

# Ambiguity and structural information in the perception of reversible figures

FRANCESCO MASULLI and MASSIMO RIANI  
*University of Genoa, Genoa, Italy*

The perspective reversals elicited by a set of drawings based on the Mach truncated pyramid are examined. We obtained each pattern of the set from the previous one by adding to it some graphic cues, which were easily integrated into one of the two competing interpretations, thus reducing step by step the ambiguity of the basic pattern. The phenomenological model, proposed to link the mean times of both alternative interpretations with their complexities, is in close agreement with our experimental data. Furthermore, different aspects of such data are well described by the model equations: The measure of the prevalence of the supported interpretation is well correlated with the difference in complexity between the two alternative interpretations; the two different trends of the mean time of the unfavored interpretation, found in the data obtained from different observers as a function of the various patterns of the set, are well fitted by the model without the need for any specific additional hypothesis.

The recurrent perceptual changes in the appearance of an ambiguous figure (e.g., the Necker cube, the Schroeder staircase, or Rubin's vase) constitute an intriguing phenomenon in the study of visual perception. Many experiments have demonstrated that the observer's attitude and specific experimental instructions can influence, to a great extent, the way in which perceptual alternation occurs (Ammons, Ulrich, & Ammons, 1959; Girgus, Rock, & Egatz, 1977; Radilová & Radil-Weiss, 1984; Radilová, Riani, Tuccio, Radil, & Borsellino, 1983; Riani, Tuccio, Borsellino, Radilová, & Radil, 1986). Other experiments have focused attention on the passive aspects of the phenomenon (Babich & Standing, 1981; Hochberg, 1950; Howard, 1961; Orbach, Ehrlich, & Heath, 1963).

Two well-known competing theories have been developed to explain perceptual fluctuations: one is based on a passive process of neuronal fatigue (Attneave, 1971; Kohler & Wallach, 1944), and another is based on an active process of hypothesis testing (Gregory, 1974; Vickers, 1972). Neither theory can account for the varied phenomenology or the significance of the phenomenon; therefore, in recent years, there has been a renewal of interest in experimental works (Phillipson & Harris, 1984; Radilová, Radil, & Havranek, 1984; Reisberg & O'Shaughnessy, 1984; Riani, Oliva, Selis, Ciurlo, & Rossi, 1984; Toppino & Long, 1987; von Grünau, Wiggins, & Reed, 1984; Wilton, 1985). Some attempts

have been made to merge the two theories (Long, Toppino, & Kostenbauder, 1983) and to explain the multi-stable perception phenomenon in the context of a computational approach (Hoffman, 1983; Hoffman & Richards, 1984) or of neural network models (Aicardi & Borsellino, 1987; Kawamoto & Anderson, 1985).

Riani et al. (1986) proposed that the ambiguity of a pattern and the stability of its alternative interpretations are correlated with the degree of competitiveness of such interpretations. This degree of competitiveness might, in turn, be related to the complexities of the two alternative percepts.

In the present work, we took this approach further to reach an analytical formulation of the dependence of the mean durations of the two interpretations on their complexities, measured according to Leeuwenberg's (1971) structural information theory, which subsequently was developed, in terms of experimental and mathematical aspects, by several other researchers (Buffart & Leeuwenberg, 1983; Buffart, Leeuwenberg, & Restle, 1983; Burigana & Lucca, 1985; Leeuwenberg & Buffart, 1983; Martinoli, Masulli, & Riani, 1988; van der Helm & Leeuwenberg, 1986). Some preliminary results in the study of perceptual fluctuations in perspective alternation have been analyzed by Riani and Masulli (1983), using this theory. According to the structural information theory, the different interpretations, or percepts, associated with an ambiguous figure (i.e., its different semantic meanings) can be represented by different "strings" of syntactic elements, each giving rise to a coded representation of a percept. The syntactic elements of the structural information theory are angles, lines, and some operators, all describing the regularities that are present in figures. The end codes of the theory are the shortest strings; they can be obtained through the application of syntactic operators to the initial strings, in which all the angles and

---

This research was supported by the Ministero della Pubblica Istruzione and by the National Group of Cybernetics and Biophysics of the Consiglio Nazionale delle Ricerche. We would like to thank M. Scotto for useful discussions and the two anonymous referees for their helpful comments. We are grateful to E. Gaggero for the design and realization of the tape-recording apparatus. Requests for reprints may be sent to M. Riani, Dipartimento di Fisica dell'Università di Genova, via Dodecaneso 33, 16146 Genova, Italy.

lines of a semantic interpretation of the figure are represented. In the Appendix, we give a brief outline of this theory, with some details concerning its application to our experiment.

In the present study, we analyzed the phenomenon of perceptual alternation, with regard to the principle of *perceptual economy*, or minimum tendency, that is present in various theoretical approaches to the psychology of perception (see Musatti, 1924, for a pioneer work; and Hatfield & Epstein, 1985, for a recent review). According to this principle, perceived objects tend to appear as the simplest phenomenological shapes (Hochberg & Brooks, 1960; Hochberg & McAllister, 1953). In order to make a pattern appear ambiguous, it is necessary that two equally simple shapes, or interpretations, be associated with it. The two interpretations should correspond to two different, but nearly equivalent, minima in the perceptual space of the observer. Under continuous viewing, the perception of the pattern then will oscillate between the two alternative percepts with equal probabilities. In other words, the lengths of the time intervals in which the two percepts, *A* and *B*, are alternately viewed will give two nearly equal mean values,  $\bar{t}_A$  and  $\bar{t}_B$ , although such individual time intervals will be stochastically distributed as a gamma distribution (Borsellino, De Marco, Allazetta, Rinesi, & Bartolini, 1972).

On the other hand, if a pattern has two interpretations that are not completely equivalent, it can still appear ambiguous, but the probabilities and the durations of the two percepts will be different. For example, in Fisher's (1967) set of "Man or Girl" drawings, the first and final figures were perceived as totally nonambiguous (the first represents a man's face, and the final figure is the outline of a woman), while the two central ones were perceived as either alternative interpretation, with equal probability. For the remaining figures, one of the two interpretations ("man" or "girl") is clearly the dominant choice, although the alternative one is also considered.

The lack of balance in the perceptual duration of the two alternative percepts can be regarded as an a posteriori measure of the nonequivalence of the percepts. Hence, one could derive the degree of ambiguity of a pattern from the a posteriori probabilities of the alternating percepts, *A* and *B*, namely,  $p_A = \bar{t}_A / (\bar{t}_A + \bar{t}_B)$  and  $p_B = \bar{t}_B / (\bar{t}_A + \bar{t}_B)$ , as the  $p_B/p_A$  ratio between the two probabilities, which obviously is equal to the  $\bar{t}_B/\bar{t}_A$  ratio. In fact, this ratio ranges from 1 for a pattern with two alternative interpretations with equal probability ( $p_A = p_B = 1/2$ ) to 0 for an unambiguous pattern with only one accepted interpretation ( $p_A = 1, p_B = 0$ ).

To estimate a priori the ambiguity of a pattern, one should measure the degree of relative simplicity of the alternative interpretations to obtain a quantitative evaluation of the differences in the minima of the perceptual space associated with such interpretations.

We tried to achieve such a measure by applying the structural information theory to evaluate the complexities of the different percepts elicited by a set of patterns

for which the two competitive interpretations were increasingly unbalanced. We obtained the set of patterns by adding to the basic pattern some details easily integrated into only one interpretation, thus favoring this interpretation.

To attain this goal, we (Riani et al., 1986) assumed that the mean duration of the two alternative percepts, *A* and *B*, of an ambiguous pattern depends on the interaction of

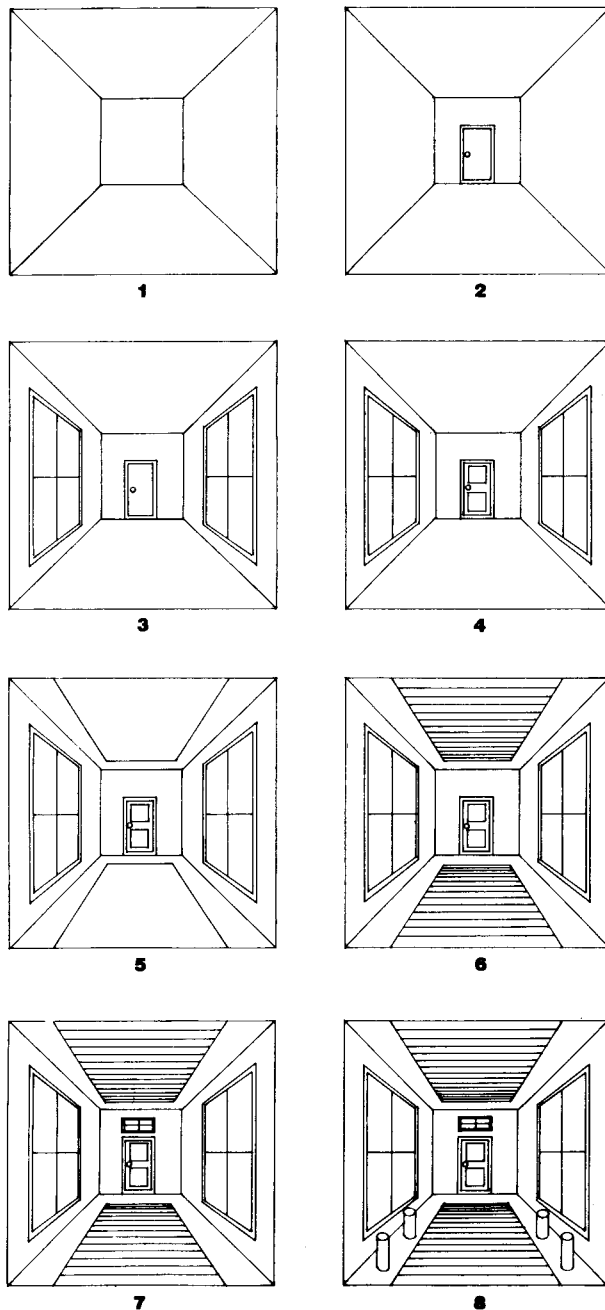


Figure 1. The set of eight stimuli used in the experiment, denoted by the numbers 1 to 8. For the six stimuli, denoted in the paper by the letters A, B, C, D, E, and F, there exists the following correspondence: A=1, B=2, C=3, D=5, and E=6, while Stimulus F corresponds to Stimulus 8, without the little window over the front door.

two competitive contributions: (1) the resistance offered by the *actual* percept against its replacement with the alternative one and (2) the ability of the *nonactual* percept to impose itself as an actual percept. In other words,

the mean time of the duration of a percept (*A* or *B*) can be considered to be proportional to its stability ( $S_A$  or  $S_B$ ), and inversely proportional to the force of the antagonistic percept ( $F_B$  or  $F_A$ ) (i.e.,  $\bar{t}_A \propto S_A/F_B$  or  $\bar{t}_B \propto S_B/F_A$ ).

**Table 1**  
Mean Times and Standard Errors (in sec) of the Alternative Interpretations Related to the Two Sets of Ambiguous Patterns (1 to 8 or A to F).

		Patterns								
		1=A	2=B	3=C	4	5=D	6=E	F	7	8
<b>Subjects</b>										
1	$\bar{t}_A$	5.07	5.46	6.99	-	7.21	8.18	5.83	-	-
	$e_A$	.13	.14	.22	-	.28	.46	.36	-	-
	$\bar{t}_B$	8.37	9.84	16.35	-	30.71	50.70	306.80	-	-
	$e_B$	.26	.27	.62	-	1.75	3.84	39.30	-	-
2	$\bar{t}_A$	13.83	38.85	40.29	-	49.47	49.58	70.00	-	-
	$e_A$	.86	4.76	4.40	-	5.74	7.54	10.50	-	-
	$\bar{t}_B$	16.24	42.52	62.38	-	86.67	100.78	114.80	-	-
	$e_B$	1.03	3.63	6.48	-	8.25	11.38	19.90	-	-
3	$\bar{t}_A$	13.64	10.41	10.36	-	6.84	4.02	5.17	-	-
	$e_A$	.62	.67	.53	-	.60	.44	.40	-	-
	$\bar{t}_B$	19.31	22.43	27.38	-	37.81	62.92	53.96	-	-
	$e_B$	.89	1.51	1.36	-	2.06	7.03	3.32	-	-
4	$\bar{t}_A$	3.31	3.29	3.03	-	3.09	3.21	2.57	-	-
	$e_A$	.07	.08	.11	-	.20	.23	.29	-	-
	$\bar{t}_B$	3.70	3.93	14.96	-	43.03	75.90	108.80	-	-
	$e_B$	.06	.08	.93	-	3.77	12.03	19.60	-	-
5	$\bar{t}_A$	4.26	4.41	3.88	-	3.01	3.04	2.34	-	-
	$e_A$	.25	.20	.18	-	.16	.13	.10	-	-
	$\bar{t}_B$	10.22	8.25	9.82	-	13.89	12.16	23.84	-	-
	$e_B$	.63	.37	.47	-	.88	.62	1.66	-	-
6	$\bar{t}_A$	9.50	14.39	15.83	22.46	17.97	19.26	-	31.55	30.25
	$e_A$	.36	.74	.79	2.57	1.56	1.39	-	3.18	2.56
	$\bar{t}_B$	10.96	17.21	20.36	34.79	27.88	27.16	-	42.91	56.18
	$e_B$	.50	.96	1.19	2.99	2.26	2.24	-	4.29	5.31
7	$\bar{t}_A$	2.82	2.76	2.44	2.75	2.32	2.39	-	1.94	1.66
	$e_A$	.04	.06	.07	.09	.06	.07	-	.07	.07
	$\bar{t}_B$	3.53	5.71	13.34	14.60	17.63	14.83	-	33.12	93.67
	$e_B$	.07	.20	.69	.90	.97	.73	-	2.52	11.32
8	$\bar{t}_A$	4.92	5.26	4.89	5.04	8.47	10.51	-	7.56	12.47
	$e_A$	.16	.20	.18	.20	.53	.86	-	.42	1.01
	$\bar{t}_B$	9.81	14.10	15.74	13.46	21.43	22.74	-	21.95	27.19
	$e_B$	.41	.66	.78	.52	1.31	.95	-	1.27	1.72
<b>Group Means</b>										
M3	$\bar{t}_A$	5.75	7.47	7.72	10.08	9.59	10.72	-	13.68	14.79
	$e_A$	1.39	2.50	2.91	4.40	3.22	3.44	-	6.42	5.89
	$\bar{t}_B$	8.10	12.34	16.48	20.95	22.31	21.58	-	32.66	59.01
	$e_B$	1.63	2.43	1.46	4.90	2.12	2.55	-	4.28	13.61
M5	$\bar{t}_A$	8.02	12.48	12.91	-	13.92	13.61	17.18	-	-
	$e_A$	1.17	3.35	3.48	-	4.47	4.52	6.61	-	-
	$\bar{t}_B$	11.57	17.37	26.18	-	42.42	60.49	121.64	-	-
	$e_B$	1.40	3.50	4.75	-	6.05	7.33	24.66	-	-
<b>Complexities</b>										
$C_A$		8	17	36	39	44	60	77	72	89
$C_B$		6	15	28	31	34	38	48	50	60

Note—Individual subjects are denoted by natural numbers. *M3* and *M5* denote the results of the group means obtained by averaging the mean times of Subjects 6, 7, and 8 for the set of eight patterns (*M3*) and of Subjects 1 to 5 for the set of six patterns (*M5*). In the last two rows, the complexities of the two alternative interpretations of the different patterns are reported.

The dominant percept will be the one with higher stability and force. If both the stability and the force of a percept are assumed to be monotonic functions of the complexity of the percept itself, according to the minimum tendency principle, a relationship will result between the mean times of the two percepts and their complexities. To determine this relationship, we performed the following experiment.

## METHOD

### Apparatus

The subjects were seated in a room illuminated by daylight, and were separated from the experimenter by a screen. An additional 60-W lamp illuminated the stimulus pattern from above and behind the subject's head. Each reversible figure, drawn in black india ink in the center of a white sheet of paper (70×70 cm<sup>2</sup>), was placed in front of each subject at a distance of 80 cm. The subject reported each perspective reversal by moving the index finger of his/her right hand horizontally through a thin beam that projected to a phototransistor. We recorded the output signals of the phototransistor and sent them to a microcomputer, using a suitable program for data acquisition and processing (Carlini & Masulli, 1983), in order to measure the time intervals corresponding to the two different percepts in the sequence of perspective reversals. Subsequently, we separately processed the time intervals related to the *A* and *B* percepts, thus obtaining the mean time and the standard deviation of each percept for all the patterns and subjects.

### Stimuli

During the experiment, two sets of six and eight ambiguous stimuli based on the Mach truncated pyramid were used. The two alternative interpretations of this pattern, labeled as Percept *A* and Percept *B*, corresponded to a convex pyramid and to a concave room. We obtained the stimuli of each set by adding to the basic pattern some graphic details (e.g., windows and carpets) that were easily integrated into the pattern interpreted as a room. In the pyramid interpretation, the subjects perceived "windows" and "carpets" as ornaments on the pyramid sides and the "wastepaper baskets" of the last figure as chimney pots. Figure 1 shows the complete set of eight patterns. The original set of six patterns was made up of five of these patterns (*A* = 1, *B* = 2, *C* = 3, *D* = 5, *E* = 6) and one modified pattern (*F* = 8, without the little window over the front door). The sides of the outer square of each stimulus were 25 cm long and the sides of the inner square were 8 cm long; the line width was 0.8 mm.

All patterns of each set were interpreted by all of the subjects as reversible perspective ones. In fact, the average perceptual duration of the pyramid interpretation was always significantly different from zero for all the patterns and subjects.

Note that the graphic details added to the small square of the Mach pyramid differ from those added to the surrounding trapezoids in their perceptual qualities during the phenomenon of perspective reversal. In fact, the former details can be reported as identical geometrical shapes in the two competitive global interpretations (i.e., room and pyramid). The object added to the lateral or upper and lower sides, however, can be reported in different ways, depending on the two alternative percepts (i.e., as rectangles in the room interpretation and as trapezoids in the pyramid interpretation). For this reason, as explained in greater detail in the Appendix, the objects added to the small central square exhibit the same complexities in both interpretations, while the others increase the complexities of the two interpretations to a different degree; therefore, the details added to the lateral sides can be expected to have different unbalancing effects on the perceptual minima from those due to the details added to the central square.

As a matter of fact, the complete set of eight patterns was obtained by adding to the original set two patterns with added details in the small square only. The two patterns were added to verify that this kind of cue has no unbalancing effect, as suggested by the analysis of the results on Patterns 1 and 2.

### Subjects and Procedure

Eight university students took part in the experiment and were paid for their participation. Three females and 2 males were tested with the set of six patterns, and 3 males were tested with the complete set of eight patterns. All of the subjects had acquired some experience with the phenomenon of perceptual reversals, having participated in previous experiments with Necker cube alternation in our laboratory. The subjects' mean age was about 22 years.

For each pattern, the subjects participated in at least two experimental sessions, each lasting about 1 hr, on different days. Each session was divided into two 20-min runs, usually separated by a rest period of 10 to 15 min. The observers looked binocularly at the figures and were instructed to assume a passive attitude toward perspective alternation, that is, without making any effort to keep one percept or to favor reversals.

Because we wanted to consider only the stationary phase of the phenomenon and to avoid any learning phase, thus allowing us to collect experimental data in such conditions as to obtain the highest reliability and reproducibility (Borsellino et al., 1972; De Marco et al., 1977), the first 20-min run of each pattern was disregarded in the processing and was used for practice only. Furthermore, in statistical analysis, we disregarded the first measurements in all other runs. In this way, we were almost sure that our data referred to the stationary phase of the reversal phenomenon. We checked all sessions for stationariness using the method utilized by De Marco et al. (1977), and we found that the stationary phase had already been reached by all but a few subjects on the last pattern of each set. In this case, we rejected these sessions, and the subject was tested again on the following days, until at least three runs were obtained in the stationary phase of the phenomenon. Using this procedure, we collected a different number of reversals for each pattern, because the reversal rate decreases with the decreasing ambiguity of the different patterns. We collected at least 30 perspective reversals (during other sessions, if necessary) to obtain a reasonable amount of statistical reliability on all mean times.

No fixation point was present in the patterns and no particular pattern-scanning strategy was suggested to the subjects. In this way, we collected data in a more natural condition, without suggesting any precise point to the subjects and without tiring them with the additional task of keeping their eyes on a fixation point.

Table 2  
Values of the Parameters Used in the Model

Subjects	Parameters			
	j	a	k	b
1	4.84	2.49 10 <sup>-2</sup>	3.51	5.88 10 <sup>-2</sup>
2	16.20	-.85 10 <sup>-2</sup>	13.10	2.24 10 <sup>-2</sup>
3	19.25	3.55 10 <sup>-2</sup>	13.18	3.61 10 <sup>-2</sup>
4	2.28	5.30 10 <sup>-2</sup>	2.44	8.59 10 <sup>-2</sup>
5	7.76	2.79 10 <sup>-2</sup>	4.33	3.13 10 <sup>-2</sup>
6	10.58	-.49 10 <sup>-2</sup>	8.79	1.52 10 <sup>-2</sup>
7	3.26	3.51 10 <sup>-2</sup>	2.54	5.27 10 <sup>-2</sup>
8	9.30	-.61 10 <sup>-2</sup>	3.91	7.13 10 <sup>-3</sup>
Group Means				
<i>M3</i>	7.61	3.44 10 <sup>-4</sup>	4.96	2.06 10 <sup>-2</sup>
<i>M5</i>	9.26	1.59 10 <sup>-2</sup>	7.14	4.27 10 <sup>-2</sup>

Note—*a* and *b* are expressed in (unit of complexity)<sup>-1</sup>; *j* and *k* are expressed in seconds.

**RESULTS**

In Table 1, the mean times and the standard errors of the two alternative interpretations ( $A = \text{pyramid}$ ;  $B = \text{room}$ ) are given for the 5 subjects who were tested with six patterns and for the 3 subjects who were tested with the complete set of eight patterns. Table 1 also presents the results that were obtained when we averaged

the mean times of the 3 subjects on the eight patterns ( $M3$ ) and those of the 5 subjects on the six patterns ( $M5$ ). We will refer to  $M3$  and  $M5$  as the group means, and the related results will be used to describe the general trends of the experimental data, averaging the differences in the individual performances. In this way, the group means represent the global behavior of our experimental data, but they reveal no details of the individual trends. The

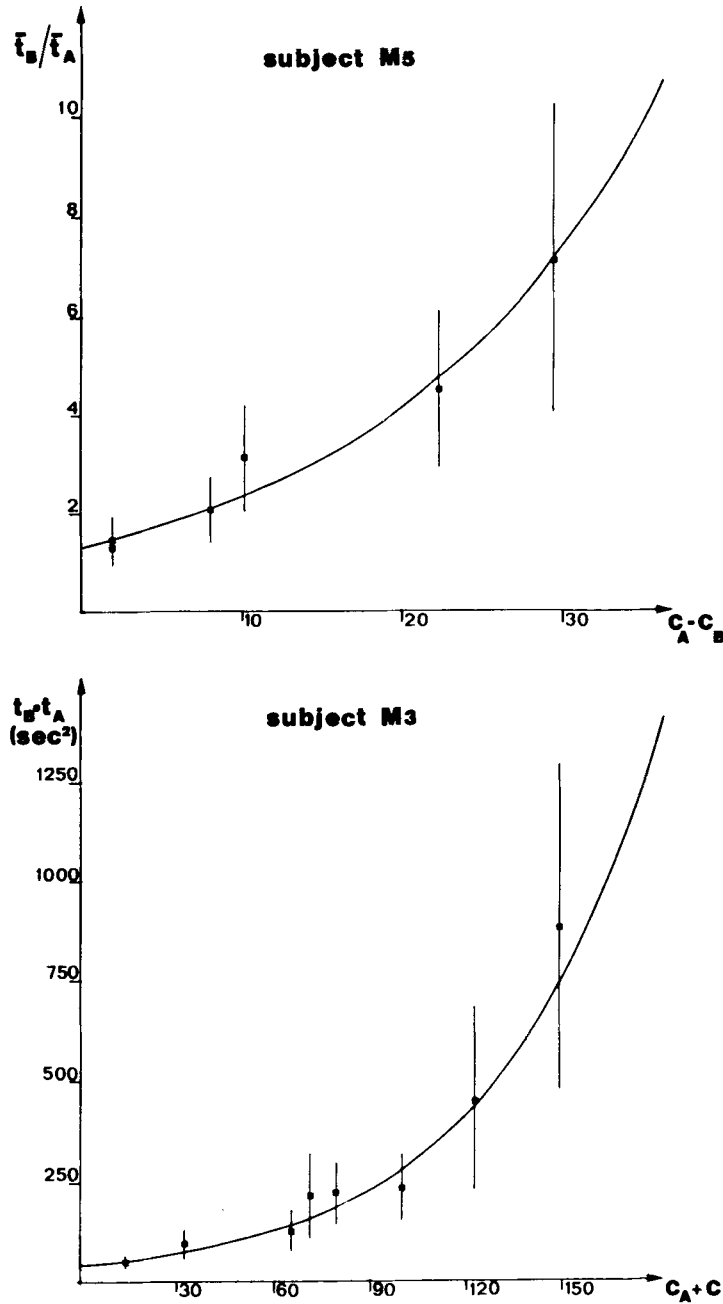
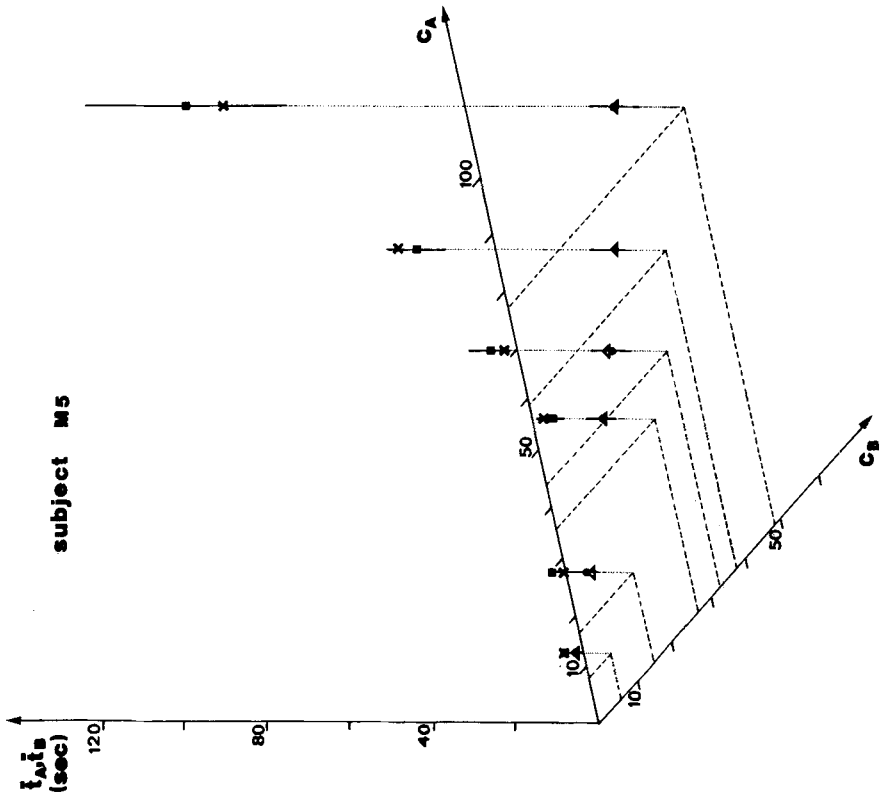
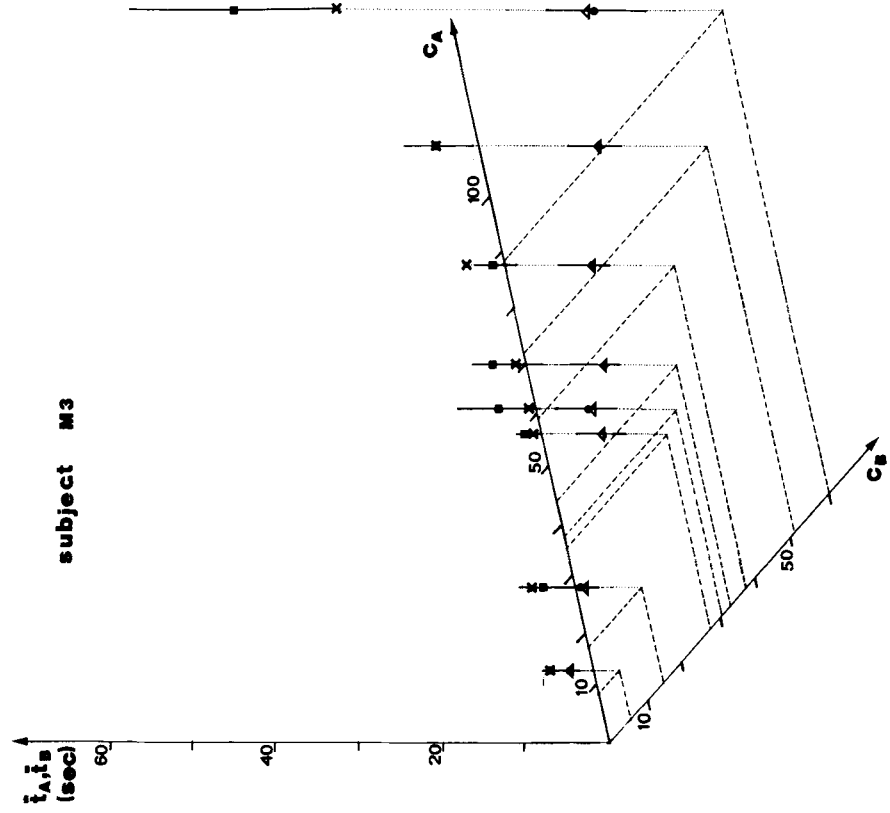


Figure 2. Upper: The fit of the experimental values of the  $\bar{t}_B/\bar{t}_A$  ratio as a function of the difference in the complexities,  $C_A$  and  $C_B$ , of the two alternative percepts, A and B, for the group mean  $M5$ . Lower: The fit of experimental values of the  $\bar{t}_A \cdot \bar{t}_B$  product as a function of the sum of the complexities of the two alternative percepts for the group mean  $M3$ . The complexities,  $C_A$  and  $C_B$ , of the two different percepts have been obtained by the structural information theory.



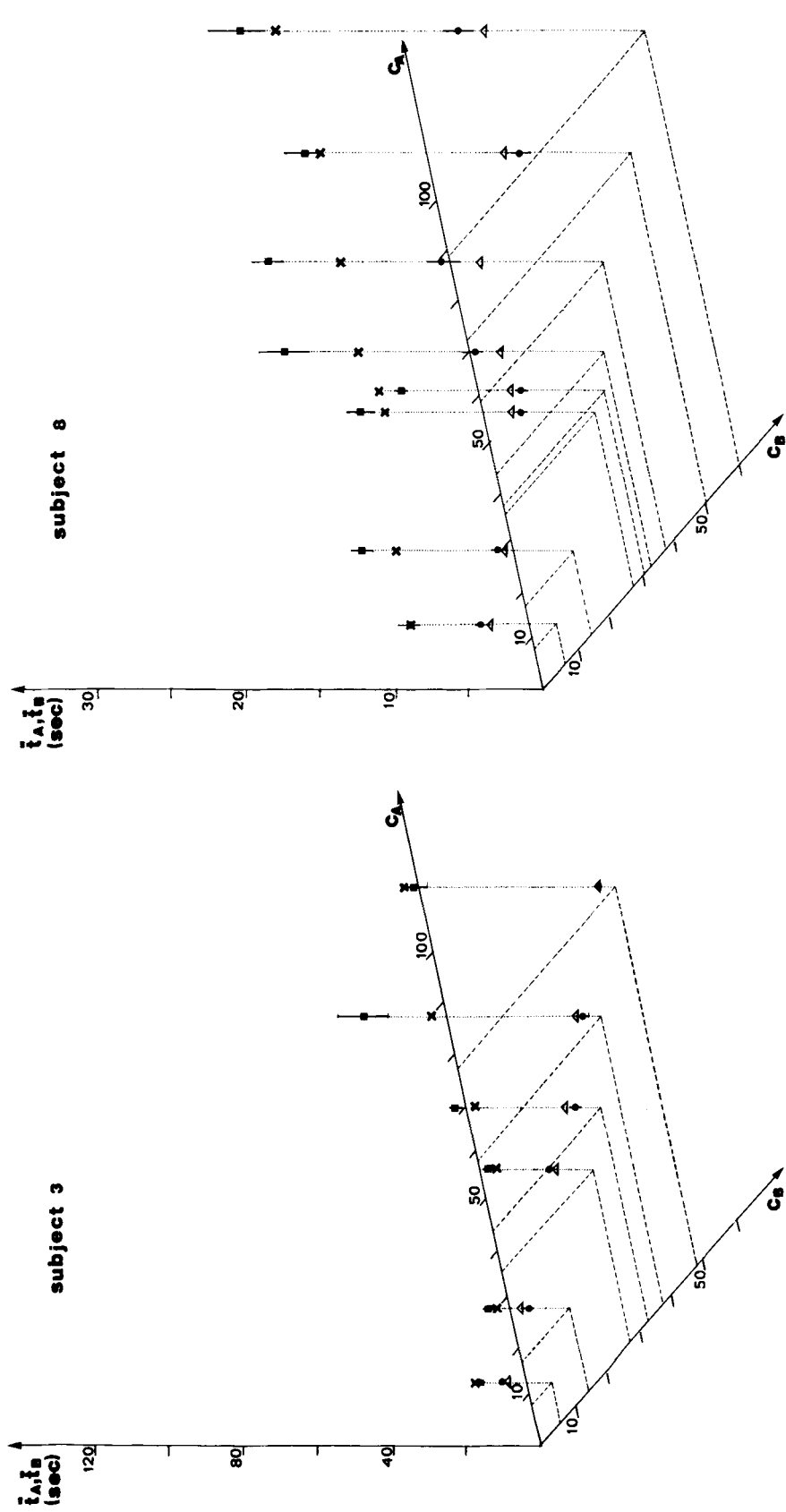


Figure 3. Comparison between the experimental values ( $\bar{t}_A = \phi, \bar{t}_B = \psi$ ), and the theoretical values ( $\bar{t}_A = \Delta, \bar{t}_B = X$ ), derived from the model, of the mean times of the two alternative percepts ( $\bar{t}_A, \bar{t}_B$ ) for the patterns used in the experiment. The two upper graphs show the comparisons for the two group means, M3 and M5, in which every individual fluctuation has been canceled by the averaging procedure; the comparisons are quite favorable. The lower graphs show comparisons of 2 subjects. The trend of the experimental data is well described by the theoretical values, although some spread in the data is evident. The 2 subjects have been selected to illustrate the two different trends of the mean time,  $\bar{t}_A$ , of the unfavored percept ("pyramid"), as a function of the patterns of the sets. The decrease in the  $\bar{t}_A$  values for Subject 3 and the increase for Subject 8 are very well described by the model, without the need for any additional hypothesis.

standard errors of the group means are calculated from the differences between the individual mean times and the total average time.

In Table 1, note the marked differences in the individual reversal rates. This high variability concerning the phenomenon of perceptual alternation, which previously has been reported (Borsellino et al., 1982; Borsellino et al., 1972; Künnapas, 1969; Washburn, Mallay, & Naylor, 1931), is due to the presence of different kinds of subjects in a normal population: some exhibit fast and regular reversal rates, others exhibit slow and irregular perceptual alternations.

Data analysis points out that the mean duration,  $\bar{t}_B$ , of the room interpretation of the various patterns increased with the addition of details. The increase was very regular for the group means and more fluctuating, although still very evident, for each subject.

With regard to the dependence of  $\bar{t}_A$  on the patterns, two kinds of behavior could be seen: For some subjects, the mean times of the pyramid interpretation decreased regularly with the addition of details supporting the room percept, while, for others, there was an increase in the mean times of both interpretations. For these subjects, the increase in  $\bar{t}_A$  was slower than it was in  $\bar{t}_B$ ; nevertheless, the increase was clearly greater than was found in the statistical scatter of data. Thus, the introduction of cues supporting one of the two alternative percepts can have two distinct effects on perceptual alternation. For one behavior, these cues act differently on the two percepts, increasing the mean time of the favored interpretation and decreasing the mean time of the unfavored one. For the other behavior, the added details seem to introduce a stabilizing effect on both percepts, even if the most conspicuous outcome is an evident loss of balance in favor of the room percept. In their study on the effect of a stereoscopic cue on the perception of ambiguous patterns, Borsellino, Oliva, Riani, and Tuccio (1979) reported a similar twofold behavior.

To test the hypothesis that the details added to the front square of some patterns have no unbalancing effect on perceptual alternation, we made a careful analysis of the  $\bar{t}_B/\bar{t}_A$  ratio for the different patterns and for each subject. As we mentioned earlier, this ratio is equal to the  $p_B/p_A$  ratio between the a posteriori probabilities of the alternating percepts and can be considered a measure of the ambiguity of a pattern or, similarly, of the balance of the alternative interpretations.

By performing the analysis of variance for each subject, with the  $\bar{t}_B/\bar{t}_A$  ratio as the controlled factor, we obtained the  $F$  values of the Snedecor test, which ranged from  $[F(7,16) = 4.34, \text{ for Subject 6}]$  to  $[F(5,12) = 50.35 \text{ for Subject 5}]$ . All of the  $F$  values indicated that the  $\bar{t}_B/\bar{t}_A$  ratio was dependent on the pattern of the set, with a probability higher than 0.990 for Subject 6 and higher than 0.995 for each of the other subjects. However, the experimental  $\bar{t}_B/\bar{t}_A$  ratios obtained for the couples of Patterns 1-2, 3-4, and 6-7 were equal, within standard errors, with a good efficiency  $E$ . In fact, the ratio between

the number of favorable cases and the total number of cases derived from such patterns was

$$E = 12/14 = 85.7\%$$

at a  $3 \sigma/\sqrt{N}$  level of confidence or

$$E' = 10/14 = 71.4\%$$

at a  $2 \sigma/\sqrt{N}$  level of confidence.

Therefore, considering that these are ratios of experimental data, with large statistical errors, we have concluded that our hypothesis (of no unbalancing effect of details added to the front square) has been verified, at least to a first approximation.

## DISCUSSION

### The Phenomenological Model

We now analyze our experimental data under the hypothesis that perceptual alternation is the result of the competition of two contributions linked to the stability ( $S$ ) of the actual percept and to the force ( $F$ ) of the antagonistic percept. In this context, the mean time of the room percept,  $\bar{t}_B$ , will be directly proportional to  $S_B$  and inversely proportional to  $F_A$ , while  $\bar{t}_A$  will be proportional to  $S_A/F_B$ .

If the degree of stability and the force of both alternative interpretations are exponentially decreasing functions of their complexities, then

$$\begin{aligned} S_A &= m \exp(-aC_A) & F_A &= q \exp(-bC_A) \\ S_B &= p \exp(-aC_B) & F_B &= n \exp(-bC_B) \end{aligned} \quad (1)$$

where  $a$ ,  $b$ ,  $m$ ,  $n$ ,  $p$ , and  $q$  are constants related to the two alternative percepts, pyramid and room.

Note that parameters  $m$ ,  $n$ ,  $p$ , and  $q$  represent the factors of the stability and force not directly dependent on complexity. Ideally, these parameters are functions of all other variables affecting the phenomenon of perceptual alternation. These variables can be both of a physical nature (e.g., visual angle, Borsellino et al., 1982; drug consumption, Phillipson & Harris, 1984), and of the perceptual or psychological kind (e.g., familiarity, Girgus et al., 1977, etc.).

The assumption that the trends of the degree of stability and force are exponential functions of the complexity of the related percepts is supported by the fact that the perceptual processes involved in pattern recognition are similar to the procedures employed in optimization problems. For such problems, all known methods for determining an optimal route to their solution require a computing effort that increases exponentially with the number ( $N$ ) of the variables involved (Kirkpatrick, Gelatt, & Vecchi, 1983). Hence, we deem it reasonable to hypothesize that, when one increases the complexity of a percept (i.e., the number of different items to be connected in order to give rise to the corresponding interpretation of the pattern), the related force and degree of stability decrease exponentially, as we assumed in Set of Equations 1. Moreover, an exponential function is the simplest form of relation (linking a percept's force and degree of stability to its com-



plexity) by which the ratio  $\bar{t}_B/\bar{t}_A$  between the mean times (i.e., the measure of the unbalancing between percepts) becomes a function of the difference in the complexities of the alternative percepts, thus supporting the experimental results obtained for the couples of Patterns 1-2, 3-4, and 6-7, as previously reported.

In accordance with the above hypotheses, we can express  $\bar{t}_A$  and  $\bar{t}_B$  in terms of these stabilities and forces, using Set of Equations 1:

$$\begin{aligned} \bar{t}_A &= S_A/F_B = m/n \exp(-aC_A + bC_B) \\ &= k \exp(-aC_A + bC_B) \end{aligned} \quad (2)$$

$$\begin{aligned} \bar{t}_B &= S_B/F_A = p/q \exp(-aC_B + bC_A) \\ &= j \exp(-aC_B + bC_A). \end{aligned} \quad (3)$$

Both expressions should be fitted by experimental data, using the same values of the parameters a and b, as in the model. To facilitate the best-fit procedure, and to avoid possible problems that can arise from the above constraint, we chose to fit first the  $\bar{t}_B/\bar{t}_A$  ratio between the measured mean times, and then the product of the same quantities,  $\bar{t}_A \cdot \bar{t}_B$ . In fact, as a consequence of the model hypotheses, the following relationships will be exponential functions of only two distinct parameters:

$$\begin{aligned} \bar{t}_B/\bar{t}_A &= j/k \exp[(a+b)C_A - (a+b)C_B] \\ &= l \exp[r(C_A - C_B)] \end{aligned} \quad (4)$$

and

$$\begin{aligned} \bar{t}_A \cdot \bar{t}_B &= k \cdot j \exp[(b-a)C_A + (b-a)C_B] \\ &= d \exp[s(C_A + C_B)]. \end{aligned} \quad (5)$$

The results are very well fitted by the model in both cases for the group means, as shown in Figure 2, while for the individual subjects the fit is rougher but satisfactory enough.

The values of the parameters of Equations 4 and 5 were estimated for each subject and for the group means, M3 and M5, with a best-fit technique, using the program MINUIT (James & Ross, 1977). We then calculated the other parameters of the model, as given in Table 2, which allowed us to compare the experimental values of  $\bar{t}_A$  and  $\bar{t}_B$  with the theoretical ones obtained by the Model Formulas 2 and 3. In fact, the knowledge of the values of the parameters j, k, a, and b allowed us to calculate the theoretical values of  $\bar{t}_A$  and  $\bar{t}_B$  as exponential functions of the complexities of both alternative percepts.

Figure 3 shows the results of this comparison for group means, M3 and M5, and for 2 individual subjects. As in the comparison of the ratio and product of the mean times, the agreement between theoretical and experimental data is quite good for the group means, whose fluctuations are smoothed, and good enough for all the subjects. We must point out, however, that the comparison between the experimental data and the theoretical values of the mean times,  $\bar{t}_A$  and  $\bar{t}_B$ , is statistically more meaningful than is the comparison of the  $\bar{t}_B/\bar{t}_A$  ratios, because of the higher

percent errors affecting the calculated values of these ratios.

Thus, despite the smaller statistical errors in the data and despite the fact that the theoretical values of  $\bar{t}_A$  and  $\bar{t}_B$  were not the outcome of a specific fitting procedure but of a mathematical deduction based on the model, the good result of the comparison of the mean times suggests that the general aspects of the phenomenon of perceptual alternation, as pointed out by our experiment, are very well simulated by the proposed model.

Furthermore, in the case of individual subjects (that is, in the most fluctuating situations), the model provides more proof of its capability for representing experimental data. In fact, as mentioned earlier, the subjects showed two different trends of the mean time,  $\bar{t}_A$ , related to the unfavored percept for the various patterns—more precisely,  $\bar{t}_A$  decreased for 5 subjects and increased for the other 3 subjects.

This different effect of the addition of cues supporting one of the two percepts on the mean durations of the alternative one is represented in Figure 3, for which we have chosen 2 subjects belonging to the two groups showing different  $\bar{t}_A$  trends:  $\bar{t}_A$  decreases for Subject 3 and increases for Subject 8. Experimental data are very well described by the model for these and all other subjects. It is worth noting again that this agreement was achieved without using any specific fitting procedure for the mean times and any additional hypothesis.

Therefore, we can conclude that all of the aspects of the phenomenon of perceptual alternation, tested in our experiment, are well described by the proposed model.

## REFERENCES

- AICARDI, F., & BORSELLINO, A. (1987). Statistical properties of flip-flop processes associated to the chaotic behavior of systems with strange attractors. *Biological Cybernetics*, *55*, 377-385.
- AMMONS, R. B., ULRICH, P., & AMMONS, C. H. (1959). Voluntary control of perception of depth in a two-dimensional drawing. *Proceedings of the Montana Academy of Sciences*, *19*, 160-168.
- ATTNEAVE, F. (1971). Multistability in perception. *Scientific American*, *225*, 62-71.
- BABICH, S., & STANDING, L. (1981). Satiation effects with reversible figures. *Perceptual & Motor Skills*, *52*, 203-210.
- BORSELLINO, A., CARLINI, F., RIANI, M., TUCCIO, M. T., DE MARCO, A., PENENGO, P., & TRABUCCO, A. (1982). Effects of visual angle on perspective reversal for ambiguous patterns. *Perception*, *11*, 263-273.
- BORSELLINO, A., DE MARCO, A., ALLAZETTA, A., RINESI, S., & BARTOLINI, B. (1972). Reversal time distribution in the perception of visual ambiguous stimuli. *Kybernetik*, *10*, 139-144.
- BORSELLINO, A., OLIVA, G. A., RIANI, M., & TUCCIO, M. T. (1979). Effects of a stereoscopic cue on the perception of ambiguous patterns representing solid bodies in two different perspectives. In M. Barbi & V. Passarelli (Eds.), *Proceedings of 5th Congress of National Group of Cybernetics & Biophysics* (pp. 285-292). Pisa, Italy: GNCCB.
- BUFFART, H., & LEEUWENBERG, E. (1983). Structural information theory. In H. G. Geisler, H. Buffart, E. Leeuwenberg, & V. Sarris (Eds.), *Modern issues in perception* (pp. 48-72). Amsterdam: North-Holland.
- BUFFART, H., LEEUWENBERG, E., & RESTLE, F. (1983). Analysis of ambiguity in visual pattern perception. *Journal of Experimental Psychology: Human Perception & Performance*, *9*, 980-1000.

- BURIGANA, L., & LUCCA, A. (1985). *La teoria dell'informazione strutturale: Una ricostruzione critica* (Report No. 114). Padua, Italy: University of Padua, Institute of Psychology.
- CARLINI, F., & MASULLI, F. (1983). R.E.D., un sistema per la raccolta e la elaborazione dati di esperienze di psicofisica. In F. Conti & S. Vallergera (Eds.), *Proceedings of 7th Congress of National Group of Cybernetics & Biophysics* (pp. 142-144). Camogli, Italy: GNCB.
- DE MARCO, A., PENENGO, P., TRABUCCO, A., BORSELLINO, A., CARLINI, F., RIANI, M., & TUCCIO, M. T. (1977). Stochastic models and fluctuations in reversal time of ambiguous figures. *Perception*, **6**, 645-656.
- FISHER, G. H. (1967). Preparation of ambiguous stimulus materials. *Perception & Psychophysics*, **2**, 421-422.
- GIRGUS, J. J., ROCK, I., & EGATZ, R. (1977). The effect of knowledge of reversibility on the reversibility of ambiguous figures. *Perception & Psychophysics*, **22**, 550-556.
- GREGORY, R. L. (1974). Choosing a paradigm for perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception: Vol. 1. Historical and philosophical roots of perception* (pp. 255-283). New York: Academic Press.
- HATFIELD, G., & EPSTEIN, W. (1985). The status of the minimum principle in the theoretical analysis of visual perception. *Psychological Bulletin*, **97**, 155-186.
- HOCHBERG, J. (1950). Figure-ground reversal as a function of visual satiation. *Journal of Experimental Psychology*, **40**, 682-686.
- HOCHBERG, J., & BROOKS, W. (1960). The psychophysics of form: Reversible perspective drawings of spatial objects. *American Journal of Psychology*, **73**, 337-354.
- HOCHBERG, J., & MCALISTER, E. (1953). A quantitative approach to figural "goodness." *Journal of Experimental Psychology*, **46**, 361-364.
- HOFFMAN, D. D. (1983). The interpretation of visual illusions. *Scientific American*, **249**, 154-162.
- HOFFMAN, D. D., & RICHARDS, W. A. (1984). Parts of recognition. *Cognition*, **18**, 65-96.
- HOWARD, I. P. (1961). An investigation of a satiation process in reversible perspective of revolving skeletal shapes. *Quarterly Journal of Experimental Psychology*, **13**, 19-33.
- JAMES, F., & ROSS, M. (1977). *Minuit: A system for function minimization and analysis of the parameter errors and correlations* (Internal Rep. D 506). Geneva: CERN.
- KAWAMOTO, A. H., & ANDERSON, J. A. (1985). A neural network model of multistable perception. *Acta Psychologica*, **59**, 35-65.
- KIRKPATRICK, S., GELATT, D. D., & VECCHI, M. P. (1983). Optimization by simulated annealing. *Science*, **220**, 671-680.
- KOHLER, W., & WALLACH, H. (1944). Figural aftereffects: An investigation of visual processes. *Proceedings of the American Philosophical Society*, **88**, 269-357.
- KÜNNAPAS, T. (1969). Figural reversal rate and personal tempo. *Scandinavian Journal of Psychology*, **10**, 27-32.
- LEEUWENBERG, E. (1971). A perceptual coding language for visual and auditory patterns. *American Journal of Psychology*, **84**, 307-349.
- LEEUWENBERG, E., & BUFFART, H. (1983). An outline of coding theory. Summary of some related experiments. In H. G. Geisler, H. Buffart, E. Leeuwenberg, & V. Sarris (Eds.), *Modern issues in perception* (pp. 25-47). Amsterdam: North-Holland.
- LONG, J. M., TOPPINO, T. C., & KOSTENBAUDER, J. F. (1983). As the cube turns: Evidence for two processes in the perception of a dynamic reversible figure. *Perception & Psychophysics*, **34**, 29-38.
- MARTINOLI, O., MASULLI, F., & RIANI, M. (1988). Algorithmic information of image. In V. Cantoni, V. Di Gesù, & S. Levisaldi (Eds.), *Image analysis and processing II* (pp. 287-293). New York: Plenum.
- MUSATTI, C. L. (1924). Sui fenomeni stereocinetici. *Archivio Italiano di Psicologia*, **3**, 105-120.
- ORBACH, J., EHRLICH, D., & HEATH, H. A. (1963). Reversibility of the Necker cube: I. An examination of the concept of "satiation of orientation." *Perceptual & Motor Skills*, **17**, 439-458.
- PHILLIPSON, O. T., & HARRIS, J. P. (1984). Effects of chlorpromazine and promazine on the perception of some multistable visual figures. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **36A**, 291-308.
- RADILOVÁ, J., RADIL, T., & HAVRANEK, T. (1984). Sequential dependences of perceptual interpretations of a repetitively illuminated reversible figure. *International Journal of Psychophysiology*, **2**, 45-50.
- RADILOVÁ, J., & RADIL-WEISS, T. (1984). Subjective figures reversal in two- and three-dimensional perceptual space. *International Journal of Psychophysiology*, **2**, 59-62.
- RADILOVÁ, J., RIANI, M., TUCCIO, M. T., RADIL, T., & BORSELLINO, A. (1983). Reversible figures perceived in two- and three-dimensional subjective space. *Studia Psychologica*, **25**, 76-79.
- REISBERG, D., & O'SHAUGHNESSY, M. (1984). Diverting subjects' concentration slows figural reversals. *Perception*, **13**, 461-468.
- RIANI, M., & MASULLI, F. (1983). Informazione strutturale e ambiguità percettiva. In F. Conti & S. Vallergera (Eds.), *Proceedings of 7th Congress of National Group of Cybernetics & Biophysics* (pp. 150-152). Camogli, Italy: SIBPA-GNCB.
- RIANI, M., OLIVA, G. A., SELIS, G., CIURLO, G., & ROSSI, P. (1984). Effect of luminance on perceptual alternation of ambiguous patterns. *Perceptual & Motor Skills*, **58**, 267-274.
- RIANI, M., TUCCIO, M. T., BORSELLINO, A., RADILOVÁ, J., & RADIL, T. (1986). Perceptual ambiguity and stability of reversible figures. *Perceptual & Motor Skills*, **63**, 191-205.
- TOPPINO, T. C., & LONG, G. M. (1987). Selective adaptation with reversible figures: Don't change that channel. *Perception & Psychophysics*, **42**, 37-48.
- VAN DER HELM, P. A., & LEEUWENBERG, E. (1986). Avoiding explosive search in automatic selection of simplest pattern codes. *Pattern Recognition*, **19**, 181-191.
- VICKERS, D. (1972). A cyclic decision model of perceptual alternation. *Perception*, **1**, 31-48.
- VON GRÜNAU, M. W., WIGGIN, S., & REED, M. (1984). The local character of perspective organization. *Perception & Psychophysics*, **35**, 319-324.
- WASHBURN, M. F., MALLAY, H., & NAYLOR, A. (1931). The influence of the size of an outline cube on the fluctuations of its perspective. *American Journal of Psychology*, **43**, 484-489.
- WILTON, R. N. (1985). The recency effect in the perception of ambiguous figures. *Perception*, **14**, 53-61.

## APPENDIX

### Estimate of Percept Complexity Using the Structural Information Theory

This Appendix contains a short outline of the structural information theory and describes our application of this theory to estimate the complexities of the alternative interpretations of the ambiguous patterns used in the experiment.

The estimate of the complexity of a pattern using the structural information theory begins with a linguistic description (called *primitive code*) of the pattern and continues with the search of the most synthetic pattern code, since the degree of complexity is associated with the length of such a code.

The pattern-coding procedure starts from an assigned point and records all lines and angles encountered while following the pattern contours in an assigned direction. The lines (denoted by the letters l to z), the angles (denoted by the letters a to h), and the so-called *grain* that represents a unitary line (denoted by  $\delta$ ) are the primitives of the structural information theory.

In the primitive code, some semantic operators can be used; the principal ones represent (1) the projection in the three-dimensional perceptual space  $\wedge$ , (2) the negative angles  $-$ , (3) the invisible lengths  $\parallel$ , and (4) the continuation operator  $@$ ; (X), where X is the iterated substructure.

The string of symbols associated with the pattern in this way is called primitive code. In addition, a set of syntactic operators is applied recursively to the pattern code to detect its regularities. A basic set of syntactic operators of the structural information theory is given in Table A1.

**Table A1**  
**A Basic Set of Syntactic Operators**

Syntactic Operator	Code	Reduced Code
Iteration	ab ab ab	3*(ab)
Symmetry	a bc d bc a a bc bc a	S[(a)(bc),(d)] S'[(a)(bc)]
Alternation	ab acd ae aa ca dea	<(a)>\<(b)(cd)(e)> <(a)(c)(de)>\<(a)>
Reflection	abcd	#(dcba)

Note—In the estimate of code complexity, iteration and symmetry operators contain one information load. Reflection seems not to compress the code, but when it works together with a reference, it can strongly reduce the code. A reference is a symbol indicating identical parts of a code. In the code evaluation, a reference has a complexity equal to the full load of the replaced part only for its first occurrence, and equal to 1 for each subsequent occurrence.

At each recursion step, the code length is reduced. After a finite number of steps, the code can no longer be compressed by the operators. The noncompressible code is called *end code* for the actual derivation, while the other codes of the derivation are called *central codes*.

The complexity of a code is the sum of all information loads of the code constituted by the symbols representing the primitives and by the multiplicity and symmetry operators (Buffart & Leeuwenberg, 1983). The complexity of the interpretation of the pattern is the complexity of the end code, with the least informative load of all possible ones generated by means of different derivations from the same primitive code.

As an example of the application of the structural information theory, we code a simple geometrical figure, such as a trapezoid (Figure A1a). Starting from the corner marked with a circle, and following the contours in a counterclockwise direction, we encounter the lines and angles recorded in the following primitive code:

$$l_1 a_1 l_2 l_2 a_1 l_1 a_2 @;(\delta)$$

where @;(\delta) is the continuation of the grain, ending when it encounters the circle. To compress the code, we use the symmetry operator to obtain

$$S (l_1 a_1 l_2) a_2 @;(\delta),$$

where the apostrophe under the symbol of a code means that this symbol contains one information load.

This code is incompressible, and so it is the end code of this derivation; its complexity is 5. Note that @;(\delta) has no informative load (Leeuwenberg, 1971).

We used the structural information theory to estimate the complexities of the alternative interpretations of the stimuli included in the set used in our experiment. We independently coded the two opposite interpretations of the patterns of this set following an explicit top-down approach. First, an interpretation was represented by what we call *iconic code* or *macrocode*, which detects the disposition regularities of the substructures denoted directly by icons. Second, each of these icons (or substructures) was coded using simpler icons, and so on. This iterative process ended when the primitive codes, containing only the operators and the primitives of all substructures, were obtained. At this point, we estimated the complexities of all primitive codes and obtained from them the complexity of the whole interpretation.

In the basic stimulus, Percept A (pyramid) was described as a structure consisting of the square top of a roof, from which

four trapezoidal sides were projected into the three-dimensional perceptual space in the opposite direction with respect to the observer.

Percept B (room) was made up of a square front wall, from which an orthogonal succession of rectangles (the side walls, the ceiling, and the floor of the room) emerged in the three-dimensional perceptual space, toward the observer.

In the following, we estimate, as an example, the complexities of the interpretations of Stimulus 3 in Figure 1. Percept B of this stimulus is described using the iconic code shown in Figure A1c.

The first icon is a square structure (the front wall) containing a rectangular door with a circular doorhandle. The symbol b represents the right angle from which the side walls, the ceiling, and the floor emerge in the three-dimensional space.

The macrocode of the front wall is shown in Figure A1d. Note that, in the iconic codes, we use the semantic operator +, which expresses the coplanarity of the substructures. The square contained in the macrocode has a complexity of 2, like all regular polygons. The door has a macrocode containing a rectangular frame and a circle representing the doorhandle. The rectangular frame has central codes (Figure A1b):

$$@;(n_1 n_3 e n_3 n_2 n_3 e n_3) e \text{ } ^{\prime}n_3 -e @;(n_1 e n_2 e)$$

$$@;(<n_1 e n_2 e>\<n_3>) e \text{ } ^{\prime}n_3 -e @;(n_1 e n_2 e).$$

To reduce this code, let us define the following reference:

$$X = @;(n_1 e n_2 e) = @; (<n_1 n_2>\<e>).$$

By using this reference, we can rewrite the previous code as follows:

$$< X >\< n_3 > e \text{ } ^{\prime}n_3 -e X$$

Then, the end code is

$$< X \text{ } ^{\prime}X >\< n_3 > < e >\< n_3 X >$$

and the complexity is 7.

A circle (the doorhandle) has a primitive code analogous to those of other regular polygons [@(f,u\*\delta)], where the side of the polygon is u times the grain, and the circle has a complexity equal to 2 (Leeuwenberg, 1971). In conclusion, the front wall has a complexity equal to 11.

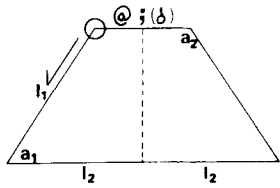
The second substructure is a rectangle that represents both the ceiling and the floor, with a complexity equal to 3.

The last substructure represents the side walls, which are rectangles with a central window. It is decomposed into the iconic code shown in Figure A1e.

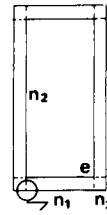
The window frame has a complexity of 7, and the cross of the windowpanes has a complexity of 5. The lateral wall has a complexity equal to 13, and the complexity of the whole Percept B is 28.

For the opposite interpretation (i.e., Percept A, pyramid), the iconic code is shown in Figure A1f. It is the roof with a door on its top, and two trapezoidal windows of the attic on two opposite sides. The roof sides are inclined at an acute angle, c, to the top of the roof.

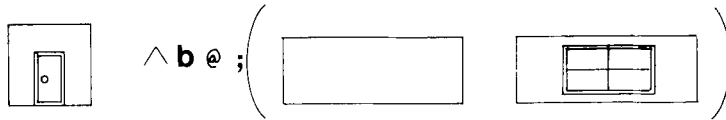
From left to right, the substructures are (1) the square structure made up of the top of the roof that, from a structural point of view, is equal to the corresponding square structure of Percept B (i.e., the front wall), and so it has the same complexity—this is a property of all stimuli of our set; and (2) the trapezoidal structures of the roof sides that are more irregular and there-



(a)



(b)



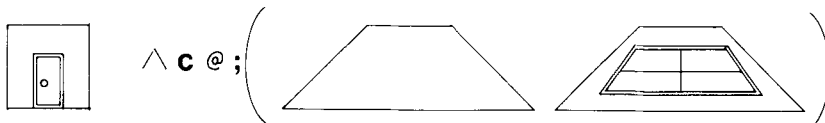
(c)



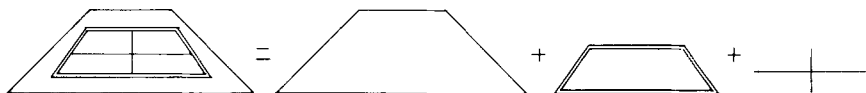
(d)



(e)



(f)



(g)

Figure A1. Some geometric structures and iconic codes analyzed in the Appendix. The circle placed on a vertex, in diagrams a and b, represents the starting point of the coding procedure; the arrow indicates the coding direction; and the other symbols stand for the angles and lines used in the codes. The iconic codes related to Stimulus 3 in Figure 1 are shown here in diagrams c, d, and e for Percept B, "room," and in diagrams f, and g for Percept A, "pyramid."

fore more complex than the corresponding walls, floor, and ceiling of the alternative interpretation.

The two roof sides without windows are trapezoids with complexities of 5, which has already been computed. The sides with the windows are macrocoded, as shown in Figure A1g. The trapezoid of this macrocode has a complexity of 1, because it is a reference to the structure already present in the super-

structure; the cross of the windowpanes has a complexity equal to 5, while the frame (i.e., the medium icon) consists of two dissimilar trapezoids with complexities of 13. In conclusion, the total complexity of Percept *A* is 36.

(Manuscript received November 30, 1987;  
revision accepted for publication September 21, 1988.)