task would be affected by the shape of both fixed and folded parts. When these parts were regularised, since they were not seen as parallel to the picture plane but slanted in depth, the imagined folding angle was higher than for shapes which did not require regularisation.

2.1 *Method*

2.1.1 *Observers.* Eight male and nine female undergraduates from the University of Verona participated in the experiment.

2.1.2 *Displays.* We chose three similar kinds of outline for both parts of each display. Criteria for selection were twofold: (i) the outlines were selected to ensure that they would satisfy Massironi's four criteria for phenomenic folding when actually folded; (ii) they were chosen so that they would be potentially informative regarding a process of regularisation. To this end we selected a parallelogram (PA, pa), a right trapezoid (RT), and an irregular quadrangle (QU, qu), respectively, for the larger and the smaller of the two outlines. In the smaller outline, however, the right trapezoid was substituted by a rectangle (re) to ensure that the folded outcome would satisfy Massironi's criteria in all combinations.

Of these three outlines, the parallelogram was chosen as an example of outline that could become more regular if interpreted as slanted in 3-D (under parallel projection any parallelogram can always be the projection of a square or a rectangle); the right trapezoid and the rectangle were chosen as examples of outlines that could not become more regular if interpreted as slanted in 3-D, whereas they presented some regularities if interpreted as parallel to the picture plane; finally, the third kind was chosen as an example of an outline that had no regular interpretation either in 3-D or when orthogonal to the line of sight. The factorial combination of the 3 kinds of outline for the larger part with the 3 kinds for the smaller part yielded a total of 9 experimental displays. Each of them was presented in 2 versions, at two different inclinations of the folding line (40° and 70°), yielding 18 distinct displays which are illustrated in figure 3. The displays were programmed and presented on a Macintosh LC computer equipped with a monochrome 60 Hz monitor. Luminance was approximately 50 $\mathrm{cd\ m^{-2}}$. The spatial resolution of the monitor was 640 × 480 (horizontal × vertical) pixels. Displays, black outlines on a white field, were presented at a distance of approximately 50 cm.

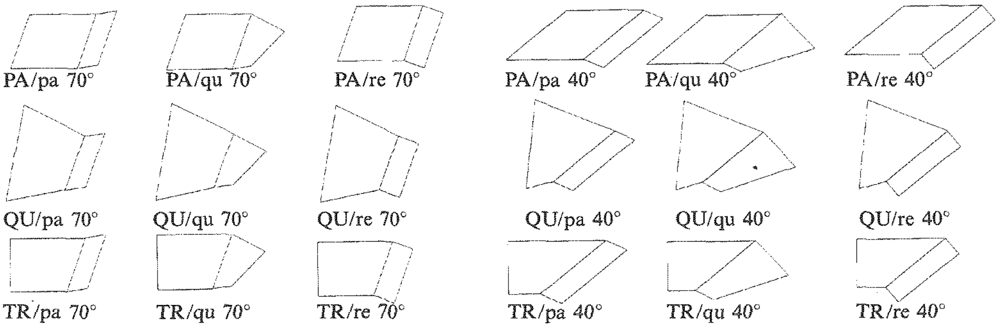

Figure 3. Stimuli of experiment 1. Fixed parts: PA = parallelogram, QU = quadrilateral, TR = trapezoid. Folded parts: pa = parallelogram, qu = quadrilateral, re = rectangle.

2.2 *Procedure*

Before the beginning of the experiment, observers were instructed concerning the meaning of a complete folding. They were presented with cardboard surfaces and encouraged to fold one part of it as if they were closing an envelope. In this way, they could become aware of the projective changes involved. Next, observers were told that they would be presented with a series of outline drawings. For each of these they were requested to perform a 'mental folding', that is to report how the depicted surface would look once folded completely, similarly to what they had just done with the cardboard. After the presentation of the display, the observer was requested to perform the mental folding. Once ready, he or she pressed a key on the keyboard. At this point, a circle made of 60 points appeared over the outline drawing. The circle was centred on the lower end of the folding edge. The observer was requested to click with the mouse on one of the points to mark the position of the smaller part once folded completely towards the larger one. The exact instruction was to mark the position of the lower outer edge, that is the edge that connected to the lower end of the folding edge. The experiment began with a few training trials to familiarise observers with the task. Next, the 18 experimental displays were presented in random order. Each was seen once with no time limits in which to complete the task.

2.3 *Results*

Average positions of the smaller parts after mental folding are summarised in figure 4. Observed angles were measured taking the folding line (FL) as the starting point. Three variables were considered: the inclination of FL (70°, 40°); the larger outline—the fixed part (PA, QU, TR); the smaller outline—the folded part (pa, qu, re). The data were subjected to a $2 \times 3 \times 3$ mixed design ANOVA which yielded (i) a main effect of larger outline, $F_{2,32} = 9.43$, $p < 0.0006$; (ii) no significant effect of the smaller outline, $F_{2,32} = 0.54$, $p > 0.5$; (iii) the larger part shape × folding edge inclination interaction, $F_{2,32} = 3.4$, $p < 0.05$. This was due to the fact that angles reported for the PA figures were smaller when the inclination was 70° (average = 78°82′) than when it was 40° (average = 90°12′), whereas this was not the case for reports of RT and QU (averages = 71°17′, 71°64′, 73°29′, 76°43′ for RT and QU, respectively). In figure 4 are also reported the significant (S) and not significant (NS) differences between the averages of the subjects' responses and the geometric 'correct' ones that were computed by means of a one group t test with $p < 0.05$.

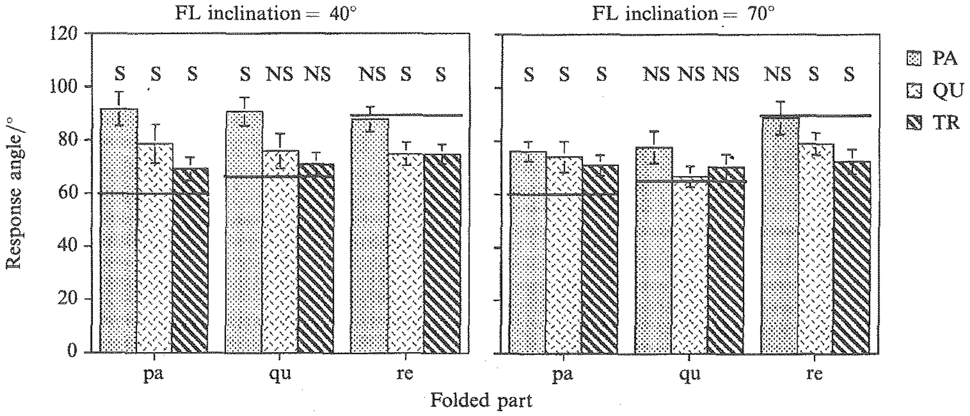

Figure 4. Means of the response angles according to fixed (PA, QU, TR) and folded (pa, qu, re) parts. Bold horizontal lines correspond to outlines interpreted as being parallel to the picture plane.

2.4 *Discussion*

Continuous horizontal lines of figure 4 represent expectations if outlines had been interpreted as being parallel to the picture plane, according to the geometric rule previously presented. Two results are noted. First, reported foldings differ systematically depending on the shape of the larger part. Parallelograms give a higher amount of the folded part rotation as compared to quadrilaterals and trapezoids. This result is closer to the expectations based on a 3-D slant interpretation; right trapezoids tend to give mean results closer to the expectations based on parallelism with the picture plane; irregular quadrangles tend to give results in between. This feature of the data is fundamentally consistent with regularisation. Parallelograms can become regular if interpreted as slanted in 3-D, and mental folding involving these outlines is more similar to the corresponding interpretation. Right trapezoids are more regular if interpreted as normal to the line of sight, and foldings are also consistent with this interpretation. Second, the normalisation effect on parallelograms is stronger at 40° than at 70°. This is also consistent with regularisation: as the inclination of the folding edge becomes more similar to 90°, the difference between the possible slanted interpretation and parallelism to the line of sight reduces (see the differences between the corresponding histograms of the two graphics of figure 4). A puzzling result is the lack of any significant effect associated with the shape of the folded part (pa, qu, re) as the histograms of figure 4 clearly show. The puzzle becomes even more intriguing when we realise that, according to the geometric rule of folding surfaces, the only difference in the results would have to depend on the shape of the folded part, no matter what the shape of the fixed one. A consequence of this effect is that the subjects' answers are some times statistically different from the geometrically 'correct' one and at other times they are not, for no explicit reason (see the S or NS marks on the histograms of figure 4). A tentative explanation could be that the fixed part, being larger and motionless, constitutes a reference schema to which the subjects refer as a stable source of information, neglecting the shape of the folded part which, being unstable, is not considered suitable.

3 Experiment 2

The results of experiment 1 showed that, in a phenomenic folding, when the fixed part was a shape perceived slanted in depth, such as a parallelogram, the folded part was mentally rotated more than when the shape was perceived to be parallel to the picture plane. This could be the reason for which some geometric complete foldings are seen as incomplete. The following experiment was aimed at assessing the strength of the incomplete folding effect as a function of possible 2-D and 3-D interpretations of the outlines. In order to verify this possibility the experimental method had to be changed, in particular: (i) the imagery task was replaced by a perceptual one, with the aim of detecting and evaluating the degrees of folding incompleteness; (ii) the fixed and the folded part were not kept independent of each other, but became two complementary parts of a single unfolded parallelogram; (iii) the stimuli, instead of being unfolded patterns which had to be mentally folded, were representations of geometrically complete foldings of the same parallelogram whose degree of phenomenic folding was to be evaluated. Two kinds of similar geometric figures that are readily perceived either as frontoparallel or slanted (parallelograms and diamonds) were selected. When the subjects were presented with two sides aligned with the horizontal axis, parallelograms tended to be perceived as rectangles slanted in depth. Conversely, when the subjects were presented with one of the diagonals aligned with the vertical axis, diamonds tended to resist a 3-D interpretation and were seen as parallel to the picture plane. Therefore, by comparing these two figures we hoped to manipulate 3-D interpretations and to observe consequent variations in the perception of folding. Having determined in experiment 1 that the inclination of FL was a second factor which influenced the subjects' performance, we tried to gain further insight into this factor by varying the position of the FLs.

3.1 *Method*

3.1.1 *Observers.* Twelve male and twelve female undergraduates from the University of Verona participated in the experiment.

3.1.2 *Stimuli and design.* The stimuli were black-on-white outlines of parallelograms and diamonds drawn on paper. The parallelograms were constructed by selecting 10 positions on the perimeter of a parallelogram having horizontal sides 9.5 cm long, oblique sides 7.2 cm long, and inside angles of 123° and 57°. As shown in figure 5a (see over), these positions were used as possible extrema of FLs and have, therefore, been connected in pairs, yielding 17 different configurations after eliminating impossible connections (such as pairs resting on the same side), reflections, and similarities. As in the previous experiment, we call stimuli derived from parallelograms the PA stimuli. Similarly, the diamonds were constructed by selecting 10 positions on the perimeter of a diamond having sides 7.2 cm long and inside angles of 123° and 57°. As shown in figure 5b, our final stimulus set consisted of 17 different configurations for diamonds too. We call these configurations the DI stimuli. Both PA and DI stimuli were complete folding, constructed according to the previously described rule, in which the folded part was always the portion of the unfolded pattern either on the right hand or on the top of the folding line. The experimental design was aimed at assessing the effect of two within-subject variables, type of figure (parallelogram or diamond) and position of FL. It must, however, be underlined that there were no grounds in this experiment for testing the interaction between these two variables, as the positions on the FLs were not comparable in the two figures.

3.2 *Procedure*

The 34 stimuli were presented in random order. For each of the stimuli, observers were asked to evaluate the width of the dihedral angle formed by the folded and the fixed parts of the folded surface. To standardise judgments, we used a comparison scale with seven angles drawn at increasing sizes from 0° to 90°, in steps of 15°, but observers were free to select intermediate positions if they so desired. There were no time limits. Before beginning the experiment, observers were shown instances of actual folded papers to illustrate the objects represented in the drawings and to clarify the task.

3.3 *Results and discussion*

Average evaluations ranged from 1.25° to 70.62° for the PA stimuli, and from 0° to 71.67° for the DI stimuli. On averaging across different FL positions, we observed a mean judgment of 30.64° for the PA stimuli and of 32.14° for the DI stimuli, a nonsignificant difference, $t_{23} = -0.071$, $p = 0.48$. To evaluate the effect of FL positions, which were not comparable in the two figures, we computed two one-way analyses of variance, which demonstrated significant effects in both figures: $F_{16,368} = 36.38$, $p < 0.001$ for the PA stimuli; and $F_{16,368} = 35.43$, $p < 0.001$ for the DI stimuli. Average evaluations as a function of FL position are summarised in table 1, discussed in the next paragraph. These results failed to demonstrate the systematic differences between the two figures, but they did suggest that evaluations varied systematically in different FL positions. But why should impressions of folding vary with FL positions?

To answer this question, we observed, after a careful inspection of the set of stimuli in relation to the data collected, that our stimuli could be divided into two groups. In the first group, both three-way intersections of edges were visible at the extrema of the FL; the folded part of the figure was therefore drawn completely inside

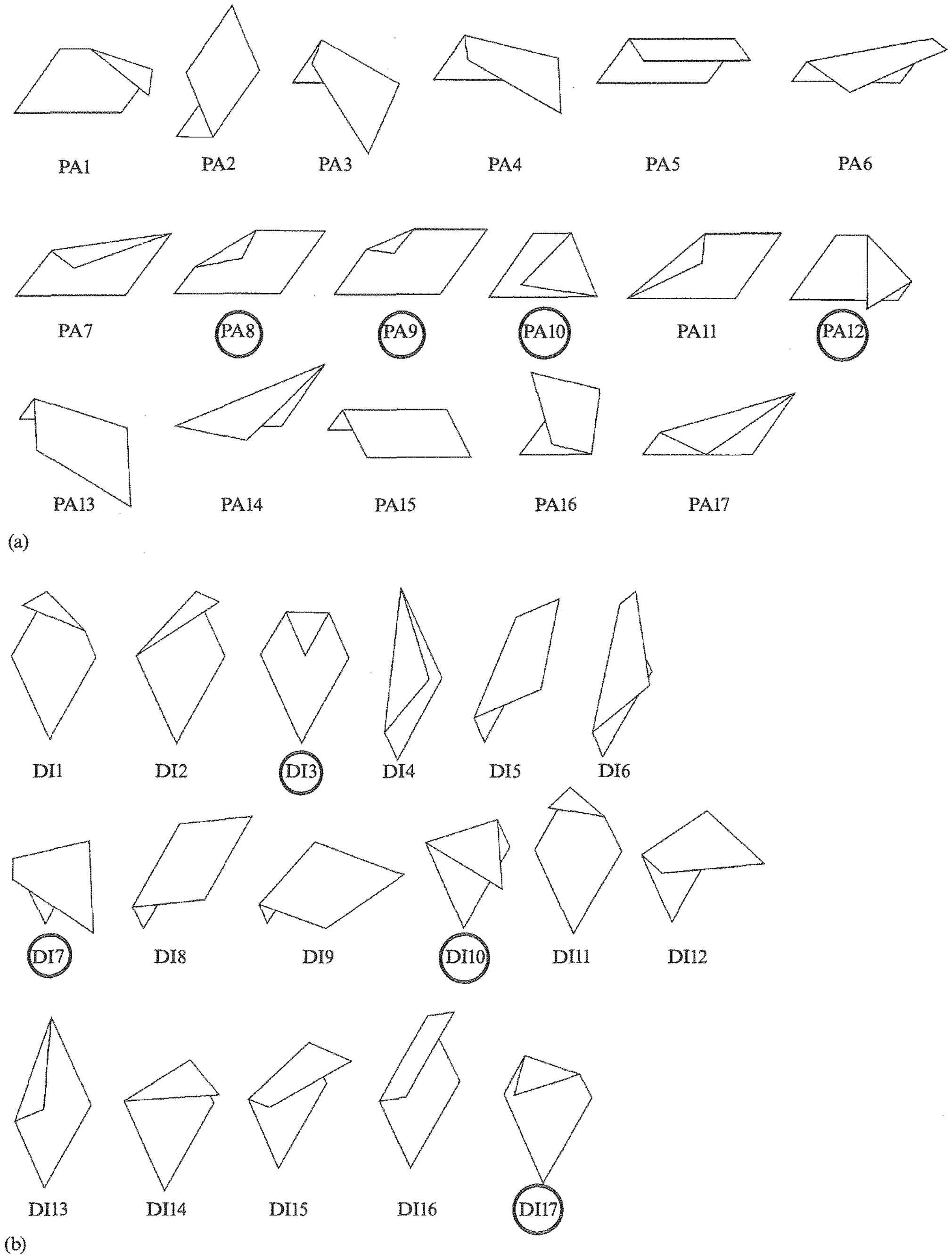

Figure 5. (a) The 17 stimuli from foldings of parallelograms. (b) The 17 stimuli from folding of DI. Circled numbers mark the cases where the folded part falls inside the fixed one (see text).

the outline of the fixed part. In these patterns, both the frontoparallel and the depth solutions were almost perfectly consistent with a complete rotation, as shown by the patterns with a circled number in figures 5a and 5b. In the second group, three-edge intersections were only partially visible at one of the FL extrema because the folded part occluded, either partially or completely, the edge of the fixed part converging to that extreme, causing the folded part not to be completely drawn inside the outline of the fixed part. In these patterns, the frontoparallel solution was projectively compatible with a complete folding, while the depth solution is not, as is shown by the other drawings in figures 5a and 5b. Thus, for the first group of stimuli we assumed that the patterns had been perceived as completely folded, whereas for the second group the patterns had been perceived as incompletely folded. On the basis of the qualitative impression that the phenomenal completeness of the folding depended on the position of the folded part relative to the fixed part (inside or outside) we considered the mean evaluations of these two kinds of stimuli separately. First, let us consider all stimuli where the folded part lay partially outside the fixed part, as summarised in table 1 (labels written in normal character). Covarying with FL position, we noted that the two segments belonging to the same side of the figure could project different angles on the picture plane. Following the hypothesis that these angles could provide a cue used by observers to evaluate the amplitude of the dihedral angle and, therefore, the phenomenal incompleteness of the folding, we looked at the correlation between the observed evaluations and these angles. This correlation was $r_{18} = 0.95$. We also regressed evaluations on angles and computed a 0.95 confidence interval on the slope of the regression line, $b = 0.862 \pm 0.142$ (see figure 6). Given that this slope is statistically indistinguishable from a unit slope, we are inclined to conclude that observers interpreted these stimuli as incomplete foldings on the basis of these angles. This conclusion supports the idea that observers tended to perceive the drawings as representations of incompletely folded rectangles in 3-D, not as frontoparallel parallelograms or diamonds folded completely. Following Perkins (1976), one could hypothesise that a minimum tendency is at work and favours the rectangle solution over less regular figures.

Next, we considered all stimuli presented in table 1 (labels in bold characters), that is all stimuli where the folded part lay inside the fixed part. We exclude from this group the PA7, PA11, PA17, DI4, and DI13 stimuli (marked with asterisks in table 1), as exceptional since they present (according to Lowe 1987 and Witkin and Tenenbaum 1983) not accidental properties, that is one of the FL extreme coincides with one of

Table 1. Average judged angles and related geometric angle values for the patterns employed in experiment 1, limited to the subgroup where the folded part falls outside the remainder of the figure (labels in normal character) and average judged angles of the subgroup where the folded part falls inside the remainder of the figure (labels in bold character). The asterisks indicate exceptional stimuli (see text).

Stimulus	Judged angle/°	Geometric angle/°	Stimulus	Judged mean/°	Geometric angle/°
PA1	18.75	24	DI1	36.45	34
PA2	68.75	69	DI2	51.25	46
PA3	13.12	6	**DI3**	**0**	
PA4	47.5	40	**DI4***	**8.33**	
PA5	68.54	67	DI5	68.13	92
(PA6)	(58.95)	(88)	(DI6)	(22.71)	(115)
PA7*	**33.75**		**DI7**	**1.25**	
PA8	**5.83**		DI8	77.08	73
PA9	**3.75**		DI9	52.08	45
PA10	**3.12**		**DI10**	**7.08**	
PA11*	**10**		DI11	16.33	15
PA12	**3.12**		DI12	15.96	16
PA13	58.75	40	**DI13***	**30**	
PA14	19.16	10	DI14	19.58	14
PA15	71.25	67	DI15	56.66	45
PA16	22.16	18	DI16	74.16	74
PA17*	**14.37**		**DI17**	**9.38**	

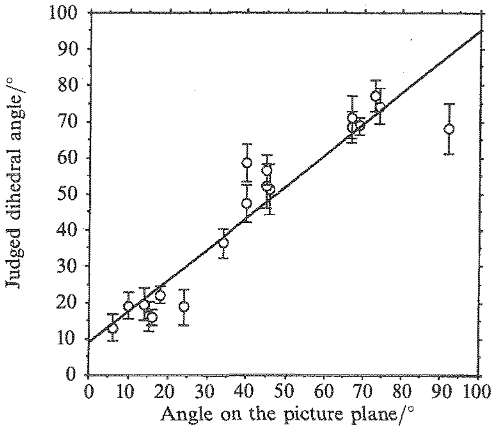

Figure 6. Linear regression of the average of the judged angles over the angles on the picture plane; explanation in the text.

the vertexes of the unfolded figure. All other stimuli yielded very low evaluations which were all in the range from 0° to 15°, corresponding to the first step of our comparison scale. On the basis of these results, we conclude that these stimuli were seen as complete or almost complete foldings, except when one of the FL extrema coincides with one of the vertexes of the unfolded figure. This result makes sense, because the folded part would lie inside the fixed part both when drawing a frontoparallel figure or when drawing a figure slanted in depth. In conclusion, and despite our failure to differentiate between parallelograms and diamonds, it seems that the impression of completeness of a represented folded surface is influenced by its 3-D interpretation in an even more fundamental way. This conclusion is supported by the comparison of the two groups of stimuli. In the subset where the folded part partially occluded the edge of the fixed part converging to one extreme of the FL, observer evaluations were consistent with the depth interpretation of the drawing and with the hypothesis that observers used the angle on the picture plane as an indicator of the extent of folding. In the subset where the folded part lay inside the fixed part, observer evaluations were consistent both with the frontoparallel and the depth interpretations. The results of this experiment seem to show that the phenomenic width of an outline folding is affected both by global and by local sources of information, namely the perceived inclination in depth of the patterns and the angles at the extrema of the FL.

4 Experiment 3

The results of experiments 1 and 2 indicated that the impression of folding might be influenced by the 3-D interpretation of the drawing. However, the second experiment also indicated that this interpretation cannot be manipulated simply by changing the type of figure. In order to find another way of manipulating this interpretation, we conducted a third experiment. The rationale for this study lies in our observation that in many of the stimuli used in the first experiment the impression of folding could be altered by rotating the drawing upside down. Figure 7 presents instances of this orientation effect. Again, the change makes sense in terms of geometrical optics. If a surface appears as frontoparallel, then its viewpoint must lie on an imaginary vertical line coming out of the centre of the surface and orthogonal to it. Conversely, if a surface appears slanted in depth, then its viewpoint must lie on a line which is not orthogonal to the surface. Therefore, when the drawing is the projection of a folded surface, it is possible that by turning the figure upside down it will force a different projective interpretation of the angles formed by the folded and the fixed parts.

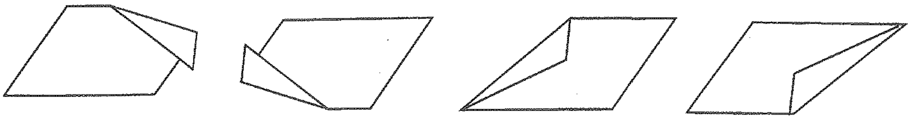

Figure 7. Effect of orientation on the amplitude of the apparent dihedral angle formed by the folding.

We also noted that the change as a function of the orientation was somewhat stronger with parallelograms than with diamonds. This observation is consistent with our previous hypothesis that diamonds are seen as frontoparallel more easily than parallelograms are.

4.1 *Method*

4.1.1 *Observers.* Nine male and nine female undergraduates from the University of Verona participated in the experiment.

4.1.2 *Stimuli and materials.* We used ten of the stimuli employed in the second experiment, five from the PA group and five from the DI group, again presented as drawings on paper. These stimuli were selected according to the following criteria: (i) extrema of the FL at adjacent sides of the figure and folded part partly inside the fixed part (PA1 and DI1 of experiment 2, figure 8a); (ii) extrema of the FL at adjacent sides of the figure but folded part completely inside the fixed part (PA8 and DI17 slightly modified from experiment 2, figure 8b); (iii) extrema of the FL connecting one of the vertexes of the figure and one side opposite to it (PA11 and DI13 of experiment 2, figure 8c); (iv) extrema of the FL at opposite sides of the figure and at equal distances from the vertexes (PA5 and DI16 of experiment 2, figure 8d); (v) extrema of the FL at opposite sides of the figure and at different distances from the vertexes (PA4 and DI15 of experiment 2, figure 8e). Each of these 10 stimuli was presented in two orientations: straight, and upside down, totalling 20 different experimental configurations.

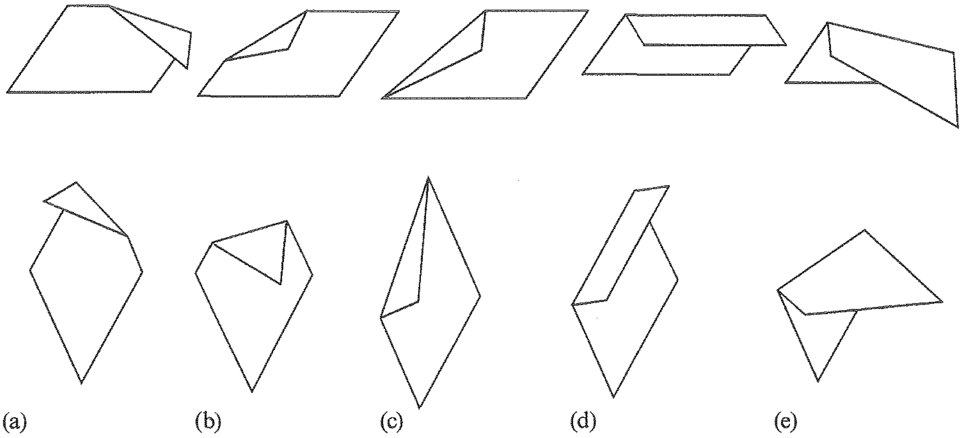

Figure 8. Stimuli in experiment 3.

4.2 *Procedure*

Each observer was shown all drawings in a randomised order. Training, instructions, and task were the same as in experiment 2, except for the rating scale that was of nine 10° steps, instead of seven 15° steps.

4.3 *Results and discussion*

Average estimates of angles formed by the folded and the unfolded parts of the figures were 41.61° for figures presented right side up, and 49.67° for those presented upside down. This difference was statistically significant, $F_{1,\,17} = 7.83, p < 0.02$, confirming that figure orientation contributed to the incomplete folding effect. However, the difference was large with parallelograms and much smaller with diamonds, as depicted in figure 9 and tested by evaluating the significance of the orientation × figure interaction, $F_{1,\,17} = 19.73$, $p < 0.0005$. A posteriori Duncan tests demonstrated that evaluations of parallelograms differed from those of diamonds both in right side up (xPA = 37.44; xDI = 45.77), $p < 0.01$, and in upside-down presentation (xPA = 53.11; xDI = 46.22), $p < 0.05$; whereas evaluations of right-side-up figures differed from those of upside-down figures with parallelograms (xPA ru = 37.44; xPA ud = 53.11), $p < 0.01$, but not with diamonds (xDI ru = 45.77; xDI ud = 46.22), $p < 0.05$.

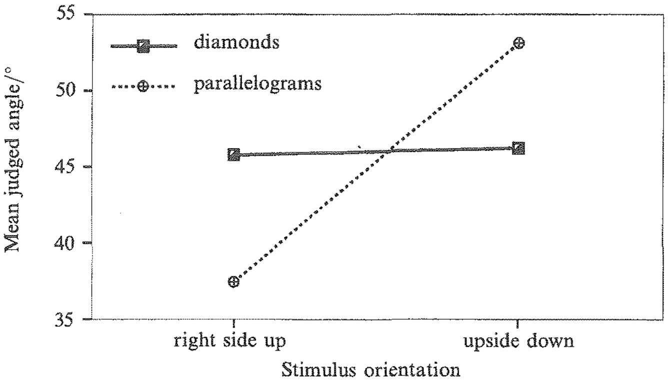

Figure 9. Effect of stimulus orientation on the amplitude of the phenomenic dihedral angle formed by the folding.

These results indicate that orientation affected the interpretation of the folding in parallelograms but not in diamonds. Thus, the tendency to interpret parallelograms as inclined in depth is also a factor in the incomplete folding effect when local factors, such as the presence of junctions that are indicative of occlusion or concavities in the silhouette of the figures, are controlled. Effects of 3-D are also possible with diamonds, but in this case they are limited to the folded part that may appear to stick out in depth from the fixed part. The fixed part is always seen as frontoparallel and, therefore, no orientation effects take place.

Our argument is best illustrated by means of projective geometry. Consider PA11, the drawing that yielded the largest difference between the two orientations (10° vs 38.89°, almost a 30° average difference). Figure 10a represents a hypothetical reconstruction of the position in space of PA11. Segment AB is a side view of the surface in one of the possible orientations relative to the viewpoint O. Segment AC is a side view of the folded part when the figure is in the right-side-up orientation, namely when this part forms a 10° angle with the unfolded part. Segment BC′ is a side view of the folded part when the figure is in the upside down orientation and the folded part is on the bottom. On the assumption that the perceived slant of the figure remains constant with the change in orientation, the visual angles subtended by the AC and BC parts remain the same ($\alpha = \beta$). Therefore, the ABC angle must be greater than the BAC′ angle, because gaze is not perpendicular to the centre of AB and the viewpoint is farther away from AC than it is from BC. Consider now DI13, the diamond display that corresponded to PA11 but yielded no difference between the orientations (an average evaluation of 41.67° in both cases). Applying the same logic as before, the spatial relations can be represented as in figure 10b, where AB is the segment and O is a viewpoint orthogonal to AB. Given the equalities between segments and visual angles, in this case BAC = ABC′.

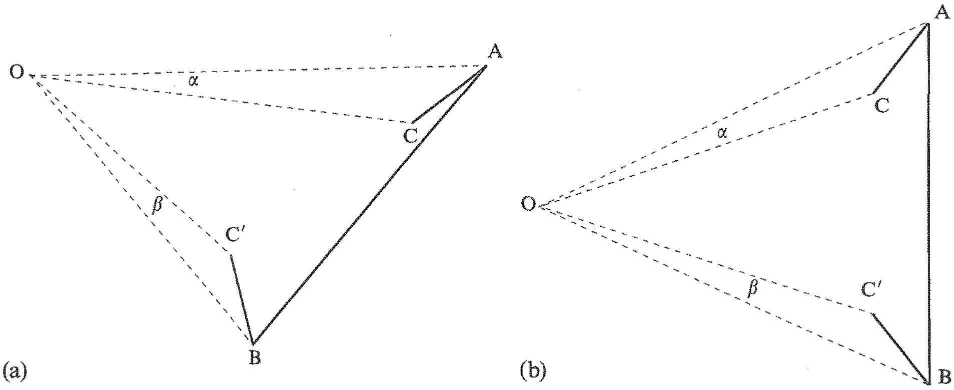

Figure 10. (a) Geometrical relations between an observer 'O' and a folded surface slanted in depth, for the PA11 stimulus depicted in figure 8c (top). (b) Geometrical relations between an observer 'O' and a folded surface slanted in depth, for the DI13 stimulus depicted in figure 8c (bottom); explanation in the text.

In sum, the results of this second experiment demonstrated that the perceived amplitude of the folding angle can be influenced by the orientation, and that this influence depends on the 3-D interpretation of the figure, which is seen preferably as frontoparallel with diamonds, and preferably as slanted in depth with parallelograms.

5 Experiment 4

The results of the third experiment confirmed that the incomplete folding effect is due in part to the 3-D interpretation of the drawing. A drawing of a folded parallelogram is more likely to be seen as a rectangle slanted in depth than as a diamond. Therefore, complete foldings of the parallelogram can be misperceived as incomplete foldings of a rectangle slanted in depth. But how do these two kinds of solution differ? Consider figure 11a (lower pattern), illustrating the trajectory of the vertexes of part of a slanted rectangle as this part is folded completely, and compare it with figure 11a (upper pattern), illustrating the trajectory of the vertexes of the same part belonging to a frontoparallel parallelogram. As can be seen from the figures, these trajectories are very different: in the case of the rectangle in depth, they are semicircular; in the case of the parallelogram, they are rectilinear. Thus, as the surface is folded over time, the trajectory of the vertexes provides unambiguous information for the type of folding being observed. This information is not available in the static view, particularly if only the endpoints of the rotation are used. If, as Perkins (1972) suggested, "the perceiver attempts to achieve an interpretation of ambiguous projections by imposing geometric regularities such as rectangularity" (page 72) then by disambiguating these projections in computer animation we should observe that the effect vanishes. Previous investigations of bending events (Jansson and Johansson 1973; Jansson and Runeson 1977) have shown that semirigid transformations of planar surfaces will induce varying percepts but will tend to preserve, if compatible with the stimulus constraints, the rigidity of the perceived figure. Thus, whenever it is compatible with the projective constraints on the stimulus, observers will prefer rigid rotations over semirigid transformations such as bending (or, in our case, folding), and semirigid transformations over elastic deformations. On the basis of these considerations, we should expect animations of frontoparallel folding parallelograms to yield to one of these two solutions: (i) The parallelogram is still seen as a rectangle in depth; in this case the folded part ought to be seen as nonrigid and performing an incomplete folding. (ii) The parallelogram is seen as an actual parallelogram parallel to picture plane; in this case the folded part ought to be seen as rigid and performing a complete folding.

5.1 *Method*

5.1.1 *Observers.* Sixteen undergraduates from the Universities of Verona and of Padua participated in this experiment.

5.1.2 *Displays.* We used computer-animated drawings of parallelograms that folded over time. The animations were built by starting from a parallelogram with 10 cm horizontal sides, 7.5 cm oblique sides, and 60° and 120° inside angles and by choosing four different positions of the folding line (see figure 11). Position 1 connected the upper with the lower side of the parallelogram, parallel to the inclined sides, and placed 25 mm from the right inclined side (figure 11a). Position 2 connected one point of the upper side (65 mm from the upper right vertex) with one point of the right side (25 mm from the upper right vertex, figure 11b). Position 3 connected one point of the upper horizontal side (25 mm from the upper left vertex) with one point of the lower side (25 mm from the lower right vertex, figure 11c). Position 4 connected one point of the left side (25 mm from the upper left vertex) with one point of the right side (25 mm from the lower left vertex, figure 11d). For each of these positions we computed two different animations: animation PA reproduced the projective transformations of parallelogram parallel to the picture plane folding completely; animation RE reproduced the corresponding projective transformations of a slanted rectangle. In the first case, the vertexes of the folding part moved along a straight line; in the second, they moved along curved paths (see figure 11). The animations were parallel rather than polar projections in order to maintain the form of the parallelogram throughout. Braunstein (1976), among others, has argued that polar projection views are neither necessary nor sufficient to perceive rigid forms in motion correctly. Animation frames were black outline figures on a white background and were generated and presented with a Macintosh LC personal computer. Animation sequences depicted surfaces that folded and unfolded continuously.

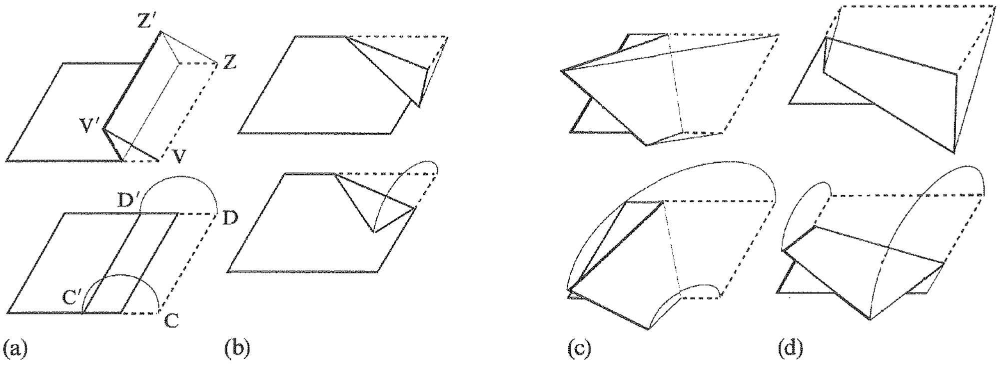

Figure 11. (a) Complete folding of a frontoparallel parallelogram compared with the isometric projection of a rectangle slanted in depth, also completely folded. In the first case, when the dashed part is rotated around the FL, the vertexes follow a rectilinear trajectory from V, Z to V′, Z′, when the corresponding dashed part is rotated, its vertexes follow a circular trajectory from C, D to C′, D′. (b), (c), (d) Trajectories of the rotating corners of frontoparallel parallelograms and of rectangles slanted in depth, according to the different positions of FL.

5.2 *Procedure*

The eight animations were shown successively in random order to each observer, who looked at them from a distance of 1 m. Observers could look at the animation as long as they wished. Their task consisted in two successive evaluations. In the first part, observers were shown each animation and were asked to report on four perceptual characteristics of what they saw: (i) the inclination in depth of the surface (0° to 90°, corresponding to the frontoparallel to the sagittal planes); (ii) the trajectory of the vertexes of the folding part, forcing a choice between straight and curved; (iii) the rigidity of the folding part, forcing a choice between rigidity or nonrigidity; (iv) the apparent amplitude of the rotation in depth of the folding part (180° for a complete folding, less if incomplete). In the second part, observers were shown the four pairs of animations that had identical positions of the folding line but which depicted the parallelogram or the rectangle in depth. The pairs were shown in succession, randomly ordered, and again they could be seen as many times as desired. In this second part, observers were requested to decide whether the difference between the two animations were due to: (i) the trajectory of the vertexes of the folding parts; (ii) the amplitude of the rotation; or (iii) the form or the inclination of the surface, that is whether the difference was due to a different form of the figure, a different inclination in depth, or both.

5.3 *Results*

Observer evaluations of the inclination in depth of the folding surface are summarised in figure 12a. On average, RE displays were judged as more slanted in depth (mean judgment 49.14°) than PA (35.7°), a significant difference, $F_{1,15} = 8.57$, $p = 0.001$. This result suggests that animating the drawings provided some information for a correct interpretation of the position in depth of the surface. However, this information is not sufficient to disambiguate completely the depth interpretation of parallelograms, and serves only to reduce the tendency to regularise them. As can be seen in the graph, this reduction took place systematically, at all positions of the folding line. Average evaluations also differed significantly as a function of the position of the folding line, $F_{3,45} = 4.62$, $p < 0.01$. The figure × position interaction was not significant, $F_{3,45} = 1.84$, $p = 0.15$.

Observer evaluations of the amplitude of the rotation of the folding part are summarised in figure 12b. These evaluations were significantly greater for rectangles than for parallelograms, $F_{1,15} = 116.48$, $p < 0.0001$, indicating that rectangles were always seen as folding completely, whereas parallelograms tended to be seen as folding only incompletely, to various extents depending on the position of the folding line, $F_{3,45} = 20.68$, $p < 0.0001$. We interpret this as evidence that this position does not influence the apparent amplitude of the rotation in the RE stimuli, whereas it does in the PA stimuli. These results are consistent with our previous conclusion that parallelograms are seen as rectangles slanted in depth, not as frontoparallel parallelograms even when motion information is available. The frequency of judgments of rigidity vs nonrigidity and of rectilinearity vs curvature of the trajectories are summarised in table 2. Binomial tests of these frequencies demonstrated that observers significantly favoured rectilinearity in two pairs of the PA displays and curvilinearity in two pairs of the RE displays, other comparisons being not significant (see third column in table 2).

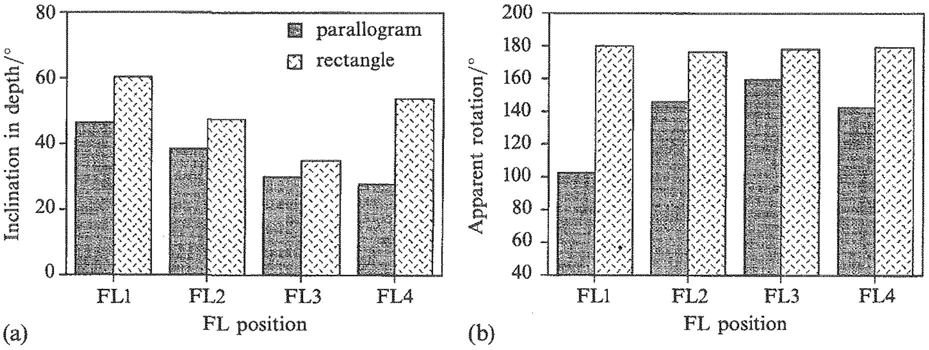

Figure 12. (a) Average observer evaluations of the inclination in depth of the folded part in experiment 4, according to the different positions of FL. (b) Average observer evaluations of apparent rotations of the folded parts according to the different positions of FL.

This result suggests that observers can judge the trajectory of the vertexes reliably only in some cases, and even in those cases this is not sufficient to induce the correct 3-D interpretation. Binomial tests on the rigidity judgments indicated that observers favoured rigidity in three cases out of four in the parallelograms, and in two cases out of four in the rectangles (see column 5 in table 2). These results indicate that observers differentiated between the two animations on the basis of the apparent trajectory of the vertexes and of the extent of the folding (complete or incomplete), not on the basis of the form or inclination in depth of the surface (see column 4 of table 3), confirming that both RE and PA displays were judged as slanted in depth, the latter display less so, however.

Table 2. Frequencies of trajectory and rigidity reports, referred to 16 subjects, after observation of the animations in experiment 4.

Stimulus	Trajectory			Rigidity		
	rectilinear	curvilinear	significance	rigid	nonrigid	significance
Parallelogram						
FL1	12	4	S	13	3	S
FL2	11	5	NS	13	3	S
FL3	10	6	NS	9	7	NS
FL4	13	3	S	13	3	S
Rectangle						
FL1	2	14	S	15	1	S
FL2	7	9	NS	13	3	S
FL3	2	14	S	7	9	NS
FL4	9	7	NS	11	5	NS

Table 3. Relative frequencies of observer selections of vertex trajectory, rotation amplitude, or inclination in depth as key elements to differentiate between pairs of stimuli cited in the first column.

Pairs of stimuli	Form of trajectory	Amplitude of rotation	Form or inclination in depth
PA/DI FL1	7	15	2
PA/DI FL2	9	8	1
PA/DI FL3	11	12	3
PA/DI FL4	9	11	4

5.4 *Discussion*

These results demonstrate that observers are not particularly sensitive to the overall consistency of the different geometric aspects of the animations. It seems that the final perceptual interpretation is achieved according to a hierarchical process that tends first to preserve the regularity of the figure by imposing the perception of a rectangle, and second to preserve its rigidity as much as possible. Once these two basic aspects are preserved, other variations in the animations are dealt with by attempting to adjust the various features of the interpretation in order to make them consistent with the two basic constraints, by varying the apparent slant of the figure, for instance, or neglecting the trajectory of the vertexes. This scheme is remarkably similar to the five constraints proposed by Ullman (1984) for the recovery of 3-D structure from motion: "(i) At each instant there should exist an estimation of the 3-D structure of the viewed object; this internal model of the viewed structure may initially be crude and inaccurate, and may be influenced by static sources of 3-D information. (ii) The recovery process should prefer rigid transformations. (iii) The recovery scheme should tolerate deviations from rigidity. (iv) The recovery scheme should be able to integrate information from extended viewing periods. (v) The recovery scheme should eventually recover the correct 3-D structure or a close approximation to it." (page 258)

6 Final considerations

Factors underlying the perception of phenomenic folding are easy to identify and to describe (Massironi 1988). It is, on the contrary, very difficult to discover and to set apart the factors that determine the phenomenic size of the dihedral angle formed by the fixed part and the folded part of an outline folding. The perceptual estimation of this angle seems to pose intriguing problems for the human perceiver. Despite the geometrical simplicity of the drawing, in some conditions observers make systematic errors when they are asked to evaluate the extent to which the surface is folded. This effect of incomplete folding seems to depend on many factors, both global and local, in reciprocal interaction with each other, a number of which emerged from our study:
(i) A tendency to regularise the represented figure, which is strong enough to resist integration with unambiguous motion information to the type of folding being shown. This regularisation may be conceived as a manifestation of a kind of minimum tendency (Perkins 1976). According to this view, when a minimum tendency imposes a regular interpretation for the form, size, or inclination in depth of some local structures in an image, this regular interpretation tends to propagate to adjacent structures and possibly to the whole image.
(ii) The shape of the fixed part. The subjects of experiment 1 in fact neglected the shape of the folded part, which is geometrically determinant, taking into account instead the shape of the fixed part. Therefore, the figural aspects considered by perception are not the same as those considered by geometry.
(iii) The position of the folding line. Depending on this position, the contour of the fixed part can be partially occluded by the folded part (as in the 'outside' figures in experiment 2). In the cases of such occlusions the subjects perceived the foldings as incomplete and seemed to draw the size of the folding angle from the angle visible at one extreme of the folding line.
(iv) The ineffectiveness of motion information, which made the animations of each pair of parallelograms very different from each other. Subjects of experiment 4, in fact, did not perceive either changes in the shape and position of the whole figure or in the rigidity of the folded part. With the exception of Massironi (1988) there are no reports in the literature on phenomenic folding, whereas many are available on other perceptual phenomena with which phenomenic folding has something in common. Hochberg and McAlister (1953), Hochberg and Brooks (1960), Attneave and Frost (1969), and Attneave (1972) showed, from different points of view, that 3-D perception from 2-D figures depends on a tendency to regularisation and to minimisation of the variability of angles, lengths, and orientations. Perkins (1972), and Perkins and Cooper (1980), from studies on the perception of the orthogonality between two faces of an outline parallelepiped, reached the conclusion that we cannot say the perceiver is a precise geometer, but nevertheless we cannot say that the perceiver is not a geometer at all; the perceiver can be better defined as a 'sloppy geometer' (page 117). Proffitt et al (1992) have looked at the mapping of the 2-D motion in an image to a 3-D perceptual interpretation. Based on their analysis of the stereokinetic effect and of the kinetic depth effect, they described the motion components that appear crucial for eliciting the depth interpretation, and observed that these are consistent with a perceptual heuristic for computing depth from motion, but not with a coherent computation based on geometry. Although their conclusion refers to a different class of phenomena from those investigated in this study, their final conclusion seems fully generalisable to our observations: "Clearly the visual system is doing something that is not entirely explicable in terms of a canonical geometrical analysis of the stimulus" (page 20). Wagemans and Kolinsky (1994), introducing a special issue of *Perception* on 'Perceptual organisation and object recognition', underline that the apparent conflicts between the wide range of mechanisms available to the visual system, make it very hard to foresee

how they can be embodied in one single system. Therefore, only a preliminary and open conclusion can be drawn: "Perhaps the visual system is able to use a variety of sources of information and mechanism to process them, by putting different weights on the various sources of information and compromising between different mechanisms, depending on the task" (page 380). The very puzzling question posed by phenomenic folding is: why perception of folding is, at the same time, so perceptually compelling and yet so imprecise?

Acknowledgements. This research received financial support from the Italian Ministry of University and of Scientific and Technological Research.

References

Attneave F, 1972 "Representation of physical space", in *Coding Processes in Human Memory* Eds A Melton, E Martin (Washington, DC: Winstone) pp 283 – 306

Attneave F, Frost R, 1969 "The determination of perceived tridimensional orientation by minimum criteria" *Perception & Psychophysics* **6** 391 – 396

Braunstein M L, 1976 *Depth Perception through Motion* (New York: Academic Press)

Deregowski J B, Parker D M, 1992 "Three-space inference from two-space stimulation" *Perception & Psychophysics* **51** 397 – 403

Hochberg J E, McAlister E, 1953 "A quantitative approach to figural 'goodness'" *Journal of Experimental Psychology* **46** 361 – 364

Hochberg J, Brooks V, 1960 "The psychophysics of form: reversible-perspective drawings of spatial objects" *American Journal of Psychology* **73** 337 – 354

Jansson G, Johansson G 1973 "Visual perception of bending motion" *Perception* **2** 321 – 326

Jansson G, Runeson S, 1977 "Perceived bending motions from a quadrangle changing form" *Perception* **6** 595 – 600

Kanizsa G, Gerbino W, 1981 "Il completamento amodale fra vedere e pensare" *Giornale Italiano di Psicologia* **8** 279 – 307

Leeuwenberg E, 1971 "A perceptual coding language for visual and auditory patterns" *American Journal of Psychology* **84** 307 – 349

Leeuwenberg E, Helm P van der, 1991 "Unity and variety in visual form" *Perception* **20** 595 – 622

Lowe D G, 1987 "Three dimensional object recognition from single two dimensional images" *Artificial Intelligence* **31** 355 – 395

Massironi M, 1988 "A new visual problem: Phenomenic folding" *Perception* **17** 681 – 694

Metelli F, 1960 "Morfologia dei fenomeni di completamento nella percezione visiva", in *Gestalthaftes Sehen* Ed. F Weinhandel (Darmstadt: Wissenschaftliche Buchgesellschaft) pp 266 – 278

Michotte A, 1946 *La Perception de la Causalité* (Louvain: Institut Supérieur de Philosophie)

Michotte A (Ed.), 1962 *Causalité, Permanence et Réalité Phénoménales* (Louvain: Publications Universitaires)

Navon D, 1977 "Forest before trees: the precedence of global features in visual perception" *Cognitive Psychology* **9** 353 – 383

Perkins D N, 1972 "Visual discrimination between rectangular and nonrectangular parallelopipeds" *Perception & Psychophysics* **12** 396 – 400

Perkins D N, 1976 "How good a bet is a good form?" *Perception* **5** 393 – 406

Perkins D N, Cooper R, 1980 "How the eye makes up what the light leaves out", in *The Perception of Pictures.* Volume II. *Dürer's Devices: Beyond the Projective Model of Pictures* Ed. M Hagen (New York: Academic Press) pp 95 – 130

Proffitt D R, Rock I, Hecht H, Schubert J, 1992 "The stereokinetic effect and its relations to the kinetic depth effect" *Journal of Experimental Psychology: Human Perception and Performance* **18** 3 – 21

Ullman S, 1984 "Maximizing rigidity: The incremental recovery of 3-D structure from rigid and nonrigid motion" *Perception* **13** 255 – 274

Wagemans J, Kolinsky R, 1994 "Guest editorial: perceptual organisation and object recognition—POOR is the acronym, rich the notion" *Perception* **23** 371 – 382

Witkin A P, Tanenbaum J M, 1983 "On the role of structure in vision", in *Human and Machine Vision* Eds A Rosenfeld, B Hope, J Beck (New York: Academic Press) pp 481 – 543