

# What has happened to Prägnanz? Coding, stability, or resonance

CEES VAN LEEUWEN

University of Amsterdam, Amsterdam, The Netherlands

and

MAUD VAN DEN HOF

University of Nijmegen, Nijmegen, The Netherlands

Three theoretical measures of Prägnanz were compared with four data sets. The theoretical measures were a stimulus-coding one (structural information load, SIL), a measure related to within memory processes (stability), and one based on the interaction of perception and memory (resonance). The four data sets were obtained in two experiments and involved goodness rating, grouping, and immediate and delayed recall. A complete set of seven-element binary serial patterns was used in each experiment. Both SIL and resonance were shown to correlate reliably with the data sets across tasks. The resonance measure, however, performed best. Prägnanz thus appears to be explained better by resonance than by stimulus coding or memory storage. Resonance explained all systematic variance in the recall tasks, but not in the other tasks. Regarding these, partial-correlation analyses showed that the effect of stability could be fully reduced to resonance. SIL could not be similarly reduced. Therefore, additional perceptual constraints, other than resonance, would be needed for a complete account of goodness in the judging or grouping tasks.

Simply asking subjects for the "goodness" (Garner, 1974) of a serial pattern provides the experimenter with quite consistent answers. A good series appears to be one with repeating (Restle, 1970) or symmetrically arranged substructures (Jones, 1975). Serial-pattern goodness, therefore, is likely also to manifest itself in intersubjective agreement on how the series are divided into groups (Handel & Todd, 1981).

Goodness is not necessarily identical to simplicity of a perceived structure (Hamada & Ishihara, 1988). A perceptually good stimulus could possess a regular, yet complex, structure. But a form of goodness seems also to occur in memory tasks, where "good" patterns are ones that are more easy to remember (Restle, 1970). Here, simplicity of structure may play a more important role. Unlike perceptual goodness, memory goodness may in addition be sensitive to meaningfulness, frequency of occurrence, or other factors related to the history of the perceiver.

Throughout his entire scientific career, Köhler, (1925, 1950) has maintained that there is only one generic account for what is both perceptually good and easy to remember. This generic account was called the *principle of Prägnanz* (Koffka, 1935), or the *minimum principle*

(Hatfield & Epstein, 1985; Hochberg & McAlister, 1953). Köhler sought its explanation at the neurophysiological level, in terms of an isomorphism between the mental representation structure underlying both perception and memory and the structure of an electrical potential field in the brain. Van Leeuwen (1990a, 1990b) argued that, rather than at a neurophysiological level, the principle of Prägnanz should be specified in a theory of mental representation. Not neurological evidence, but preference data could be used to evaluate such an account. If so, a numerical measure of Prägnanz would be needed to order different representational states according to preference. Correlations with observed preferences could then be used to evaluate the theory.

Such a measure was specified as *structural information load* (SIL) of expressions within a representation system for visual patterns (Leeuwenberg, 1971; Pomerantz & Kubovy, 1986) called *structural information theory* (SIT). It was introduced as a stimulus-coding language. Its coding rules have been revised by their authors more than once. The version to be discussed is the most recent one, which accords mostly to Buffart's work (see van Leeuwen, Buffart, & van der Vegt, 1988).

The corresponding measure, SIL, was introduced in 1971 as a count of the number of symbols needed for the shortest possible coding of a stimulus. Although SIL is a stimulus-coding measure, descriptions may sometimes go beyond the stimulus actually given. But this is restricted to amodal completion (Kanizsa, 1970), a phenomenon characteristic of perception. Expressions, for instance, were used to describe several alternative pattern completions, rather than the pattern as actually given (e.g.,

---

Experiment 1 of the present study was performed by the second author at the S.W.I., University of Groningen, in preparation for her master's thesis. Experiment 2 was performed by Irene Krediet and Rob van de Lubbe, whose contribution is gratefully acknowledged. Requests for reprints should be addressed to Cees van Leeuwen, Department of Psychonomy, Faculty of Psychology, University of Amsterdam, Roeterstraat 15, 1018 WB Amsterdam, The Netherlands (e-mail: ceesvl@uvapsy.psy.uva.nl).

Buffart, Leeuwenberg, & Restle, 1981). The emphasis of the measure is thus on perceptual, rather than on memory, goodness.

Leeuwenberg's approach to reduce Prägnanz to a single factor, SIL, represents a break with the traditional approach of perceptual goodness. The latter specifies goodness in terms of a combination (e.g., a weighted sum) of otherwise unrelated characteristics or features of a stimulus, such as symmetry and proximity of the elements of a pattern (Palmer, 1977), or uncertainty indicators (Vitz & Todd, 1967). Such an approach attempts to account for Prägnanz by eliminating it entirely from the theoretical vocabulary.<sup>1</sup> In this approach, parameters are needed to specify the relative weight among the various relevant characteristics. Fits are obtained with parameters adjusted to the data. It might therefore come as no surprise that these measures correlate between .80 and .90 with data (Vitz & Todd, 1967). Because a unitary measure like SIL is obtained without parameter fitting, one might expect lower correlations. However, in several studies using SIL, correlations were reported between .85 and .95.

This might raise skepticism, because the height of these correlations requires both the measure and the data to be almost error-free. Even if we grant this for the former, it is not very likely for the latter, given the necessary unconstrained character of perceptual organization tasks. Tasks that are too constraining yield trivial results; if the instruction tells a perceiver to rate series for symmetry, ratings will, of course, highly correlate with measures of symmetry, irrespective of whether it captures the essence of Prägnanz. If, on the other hand, the instruction tells the perceiver to rate for goodness, it is left open to what extent the symmetry of the patterns will contribute to the ratings. Intuitively, the contribution is expected to vary from subject to subject and from moment to moment. These circumstances may set an upper limit to the performance of any unitary, parameter-free measure. If, therefore, such high correlations are reported, they are likely to be called *inflated*, either because the measure has been adapted to the stimulus set or the stimulus set has been adapted to the measure. In follow-up research, correlations are then likely to fall back on lower values. If such results are unreported because the weaker, but still significant, correlations are viewed as merely bad experimenting, this leaves for the measure an unwarranted impression of reliability. In the face of such problems, it must be agreed that correlations are not sufficient to establish any measure (Restle, 1970).

Yet, a failure to obtain a correlation can be evidence against a measure. This is why we have chosen to present another correlation study; it could be read as an attempt to eliminate certain alternative measures proposed within the tradition of SIT. This representation system will be used in order to allow a comparison of different measures to be relevant for the alternative processing assumptions on which these measures are based. Such a comparison would be meaningless if the representation system would vary across measures. We will avoid some of the sins of

the earlier mentioned correlation studies. First, to avoid commitment to one specific model and the corresponding inclination to adapt the model to the data, several measures will be compared. Second, among these will be a baseline model based on common sense, which will be used to find out whether the measures perform not only better than chance, but also better than common sense. Third, we will avoid selection of stimuli in adaptation to the model by using a full variation of stimuli of a certain type. Fourth, variation of tasks is introduced to investigate to what extent the proposed measures capture perceptual and memory aspects of goodness.

### Structural Information Theory

SIL is the first measure to be discussed. As mentioned, it is based on the expressions of Leeuwenberg's SIT. These describe identity relations between elements of a pattern. They do so by means of operators. There are three types of operators: iteration, symmetry, and concatenation. Formulas 1a-5b show examples of expressions and their evaluations. The operator is on the left side of the " $\rightarrow$ "; the evaluation is on the right side. The operator *iteration* expresses the identity of adjacent elements. The expression " $3*a$ ," for instance, specifies the identity of the three *as* in the sequence *a a a*. Formula 1a shows a simple example. Formula 1b shows an iteration of a group of elements, called a *substructure*. Substructures in an expression are treated as units and are denoted by parentheses. A variant of iteration is *continuation*, which indicates an indefinite amount of repetitions. Continuation is indicated with the symbol " $\langle \dots \rangle$ ." An example is given in Formula 1c.

The operator *symmetry* expresses the identity of a substructure with the same substructure reversed, as shown in Formulas 2a-2d. There are two symmetry operators, the *S* operator shown in Formulas 2a-2b and the *S'* operator shown in Formulas 2c-2d. The evaluation of the *S'* operator contains the last argument only once as a center, whereas the evaluation of the *S* operator does not have such a center. Notice the difference in evaluation between Formulas 2a and 2b and between Formulas 2c and 2d due to the parentheses. A variant of the symmetry operator is *reversal*, indicated by the symbol # in Formulas 2e and 2f.

The operator *concatenation* connects two or more structures following each other. Simple examples are shown in Formulas 3a-3b. Usually, the backslashes and parentheses of the simple concatenations are omitted. The relevance of the concatenation operator lies in its distributive variant. It allows the description of the identity of nonneighboring elements or substructures, as shown in Formulas 4a-4e. The rule for the distributive iteration prescribes that elements alternate from the leftmost square brackets and from the rightmost ones; this is repeated until the last symbol from the left follows the last symbol from the right. This may lead to quite expanded evaluations, especially if there is more than one element in both the leftmost and the rightmost brackets, as in Formula 4e-4f. The iteration too has a distributive variant. This is shown

in Formulas 5a and 5b, where the number of iterations in the first argument is distributed over the substructures in the second argument.

As shown in Formulas 6-9, arguments of an operator may also be expressions containing operators. This nesting of operators in the language is in accordance with the assumption underlying the code that representations have a hierarchical character. Nested expressions can be evaluated stepwise using the evaluation rules for iteration, symmetry, and concatenation, until the resulting series contains no more operators. A series of variables only is called a *pattern*. Patterns are related to a stimulus by a so-called "semantic mapping." For instance, the elements of a pattern may be mapped one to one onto the elements of a sequence of letters, of tones, or of colored dots, or even onto the elements of a contour of a visual stimulus.

- (1a)  $3*a \rightarrow a a a$
- (1b)  $3*(a b) \rightarrow a b a b a b$
- (1c)  $\langle\langle a b \rangle\rangle \rightarrow \dots a b a b a b a b a \dots$
- (2a)  $S[a b c] \rightarrow a b c c b a$
- (2b)  $S[(a b) c] \rightarrow a b c c a b$
- (2c)  $S'[a b c] \rightarrow a b c b a$
- (2d)  $S'[(a b) (c d)] \rightarrow a b c d a b$
- (2e)  $\#[a b c] \rightarrow c b a$
- (2f)  $\#[a (b c)] \rightarrow b c a$
- (3a)  $a \backslash b \backslash c \rightarrow a b c$
- (3b)  $(a b) \backslash (c d) \rightarrow a b c d$
- (4a)  $\langle a \rangle \backslash \langle b c d \rangle \rightarrow a b a c a d$
- (4b)  $\langle a \rangle \backslash \langle b (d e) \rangle \rightarrow a b a d e$
- (4c)  $\langle a b \rangle \backslash \langle c \rangle \rightarrow a c b c$
- (4d)  $\langle (a b) c (d e) \rangle \backslash \langle (f g) \rangle \rightarrow a b f g c f g d e f g$
- (4e)  $\langle a b \rangle \backslash \langle c d e f \rangle \rightarrow a c b d a e b f$
- (4f)  $\langle a b \rangle \backslash \langle c d e \rangle \rightarrow a c b d a e b c a d b e$
- (5a)  $\langle 2 \rangle * \langle a b \rangle \rightarrow 2*a 2*b \rightarrow a a b b$
- (5b)  $\langle 2 \rangle * \langle (a b) c \rangle \rightarrow 2*(a b) 2*c \rightarrow a b a b c c$
- (6)  $\langle\langle a \rangle \backslash \langle b c \rangle\rangle \backslash \langle p \rangle \rightarrow \langle a b a c \rangle \backslash \langle p \rangle \rightarrow$   
 $a p b p a p c p$
- (7)  $S'[\langle (a b) \rangle \backslash \langle c d \rangle (e f)] \rightarrow$   
 $S'[a b c a b d (e f)] \rightarrow$   
 $a b c a b d e f d b a c b a$
- (8)  $2*(S[a b] c) \rightarrow 2*(a b b a c) \rightarrow$   
 $a b b a c a b b a c$
- (9)  $\langle\langle (a) \rangle \backslash \langle b c \rangle\rangle \backslash \langle (2*a) (3*c) \rangle \rightarrow$   
 $\langle (a b a c) \rangle \backslash \langle (a a) (c c c) \rangle \rightarrow$   
 $a b a c a a b a c c c c$

The evaluation may map an argument of an operator onto several variables in the pattern. Only if this is the case is it implied that an identity relation between these elements is being expressed. Because of the hierarchical character of the expressions of the code, there are certain restrictions on the identity relations between elements of patterns that could be expressed. For instance, there

is no expression for Pattern 10 that is able to describe the identity of all its *as* and *bs*. Expression i, for instance, can be evaluated into Pattern 10. The S operator in Expression i describes the leftmost two *as* and the *bs* of Pattern 10 as identical, but not the rightmost two *as*. This illustrates a consequence of the hierarchical character of the representations (viz., that they cannot describe in every case the identity relations of a pattern fully). Complementary expressions are needed for doing so. Expression ii, for instance, describes as identical the rightmost two *as* in Pattern 10 using the distributive iteration operator, but not the leftmost two.

The choice of an operator description for the coding language makes it clear that representations are hierarchical, but not its consequence (i.e., the need for complementary descriptions). The latter, however, is achieved by using the technique of abstract evaluation. For this technique, Expression iii is obtained from Expression i by renaming the variables, so that identical ones no longer occur at different positions. Expression iii is equivalent to Expression i, because both express exactly the same identity relations. Similarly, Expression iv is equivalent to Expression ii. The evaluations of the newly obtained Expressions iii and iv are Patterns 11 and 12. Identical variables in these patterns refer to the identity relations expressed in Expressions iii and iv or, equivalently, in the original Expressions i and ii, respectively. For this reason, Patterns 11 and 12 are called *abstract evaluations* of Expressions i and ii.

- (i)  $S[a b] a$
- (ii)  $a \langle 2 \rangle * \langle b a \rangle$
- (10)  $a b b a a$
- (iii)  $S[a b] c$
- (iv)  $a \langle 2 \rangle * \langle b c \rangle$
- (11)  $a b b a c$
- (12)  $a b b c c$

The existence of complementary expressions constitutes the kernel of the theory. It is assumed that the stimulus-coding system selects one of the complementary expressions as the preferred hierarchical interpretation of the pattern. The interpretation is decided on the basis of the minimum principle of the theory. The preferred interpretation is identified as the expression having the least SIL. SIL is the number of degrees of freedom in an expression, which is the sum of the unit elements, the numbers in an iteration (where continuation has no number), and each occurrence of the S or S' operator. Ambiguity results if two complementary expressions have equal SILs (Buffart, Leeuwenberg, & Restle, 1983). Examples of counting SILs are given in Expressions v-viii, showing two expressions for Pattern 13 and two for Pattern 14. Counting symmetry, but not the other operators, might seem arbitrary. Some theoretical arguments for doing so, however, have been given in Buffart and Leeuwenberg (1983). In various experiments, it is shown that the formal language of SIT

in combination with its SIL principle can be used to explain perceptual organization phenomena (see, e.g., Leeuwenberg & Buffart, 1983, for a review).

- (13)  $a b a c d e c d$   
(v)  $\langle S'[a(b)] e \rangle \langle (c d) \rangle$   
SIL = 6
- (vi)  $S'[a(b)] S'[(c d)(e)]$   
SIL = 7
- (14)  $d a a d a b d$   
(vii)  $\langle d \rangle \langle (2*a)(a b d) \rangle$   
SIL = 6
- (viii)  $\langle (d a) \rangle \langle (a b d) \rangle$   
SIL = 5

**Structural Memory**

The use of SIL as a preference measure implies commitment to the untenable view that perceptual preferences are static; it amounts to denying the influence of prior knowledge on perception. As a consequence, it is met with many counterexamples. Rock (1983), for instance, has shown that preferred interpretation depends on context. Hochberg and Peterson (1987) showed that it depends on the perceiver's intention. But the predictive success of SIL in various studies remains to be explained. Buffart (1986, 1987) and van Leeuwen et al. (1988) have therefore proposed a dynamic approach to Prägnanz, in which the influence of history and intention could be accounted for.

This approach, called *structural memory*, retained from Leeuwenberg's operator language the basic assumption that expressions must have a hierarchical character. But instead of using the operator language with all its intricacies, the constraint that expressions should be hierarchical could be expressed in terms of abstract evaluations directly. For instance, the abstract evaluation  $a b a b a b$  could be viewed as a superstructure  $x x x$ , with substructure  $a b$  substituted for each occurrence of the  $x$ :  $(a b)(a b)(a b)$ . By contrast,  $a b a b a b$  could not be viewed as having a superstructure  $x y$  with substructures  $a b a$  and  $b a b$  for  $x$  and  $y$ , respectively. This is because identities are not maintained over different substructures, so the substitutions would yield  $a b a c d c$  (variables renamed in order of appearance). There is no expression  $a b a a$  in the system, for instance, because it cannot be obtained by substitution of  $a b a$  for  $x$  into  $x a$ . Such a substitution would result in  $a b a c$ . Other attempts to obtain  $a b a a$  by substitution would also fail: substitution of  $a b$  for  $x$  in  $x a a$  would result in  $a b c c$ , and substitution of  $a b$  for the  $x$  in  $a b a$  would yield  $a b c a$ .

The absence of  $a b a a$  in the structural memory corresponds to the absence of an equivalent expression in Leeuwenberg's operator language. The expressions  $a b a c$ ,  $a b c c$ , and  $a b c a$  in structural memory are equivalent to the expressions  $\langle a \rangle \langle b a \rangle$  (as can be checked through abstract evaluation),  $a b 2*a$ , and SYMM[ $a(b a)$ ], respectively, in the operator language. With only a few exceptions, there is almost a one-to-one correspondence between the

representations in structural memory and the expressions of Leeuwenberg's language (see van der Vegt, Buffart, & van Leeuwen, 1989). Because of this, a meaningful comparison between measures could be made. Otherwise, a difference in adequacy of measures could be ascribed to the expressions used.

In a series of papers, Buffart and his colleagues (Buffart, 1986, 1987; van der Vegt et al., 1989; van Leeuwen, 1989; van Leeuwen et al., 1988) used this principle to construct a network memory model for their representation system. Abstract evaluations were included in the network as nodes. The links between the nodes are those between superstructure and structure and between substructure and structure. In addition, it was assumed that nodes with an isolated position in the network (with few links to other nodes) are difficult to activate, but once activated, their activation is less likely to be disturbed. Thus, nodes in the model which are isolated are assumed to have *stable* activation functions. Stability here indicates invariance over time as well as resistance of its characteristics over random fluctuations (Thom, 1985). Structural stability, in turn, was taken to represent Prägnanz. Therefore, isolatedness, expressed as (minus) the number of links of a node to other ones in the network, is called the (memory) *stability* measure of Prägnanz.

For illustration, suppose a serial pattern consisting of a blue cross, a red one, another blue cross, and another red one is coded by  $a b a b$ . In the network model, this node, according to the hierarchy principle, has a superstructure  $a a$  and a corresponding substructure  $a b$  (as can be checked by substitution of  $a b$  for each occurrence of the variable  $a$  in the superstructure). Therefore,  $a b a b$  is linked with both  $a a$  and  $a b$ . Another way to compose  $a b a b$  would be to alternate  $a a$  with a copy of itself. Variables are named in order of appearance, so the copy of  $a a$  becomes  $b b$ , yielding  $a b a b$  when the symbols are alternated. This composition again implies that  $a a$  is a neighbor of  $a b a b$ . There are no other superstructures and substructures for  $a b a b$ , according to the network. In other words,  $a b a b$  has no other neighbors, and the stability measure for this node amounts to 2. Note that nodes other than  $a b a b$  are possible as (partial) descriptions of the stimulus configuration of red and blue crosses at issue; the node  $a b a c$ , for instance, also fits the stimulus configuration. The node  $a b a c$  has superstructure  $a b$  and substructure  $a b a$ ; that is,  $a b a c$  could be composed by a substitution of the substructure  $a b a$  for the  $a$  in the superstructure  $a b$  (since variables in the node by convention are specified in increasing alphabetic order, because  $a b$  has already been used in  $a b a$ , the  $b$  in  $a b$  must become a  $c$ , resulting in  $a b a c$ ). Thus,  $a b$  and  $a b a$  are linked with  $a b a c$ . An alternative way to compose  $a b a c$  would be to alternate  $a a$  and  $a b$  (because an  $a$  is already used in the first series, the series  $a b$  must be renamed into  $b c$  before the composition could be made). In sum, therefore,  $a b a c$  is linked with three neighbors— $a b a$ ,  $a a$ , and  $a b$ —resulting in a value of 3 for the stability measure. For comparison, consider a

lesser pattern of crosses than the previous one: a blue cross, a red one, another blue one, and a yellow one. The most isolated node for this pattern is  $a b a c$ .

Although it is easy to determine the value of the stability measure for such short patterns, this quickly becomes extremely laborious for larger patterns, for which the corresponding nodes in the network may have more than 100 links. Our present research will therefore use a computer program (described at length in van der Vegt et al., 1989) to determine the stability measure.

When van der Vegt's model was run, it was shown to have an interesting emergent property that could be used independently to predict Prägnanz. Because of what the authors called the "confirmation" assumption, presentation of a stimulus configuration results in an increase of activation in the nodes that fit the stimulus configuration momentarily presented. The increase in activation could be identified with *resonance* (Duncker, 1945; Gibson, 1979; Shepard, 1984). Averaged over all nodes, it will be stronger, the more nodes fit. The number of fitting nodes could therefore be used as a measure for Prägnanz, on the basis of resonance, rather than stability. For illustration, if  $a b a b$  fits, this automatically implies that  $a b a c$  also fits, but not vice versa. Consider the pattern of a blue cross, a red cross, a blue cross, and a red cross. There are two nodes other than the earlier mentioned  $a b c c$  and  $a b a c$  that fit the stimulus configuration (viz.,  $a b c b$  and  $a b c d$ ). This yields 4 as the number of nodes that fit the stimulus configuration. By contrast, the lesser pattern for which the node  $a b a c$  was fitting has only one other fitting node (viz.,  $a b c d$ ).

Stability and resonance measures are not unrelated. By definition, the most stable node for a stimulus is an element of the set of nodes that resonate to a stimulus. However, the *number* of resonating *nodes* determines the resonance of a stimulus, whereas the *number* of *links* of the most stable one among these determines the stability measure. Statistically speaking, if there is a large amount of resonance (a large set of fitting nodes), it is likely that one among them is very stable (similarly, one is more likely to find the new Einstein among a relatively large class of students). But this does not imply a strict relation; in principle, for two stimuli, one could have higher resonance and the other could be more stable. There is, therefore, no reason a priori why data should correlate better with resonance than with stability.

### The Measures Represent Alternative Accounts of Prägnanz

A measure such as the number of fitting nodes is based on the assumption that goodness is not primarily determined inside of the system, rather it is based on the interface of organism and environment, where the system resonates to an external patterned configuration (Shepard, 1984). By contrast, when the stability measure is used as an account for Prägnanz, the focus is entirely on processes within memory. Stability is a typical characteristic of internal memory processes. Within the assumptions of SIT, by comparing the goodness of fit for the alternative mea-

asures SIL, stability, and resonance, it is possible to test stimulus coding, perception-memory interaction and within-memory processes as alternative accounts of Prägnanz. A fourth measure to be compared has no theoretical motivation at all, but serves as a commonsense alternative. A theoretical approach has to make more sense than just common sense, and, therefore, its predictions will have to be superior to the fourth measure; it is our baseline. The commonsense measure is calculated as follows: Identify recurrent groups of elements, and rewrite the series as a juxtaposition of such groups (a single element counts as a group by itself) in such a way as to minimize the number of groups. For instance,  $a b a b$  becomes  $(a b)(a b)$ ;  $a a a a a a$  becomes  $(a a a)(a a a)(a)$ . Count  $n$  equals the number of groups in a series, as a measure of its Prägnanz.

## EXPERIMENT 1

### Hypotheses

We may assume that the same measures that are relevant in goodness ratings are also potentially important in related tasks. Three tasks were therefore compared in Experiment 1. Together with goodness rating, we introduced a *grouping* task, in which subjects must indicate which groups of elements they perceive in the series, and a *recall* task, in which a briefly displayed series must be recalled immediately after presentation. It was expected that, in all tasks, subjects would prefer patterns for which there was a low SIL, a high stability (i.e., a low minimum number of links for its fitting nodes), and high resonance (many fitting nodes).

### Method

**Subjects.** Fifty-five undergraduate students were paid a small amount of money for participation as subjects in Experiment 1.

**Stimuli.** Series of seven circles were used, which could be either open or filled with an  $x$ . All  $2^7 = 128$  possible series of open and filled circles were included in the stimulus sets. Each series appeared on the center of the screen of an Olivetti 24 personal computer, with a stimulus width of approximately  $5^\circ$  of visual angle for the entire series. There was equal spacing between the elements; the space was approximately equal to half the size of an element. For each of the series, SIL, the value of the stability and resonance measures, and the commonsense measure were calculated. Van der Vegt's model was used for determining the stability and resonance of the series. See the Appendix.

**Procedure.** Three tasks—grouping, recall, and rating—were carried out. Eighteen subjects performed in the order grouping, recall, and rating; 17 subjects performed in the order recall, grouping, and rating. We thus had a balance in practice between the grouping and recall tasks, and the rating task was always carried out by the subjects already familiar with the patterns. In all tasks, the same 128 series were presented in a random order. Another group of 20 subjects performed only the rating task, enabling us to compare the untrained subjects with those familiar with the patterns from the earlier tasks.

In the grouping task, a movable cursor appeared directly above the series. The subjects could move the cursor to the left or the right with the arrow keys on the keyboard. By pressing the slash key, they could insert a separation mark (slash) between two elements of the series at the cursor position. A separation mark could be deleted by pressing the space bar. The subjects could insert as many

Table 1  
Correlations of Measures with Data

Theoretical Measure	Correlation with:			
	Grouping Entropy	Recall Errors	Ratings	
			Practiced	Unpracticed
SIL	.579†	.305*	.629†	.613†
Resonance	.680†	.556†	.720†	.646†
Memory Stability	.118	.167	.247*	.303*
Common Sense	.085	.027	.193	.282*

Note— $N = 128$  (series). \* $p < .01$ . † $p < .0001$ .

separation marks as they wanted and correct them freely. By pressing the return key, making the answer definitive, a trial was completed. The groupings were registered automatically. In the recall task, a beep was followed by a 200-msec presentation of the series. There was no mask. Immediately after the sequence had disappeared, the subjects typed the sequence. Two keys on opposite sides of the keyboard of the computer were used for doing so. One key showed an open circle; the other one showed a filled circle. All typed responses could be deleted before the return key was pressed. After the return key was pressed, the series as recalled was registered, together with the number of errors (wrong elements, max = 7).

In the rating task, the subjects were instructed to rate the orderliness of the series according to a 7-point scale. The subjects typed their ratings with the keyboard of the computer, completing the trial by pressing a return. If a subject did not provide a rating within 10 sec, the text "Hurry up a bit, please" appeared on the screen. The following text was displayed immediately below the series: "1 is very orderly and 7 is very disorderly," in order to remind the subjects of the extremes of their response scale. After the return key was pressed, the rating was registered.

## Results

Mean errors and ratings were calculated for the recall and the rating tasks, respectively. An *entropy* measure was calculated from the grouping responses, using the formula  $\sum(x_i/n)\ln(x_i/n)$ , in which  $x_i$  = the number of subjects who gave a grouping  $i$  to a certain series and  $n$  = the total number of subjects. A higher entropy value indicates less agreement between subjects in the grouping given to a series. Agreement had been used as an indicator of goodness by Handel and Todd (1981). Our entropy measure was intended to give a numerical value of agreement. The recall errors, ratings, and entropy for individual patterns are given in the Appendix. Entropy, errors, and ratings were compared with the theoretical measures SIL, resonance and memory stability (from van der Vegt et al., 1989), and common sense. In accordance with van Leeuwen et al. (1988), the resonance and memory-stability measure were log-transformed. The correlations with the results of all tasks are shown in Table 1. (Since all correlations were in the expected direction, absolute values of the correlations will be given for convenience in all tables.)

Table 2  
Intercorrelations of the Measures

	Resonance	Memory Stability	Common Sense
SIL	.686	.341	.214
Resonance		.399	.039
Memory Stability			.205

As shown in Table 1, Leeuwenberg's SIL correlates significantly with the results in all tasks. The correlation with errors in the recall task, however, is considerably lower than the others. Of the dynamic measures, the resonance measure, but not the memory-stability measure, correlates significantly with the results in all three tasks. Resonance is slightly better than SIL on all tasks, but significantly better only on the recall task. The memory-stability measure correlates significantly only with the goodness ratings (both practiced and unpracticed). In Table 2, however, it is shown that resonance and memory stability are correlated ( $r = .399$ ). A partial-correlation analysis revealed no significant relation between ratings and the memory-stability measure, with resonance partialled out ( $r = -.064$ , for the practiced subjects, and  $.065$ , for the unpracticed subjects). This suggests that the significance of the predictions from the stability measure was due to its correlation with resonance.

Alternatively, we could partial out SIL from the correlation between ratings and memory stability. This results in  $r = .044$ , for the practiced subjects, and  $r = .127$ , for the unpracticed subjects. These correlations are also not significant (although  $.127$  is significant at the  $.1$  level). By the same argument as before, the alternative suggestion would be that the effect of memory stability should be ascribed to SIL. To find out if our data allow a preferred choice among these alternative suggestions, we calculated, first, a partial correlation between ratings and resonance, with SIL partialled out. This correlation ( $r = .510$ , for the practiced subjects, and  $r = .399$ , for the unpracticed subjects) could be compared with correlations between ratings and SIL, with resonance partialled out. These correlations amount to  $r = .268$ , for the practiced subjects, and  $r = .306$ , for the unpracticed subjects. Thus, for the practiced subjects' ratings, the partial correlation with resonance was significantly higher than the one with SIL ( $z = 2.211$ ,  $p < .05$ ). We may therefore conclude that the stability correlations are explained better by resonance than by SIL. Finally, the commonsense measure yielded a significant correlation with the ratings in both practiced and unpracticed rating tasks. However, it was considerably lower than the correlations obtained for SIL and resonance.

## EXPERIMENT 2

In Experiment 1, the resonance measure was shown to be a much better predictor than the memory-stability mea-

**Table 3**  
Correlations of Measures with Data

Theoretical Measure	Correlation with Intermediate-Recall Errors
SIL	.411*
Resonance	.568*
Memory Stability	.174
Common Sense	.080

Note— $N = 128$  (series). \* $p < .0001$ .

sure. This result contrasts with the experiments reported earlier (van Leeuwen & Buffart, 1989; van Leeuwen et al., 1988), where the stability measure had been used successfully. An explanation for the success of the stability measure in these experiments seems to be its correlation with the resonance measure. Yet, it could also be argued that the discrepancy is a consequence of a difference in method between the experiments. Our Experiment 1 used immediate ratings or reproduction with minimal risk of confusion between the patterns, whereas the earlier reported experiments used confusion conditions (several patterns had to be remembered simultaneously). Therefore, it could be argued, the load of internal processing in these experiments was higher than in our Experiment 1, explaining the predominance of the stability measure. Experiment 2 was designed to study the effect of all measures in confusion conditions.

### Method

Sixteen students received course credit for participation as subjects in Experiment 2. These subjects performed only a modified version of the recall task of Experiment 1. Experiment 2 was run on another computer (Macintosh II), and presentation times were longer (1,000 msec). Most importantly, however, instead of the stimulus presented on a trial, the one presented on the *previous* trial was recalled (intermediate recall). This implies that the subject must face the confusion arising from the fact that two patterns must be remembered, the one to be reproduced and the one just shown.

### Results and Discussion

The number of errors in Experiment 2, listed in the Appendix, is larger than that in Experiment 1, despite longer presentation times. This shows that confusion between patterns must indeed have occurred in our procedure. Table 3 shows the correlations obtained. These are essentially the same as the ones for the recall task in Experiment 1, so

the confusion does not modify the correlations. This is contrary to the prediction of our model.

Table 4 shows the intercorrelations between the results obtained in Experiments 1 and 2. Without exception, these correlations are highly significant. This result cannot be ascribed to sameness of subjects across tasks, because three independent groups of subjects were involved: the practiced subjects in Experiment 1, who also performed the rating and grouping task, the unpracticed subjects in Experiment 1, and the subjects of Experiment 2. The less-than-perfect correlation between all the tasks supports the claim made in our introduction that no parameter-free measure could reach a perfect fit over tasks. The lowest correlation was obtained between the recall tasks in Experiments 1 and 2. This seems surprising, because these tasks seem very similar at face value. It is possible that this correlation is low merely because it is an inter-experiment one (different subjects *and* apparatus). However, this does not explain why the other interexperiment correlations are so much higher. One way or another, these recall tasks produce much random variance, which delimits the correlations, perhaps because recall strategies are essentially random processes.

Nevertheless, the fact that highly reliable correlations were obtained with the theoretical measures would imply that the systematic variance is almost fully captured by them. To investigate this, we recalculated the correlations between the data sets with two of the theoretical measures partialled out. If a measure captures all or nearly all of the systematic variance in the data, there must remain zero or near-zero partial correlation between the tasks when the measure is partialled out. As shown in Table 5, ratings and grouping contain systematic sources of variance still not accounted for by resonance SIL. A near-zero value is obtained only for the partial correlation between the two recall tasks ( $r = .044$ ) when resonance is partialled out, which would mean that this measure captures all the systematic variation in the recall tasks. By contrast, no near-zero partial correlation was obtained for SIL. So, if a unitary measure is supposed to account for all the systematic variance in a task, we might claim that this is possible for the recall tasks on the basis of resonance. We might therefore conclude from this analysis, that a unitary account of Prägnanz is not possible in terms of SIL, but seems possible for the recall data in terms of

**Table 4**  
Intercorrelations of Tasks in Experiment 1 and between Experiments 1 and 2

Condition	Ratings of Practiced Subjects	Grouping Entropy	Immediate-Recall Errors	Intermediate-Recall Errors (Experiments 1 and 2)
Ratings (Unp)	.945†	.676†	.482†	.540†
Ratings (Pr)		.742†	.565†	.571†
Entropy			.530†	.522†
Immediate-Recall Errors				.346*

Note— $N = 128$  (series). Unp = unpracticed subjects; Pr = practiced subjects.  
\* $p < .01$ . † $p < .0001$ .

Table 5  
Intercorrelations of Experiment 1 Tasks  
with Intermediate Recall from Experiment 2

Condition	Resonance Partialed Out	SIL Partialed Out
Ratings (Unp)	.276*	.400*
Ratings (Pr)	.286*	.441*
Grouping Entropy	.225*	.382*
Immediate-Recall Errors	.044	.254*

Note— $N = 126$  (Number of series-2). Unp = unpracticed subjects; Pr = practiced subjects. \* $p < .01$ .

resonance. This result may contribute to an identification of Prägnanz with resonance.

### GENERAL DISCUSSION

We investigated the same set of seven-element binary serial patterns in four tasks: rating, grouping, immediate recall, and intermediate recall. Correlations were obtained between the responses and several theoretical measures based on the economy principle. Fair correlations were obtained for Leeuwenberg's SIL measure and resonance, one of the network measures. The correlations obtained for resonance were slightly better, on the average, than those obtained for SIL. SIL performed considerably weaker than did the resonance measure, however, on the recall tasks. The height of the correlations was considered appropriate, given the principled openness of all our tasks to strategies not taken into account. This view was confirmed by our finding that only the partial correlation between two recall tasks reached zero when the effect of the resonance measure was partialled out. Thus, the recall tasks have Prägnanz as their only systematic source of variance, whereas all the other tasks may have other systematic, confounding factors.

It was assumed that if there are two tasks for which zero or near-zero partial correlation remains when a measure is partialled out, that measure captures all the systematic variance in the data. Since such a result was obtained only for resonance and not for SIL, resonance is clearly superior to SIL in the sense of having a domain where Prägnanz can be a unitary source of preference. Therefore, if it is required that a unitary account of Prägnanz be given, resonance is a better candidate than SIL. Thus, it might be concluded that Prägnanz is to be identified with the amount of resonance a stimulus elicits in memory, rather than in terms of internal memory processes (stability or decay) or stimulus coding.

In this conclusion, no preference is given to SIL, although the correlations obtained were lower than the ones reported for SIL in Leeuwenberg's papers. The lowering of correlations was expected, given the experimenter freedoms mentioned in our introduction. The fact that correlations were far from perfect is nothing to worry about, if we view a Prägnanz measure as merely a coarse approximation to the underlying process. The resonance measure, as specified here in terms of the

amount of resonating nodes, is indeed only a first approximation of the resonance expected to occur in actually processing these series by a model or a real subject.

But SIL, in contrast, is not a simplification of an implied process; Leeuwenberg's papers contain hardly a suggestion of what the processes should look like. In any case, the significance of the partial correlations after elimination of SIL in all our studies suggests that a unitary Prägnanz account based on stimulus coding does not exist. If a unitary account makes sense, it will have to start from an interactive notion of Prägnanz—as resonance. It explains all the systematic effects in the recall tasks, the other ones being random (e.g., random-walk search strategies as in Ratcliff, 1978). To account for the goodness ratings and grouping task, more specific constraints or strategies would have to be added to a model. These constraints could well be of a perceptual nature. For grouping, for instance, a possible confounding factor is perceptual "span," or length of a repeating substructure. In grouping a pattern with long substructures (e.g.,  $x\ x\ x\ o\ o\ o$ ), an additional constraint that the long substructures be taken as single groups leaves little choice for grouping, whereas it leaves much more freedom in grouping for  $x\ o\ x\ o\ x\ o\ x$ . Therefore, the amount of agreement in grouping, our entropy measure, should have to be considered relative to the amount of choice one has under this constraint, provided that it could be adequately specified.

Regarding ratings, a possible confounding factor is ambiguity. Garner (1974) has shown that ambiguity is negatively correlated with goodness ratings (Garner, 1974): good patterns have few alternatives (i.e., complementary descriptions). SIL is correlated with ambiguity (van Leeuwen & Buffart, 1989), so this might explain the significant correlations with SIL in these tasks. Structural information theory (Buffart et al., 1983) has described ambiguity in terms of the existence of *complementary* expressions with approximately equal SIL, or stability. Though a high score on the resonance measure means that many different representations respond to the pattern, resonance is not to be confused with ambiguity. The correlation between resonance and ambiguity (in terms of the proportion of the numbers of links [stability] of the two most stable representations) amounts to .37 (.21 for I-load). Ambiguity thus specified, however, did not contribute independently to the correlations.

In accordance with our claim that the emphasis of correlation studies should be critical, there are critical consequences to be drawn regarding our own earlier work. The present study shows that the memory-stability measure could be replaced by a resonance measure. The present results necessitate a major revision of the activation-spreading functions proposed in van Leeuwen et al. (1988) and implemented by van der Vegt et al. (1989). The interactive resonance processes in the model should be given a more important role than the internal stability processes. The resonance measure accords to what was labeled in the model the "confirmation" assumption. Preactivated nodes function as perceptual hypotheses regarding the



structure of the stimulus configuration presented. The preactivation is determined by the perceiver's history. The fact that resonance was only an unconfounded measure in the recall tasks, but not in the other tasks, is in accordance with this interpretation. However, only the representations that fit in the actual stimulus seem to play a role (i.e., the ones that resonate to it). It is, therefore, an interactive, rather than an internal, account of memory that seems to be favored by our data. In contrast, an internal stability measure seems to be irrelevant to Prägnanz.

## REFERENCES

- BUFFART, H. (1986). Gestalt qualities, memory structure, and minimum principles. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities: Symposium in memoriam Hermann Ebbinghaus* (pp. 189-204). Amsterdam: North-Holland.
- BUFFART, H. (1987). Seven minus two and structural operations. In E. Roskam & R. Suck (Eds.), *Progress in mathematical psychology* (Vol. 1, pp. 117-149). Amsterdam: North-Holland.
- BUFFART, H., & LEEUWENBERG, E. (1983). Structural information theory. In H.-G. Geissler, H. F. J. M. Buffart, E. L. J. Leeuwenberg, & V. Sarris (Eds.), *Modern issues in perception* (pp. 48-72). Amsterdam: North-Holland.
- BUFFART, H., LEEUWENBERG, E., & RESTLE, F. (1981). Coding theory of visual pattern completion. *Journal of Experimental Psychology: Human Perception & Performance*, *7*, 241-274.
- BUFFART, H., LEEUWENBERG, E., & RESTLE, F. (1983). Analysis of ambiguity in visual pattern completion. *Journal of Experimental Psychology: Human Perception & Performance*, *9*, 980-1000.
- CHURCHLAND, P. M. (1989). *A neurocomputational perspective*. Cambridge, MA: MIT Press.
- DUNCKER, K. (1945). On problem solving. *Psychological Monographs*, *58* (Whole No. 270).
- GARNER, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.
- GIBSON, J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- HAMADA, J., & ISHIHARA, T. (1988). Complexity and goodness of dot patterns varying in symmetry. *Psychological Research*, *50*, 155-161.
- HANDEL, S., & TODD, P. (1981). Segmentation of sequential patterns. *Journal of Experimental Psychology: Human Perception & Performance*, *7*, 41-55.
- HATFIELD, G. C., & EPSTEIN, W. (1985). The status of the minimum principle in the theoretical analysis of visual perception. *Psychological Bulletin*, *97*, 155-186.
- HOCHBERG, J. E., & MCALISTER, E. (1953). A quantitative approach to figural "goodness." *Journal of Experimental Psychology*, *46*, 361-364.
- HOCHBERG, J. E., & PETERSON, M. A. (1987). Piecemeal organization and cognitive components in object perception: Perceptually coupled responses to moving objects. *Journal of Experimental Psychology: General*, *116*, 370-380.
- JONES, M. R. (1975). Memory and rule structure in the prediction of serial patterns. *Journal of Experimental Psychology: Human Learning & Memory*, *104*, 295-306.
- KANIZSA, G. (1970). Amodale Ergänzung und Erwartungsfehler des Gestalt-psychologen. *Psychologische Forschung*, *33*, 325-344.
- KOFFKA, K. (1935). *Principles of Gestalt psychology*. London: Kegan Paul, Trench, & Trubner.
- KÖHLER, W. (1925). Komplextheorie und Gestalttheorie. Antwort auf G. E. Müllers Schrift Gleichen Namens. *Psychologische Forschung*, *6*, 358-416.
- KÖHLER, W. (1950). Psychology and evolution. *Acta Psychologica*, *7*, 288-297.
- 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: A summary of related experiments. In H.-G. Geissler, H. F. J. M. Buffart, E. L. J. Leeuwenberg, & V. Sarris (Eds.), *Modern issues in perception*. Amsterdam: North-Holland.
- PALMER, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, *9*, 441-474.
- POMERANTZ, J. R., & KUBOVY, M. (1986). Simplicity and likelihood principles. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 2, pp. 36-1-36-46). New York: Wiley.
- RATCLIFF, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59-108.
- RESTLE, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review*, *77*, 481-495.
- ROCK, I. (1983). *The logic of perception*. Cambridge, MA: Bradford Books.
- SHEPARD, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking and dreaming. *Psychological Review*, *91*, 417-447.
- THOM, R. (1985). From the icon to the symbol. In R. E. Innes (Ed.), *Semiotics: An introductory anthology* (pp. 275-291). London: Hutchinson.
- VAN DER VEGT, J., BUFFART, H., & VAN LEEUWEN, C. (1989). The "structural memory": A network model for human perception of serial objects. *Psychological Research*, *50*, 211-222.
- VAN LEEUWEN, C. (1989). PDP and Gestalt: An integration? *Psychological Research*, *50*, 199-201.
- VAN LEEUWEN, C. (1990a). Indeterminacy of the isomorphism heuristic. *Psychological Research*, *52*, 1-4.
- VAN LEEUWEN, C. (1990b). Perceptual learning systems as conservative structures: Is economy an attractor? *Psychological Research*, *52*, 145-152.
- VAN LEEUWEN, C., & BUFFART, H. (1989). Facilitation of retrieval by perceptual structure. *Psychological Research*, *50*, 202-210.
- VAN LEEUWEN, C., BUFFART, H., & VAN DER VEGT, J. (1988). Sequence influence on the organization of meaningless serial stimuli: Economy after all. *Journal of Experimental Psychology: Human Perception & Performance*, *14*, 481-502.
- VITZ, P. C., & TODD, T. C. (1967). A model of learning for simple repeating binary patterns. *Journal of Experimental Psychology*, *75*, 108-117.

## NOTE

1. For the difference between *elimination* and *reduction*, see, for example, Churchland (1989, chap. 3).

## APPENDIX

Pattern	Stab	Res	Expression (after Leeuwenberg)	I	Groups	Rating			Imm Recall Errors	Interm Recall Errors
						Unpr	Pr	Entropy		
aaaaaaa	0	337	<<a>>	1	aaa.aaa.a	1.05	1.2222	0.8581	0.1667	0.6429
aaaaaab	2	110	<<a>>b	2	aaa.aaa.b	2.15	2.1111	1.2874	0.6389	1.6429
aaaaaba	4	99	<<a>>ba	3	aa.aa.a.b.a	2.7	2.8148	1.7018	1.3611	1
aaaaabb	4	72	<<a>>2*b	3	aa.aa.a.b.b	2.7	2.4286	1.1254	1.1944	1.5357
aaaabaa	7	90	<<a>>b2*a	4	aa.aa.b.aa	3.05	2.8571	1.6918	1.0278	1.6429
aaaabab	5	60	<<<a>>a>\<b>	3	a.a.a.ab.ab	4	4.2222	1.6166	1.5278	2.9286
aaaabba	6	67	<<4*a2*b>>	4	a.a.a.ab.ba	3.95	3.8148	1.7888	1.5278	2.0714
aaaabbb	5	65	<<a>>3*b	3	aa.aa.b.b.b	1.8	2	1.1499	1.1944	1
aaabaaa	2	100	<<3*ab>>	3	aaa.b.aaa	1.45	1.6071	1.3621	0.75	1
aaabaab	5	45	a<<2*ab>>	4	a.a.ab.a.ab	3.7	3.4074	2.0071	1.1944	2.5
aaababa	5	59	3*a<<ba>>	4	a.a.ab.ab.a	3.8	4.3333	1.982	1.9722	2.3571
aaababb	7	36	<2>*<(a)(ab)>b	5	a.a.ab.ab.b	5.1	4.8889	2.0231	1.6389	1.2143
aaabbaa	2	63	<<a>>\<2>*<ba>	4	a.a.ab.baa	2.95	3.4074	1.5216	1.2778	3.5
aaabbab	7	39	<<a>>\<ba>\<b>	4	a.a.ab.ba.b	4.8	4.6296	2.0402	1.6389	1.9286
aaabbba	7	58	<<<3>*<ab>>>	3	a.a.ab.b.ba	2.85	3.5556	1.055	1.5833	2
aaabbbb	5	65	<<a>>4*b	3	a.a.a.bb.bb	2.2	2.1852	1.1728	1	1.9286
aabaaaa	6	90	#<<4*ab>>	3	aa.b.aa.aa	3.2	3.037	1.7788	0.4444	1.7857
aabaaab	4	51	#<<b3*a>>	3	aab.a.aab	3.5	3.8214	1.949	1.1944	1.9286
aabaaba	2	45	<<2*ab>>	3	aab.aab.a	3.1	3.4074	1.8344	1.3333	2.7143
aabaabb	5	36	<<2ab>>b	4	aab.aab.b	4.25	4.6296	1.7668	1.0278	2.5714
aababaa	4	49	S'[2*aba]	5	aab.a.baa	1.9	2.6667	1.9846	1.3611	2.1429
aababab	4	44	a<<ab>>	3	a.ab.ab.ab	3.6	3.8889	2.0635	2.1111	2.6429
aababba	7	34	a<ab>\<(ba)>	5	a.ab.ab.ba	4.55	4.6296	2.2076	1.5	3.4286
aababbb	5	36	a<2>*<(ab)b>	5	a.ab.ab.b.b	4.9	4.8571	2.2042	1.5	1.4286
aabbaaa	6	63	<2>*<ab>\<<a>>	4	aab.baa.a	3	2.6296	1.4804	0.6389	2.2857
aabbbaab	4	34	<<<2>*<ab>>>	3	aab.b.aab	3.45	3.4815	1.6432	1.4444	2.8571

## APPENDIX (Continued)

Pattern	Stab	Res	Expression (after Leeuwenberg)	I	Groups	Rating			Imm Recall Errors	Interm Recall Errors
						Unpr	Pr	Entropy		
aabbaba	11	35	<(2*ab)b>\<a>	5	a.ab.ba.ba	4.35	4.6786	2.1357	1.4167	2.8571
aabbabb	5	36	a<<a2*b>>	4	a.abb.abb	3.6	4.0741	1.7538	1.5556	2
aabbbaa	5	55	<<2*a3*b>>	4	aa.b.b.b.aa	1.55	1.7037	1.3946	1.5833	1.4286
aabbbab	2	38	<a(2*b)>\<(ab)>	5	a.ab.b.ba.b	5.1	4.2857	1.8954	1.7778	2.5
aabbbba	4	52	<<2*a4*b>>	4	a.ab.b.b.ba	3.2	3.7037	1.364	1.4722	2.5
aabbbbb	5	72	2*a<<b>>	3	a.a.bb.bb.b	2.6	2.5926	1.0772	1.1944	0.8571
abaaaa	2	99	ab<<a>>	3	a.b.aa.aa.a	2.7	2.6429	1.4993	0.3611	2.2857
abaaaab	2	56	<<ab3*a>>	4	ab.a.a.a.ab	3.75	3.7037	1.6829	1.0278	2.6429
abaaaba	2	53	<<ab2*a>>	4	aba.a.aba	2.05	2.1786	1.8159	1.1667	2
abaaabb	4	38	ab3*a<<b>>	5	ab.a.a.ab.b	4.7	4.25	1.7202	1.4444	2.5714
abaabaa	2	45	a<<b2*a>>	4	a.baa.baa	2.95	3.6429	2.182	1.5556	1.8571
abaabab	2	40	<<ab2*ab>>	5	a.ba.ab.ab	4.55	4.4815	2.095	1.6389	4.2857
abaabba	4	33	<a(ab)>\<(ba)>	5	ab.a.ab.ba	4.45	4.7778	2.1703	1.4167	3
abaabbb	2	39	<a>\<ba>3*b	5	ab.a.ab.b.b	4.8	4.5714	2.0009	1.5556	1.8571
ababaaa	2	59	2*(ab)<<a>>	4	ab.ab.a.a.a	3.95	4.2593	1.4019	1.4167	2.1429
ababaab	2	38	<2>*<(ab)a>b	5	ab.ab.a.ab	4.5	4.8519	2.2361	1.9167	2.7857
abababa	2	69	<<ab>>	2	ab.ab.ab.a	1.45	1.7407	1.4878	1.7778	1.5
abababb	2	44	<<ab>>b	3	ab.ab.ab.b	4.3	4.4074	1.9205	1.9444	2.0714
ababbba	4	35	<<<ab>\<(ba)>>>	4	ab.ab.ba.a	4.5	4.963	2.0576	1.6667	2.4286
ababbab	2	40	<a<<b>>>\<(ba)>	4	ab.ab.ba.b	3.55	4.1481	2.3448	1.6389	2.6429
ababbba	2	45	<<<2>*<(ab)b>>>	4	ab.ab.b.ba	5	5.2963	2.1808	1.7222	2.1429
ababbbb	2	60	<a>\<b(<<b>>)>	3	ab.ab.b.b.b	4	4.4444	1.786	1.0833	1.0714
abbaaaa	2	67	a2*b<<a>>	4	ab.ba.a.a.a	3.4	3.2963	1.5067	0.6667	2.2857
abbaaab	7	37	<<S[ab]a>>	4	ab.ba.a.ab	4.55	4.0741	2.2492	1.1389	2.3571
abbaaba	4	33	<(ab)>\<(ba)a>	5	ab.ba.a.ba	4.1	4.2963	2.095	1.8333	2.2857
abbaabb	4	34	<<S[ab]>>	3	abb.a.abb	2.9	3.6296	2.0037	1.8889	2.1429

## APPENDIX (Continued)

Pattern	Stab	Res	Expression (after Leeuwenberg)	I	Groups	Rating			Imm Recall Errors	Intern Recall Errors
						Unpr	Pr	Entropy		
abbabaa	4	34	<(ab)><b(<<a>>)>	4	ab.ba.ba.a	5.15	4.7037	2.124	2.1111	2.2857
abbabab	2	38	<a<<b>>>><b<<a>>>	4	ab.ba.ba.b	4.55	4.6667	1.927	1.8611	2.2857
abbabba	2	42	<<a2*b>>	3	abb.abb.a	1.75	2.6667	2.0688	1.4444	1.5
abbabbb	6	45	S[ab]<<b>>	4	abb.abb.b	4.25	4	2.0764	1.4444	1.8571
abbbaaa	2	58	<<#<3>*ab>>>	3	ab.b.ba.a.a	3.05	3.2963	1.3154	0.8611	2.0714
abbbaab	6	37	<<S'[abb]>>	4	ab.b.ba.ab	4.75	4.4815	1.9627	2	2.5
abbbaba	2	45	<<<<a><ba>>><b>>>	4	ab.b.ba.ba	5.25	5.037	2.2937	1.4722	3
abbbab	4	51	<<a3*b>>	3	abb.b.abb	4.6	4.3571	2.0185	1.5833	2.7857
abbbaa	4	52	<<#2*a4*b>>	4	abb.bba.a	3.6	3.4815	1.8092	1.25	1.6429
abbbbab	4	56	<<a4*b>>	3	abb.bba.b	4.55	4.1481	1.9022	2.1389	2.4286
abbbbba	2	72	<<a5*b>>	3	abb.b.bba	1.5	1.5926	1.2958	1.1389	1
abbbbb	2	110	a6*b	3	a.bbb.bbb	1.95	2.3929	1.1223	0.8611	0.7143
baaaaa	2	110	#b6*a	3	b.aaa.aaa	1.95	2.1111	1.329	0.6389	2.0714
baaaaab	2	72	#<<b5*a>>	3	baa.a.aab	1.2	1.5926	1.4546	0.4444	1.8571
baaaaba	4	56	#<<b4*a>>	3	baa.aab.a	4.1	3.6667	2.0769	1.3333	2.7143
baaaabb	4	52	<<2*b4*a>>	4	baa.aab.b	4.15	3.7037	1.5972	1.3056	1.8571
baaabaa	4	51	#<<b3*a>>	3	baa.a.baa	3.35	3.6667	1.8948	1.3611	3.2857
baaabab	2	45	#<<<<b><ab>>><a>>>	4	ba.a.ab.ab	4.6	4.3571	2.0693	1.5278	2.8571
baaabba	6	37	#<<S'[baa]>>	4	ba.a.ab.ba	4.4	4.4074	2.1101	1.7222	2.7857
baaabbb	2	58	<<<3>*ba>>>	3	ba.a.ab.b.b	3.7	3.7407	1.3682	1.4167	2
baabaaa	6	45	#S[ba]<<a>>	4	baa.baa.a	3.15	3.5926	1.6373	0.6111	1.9286
baabaab	2	42	#<<b2*a>>	3	baa.baa.b	1.45	2.1481	1.7317	1.0556	1.8571
baababa	2	38	#<b<<a>>>><a<<b>>>	4	ba.ab.ab.a	4.5	4.8148	2.0924	1.8333	2.6429
baababb	4	34	#<(ba)><a(<<b>>)>	4	ba.ab.ab.b	4.55	4.6667	2.0986	1.8333	1.9286
baabbaa	4	34	#<<S[ba]>>	3	baa.b.baa	3.75	3.5926	1.9602	1.5	1.8929
baabbab	4	33	#<(ba)><(ab)b>	5	ba.ab.b.ab	4.05	4.2593	2.0688	1.75	2.6429

## APPENDIX (Continued)

Pattern	Stab	Res	Expression (after Leeuwenberg)	I	Groups	Rating			Imm Recall Errors	Intern Recall Errors
						Unpr	Pr	Entropy		
baabbba	7	37	#<<S[ba]b>>	4	ba.ab.b.ba	4.9	4.3704	2.1611	1.75	2.9286
baabbbb	2	67	#b2*a<<b>>	4	ba.ab.b.b.b	4.3	3.8519	1.6027	1.3333	1.8571
babaaaa	2	60	#<b>\<a(<<a>>)>	3	ba.ba.a.a.a	4.05	3.8889	1.7826	0.7778	1.9286
babaaab	2	45	#<<<2>*(<ba>a)>>	4	ba.ba.a.ab	4.5	4.5926	2.148	1.2778	2.1429
babaaba	2	40	#<b<a>>\<(ab)>	4	ba.ba.ab.a	4.6	4.1111	2.2253	1.6667	2.2143
babaabb	4	35	#<<<ba>\<(ab)>>>	4	ba.ba.ab.b	4.2	4.5	1.9906	1.5556	2.4286
bababaa	2	44	#<<ba>>a	3	ba.ba.ba.a	3.1	4.2593	1.8763	2.1944	1.7857
bababab	2	69	#<<ba>>	2	ba.ba.ba.b	1.65	2.1429	1.8077	2.25	2.2857
bababba	2	38	#<2>*(<ba>b>a	5	ba.ba.b.ba	4.65	4.9259	2.223	1.9444	2.2857
bababbb	2	59	#2*(<ba><<b>>	4	ba.ba.b.b.b	4.75	4.4074	1.8387	2.2778	1.9286
babbaaa	2	39	#<b>\<ab>3*a	5	ba.b.ba.a.a	5	4.963	2.2837	0.9167	2.4286
babbaab	4	33	#<b(<ba>)>\<(ab)>	5	ba.b.ba.ab	3.9	4.3704	2.1038	1.1389	2.5357
babbaba	2	40	#<<ba2*ba>>	5	b.ab.ba.ba	4.4	4.2963	2.0579	1.7222	2.1429
babbabbb	2	45	#b<<a2*b>>	4	b.abb.abb	3.3	4.1852	2.2193	1.4444	2
babbbaa	4	38	#ba3*b<<a>>	5	ba.b.b.ba.a	4.5	4.8889	1.8741	1.4722	2.0714
babbbab	2	53	#<<ba2*b>>	4	bab.b.bab	2.3	2.7407	1.684	1.0556	1.3571
babbbba	2	56	#<<ba3*b>>	4	ba.b.b.b.ba	4.25	4.3333	1.9675	1.6667	1.9286
babbbbbb	2	99	#ba<<b>>	3	b.a.bb.bb.b	3.2	3.037	1.5683	1.0556	1.7143
bbaaaaa	5	72	#2*b<<a>>	3	b.b.aa.aa.a	2.65	2.5556	1.2521	0.7222	1.3571
bbaaaab	4	52	#<<2*b4*a>>	4	b.ba.a.a.ab	3.85	3.6667	1.5608	1.3611	1.9286
bbaaaba	2	38	#<b(2*a)>\<(ba)>	5	b.ba.a.ab.a	4.6	4.3704	1.9344	1.7222	1.8571
bbaaabb	5	55	#<<2*b3*a>>	4	bb.a.a.a.bb	1.6	1.6786	1.4981	1.3333	2.1429
bbaabaa	5	36	#b<<b2*a>>	4	b.baa.baa	4.35	4.7143	1.9898	2.0556	2.7857
bbaabab	11	35	#<(2*ba)>\<b>	5	b.ba.ab.ab	4.65	4.5185	1.8728	1.8333	2.8571
bbaabba	4	34	#<<<2>*(<ba>>>	3	bba.a.bba	3.25	3.6429	1.527	1.9167	2.1429
bbaabbb	6	63	#<2>*(<ba>\<<b>>	4	bba.abb.b	3.2	3.4444	1.4646	1.4444	2.2143

## APPENDIX (Continued)

Pattern	Stab	Res	Expression (after Leeuwenberg)	I	Groups	Rating			Imm Recall Errors	Interm Recall Errors
						Unpr	Pr	Entropy		
bbabaaa	5	36	#b<2>*(ba)a>	5	b.ba.ba.a.a	4.65	4.6071	1.9476	1.25	2
bbabaab	7	34	#b<ba>\<(ab)>	5	b.ba.ba.ab	4.9	5.0741	2.3594	1.7778	2.6429
bbababa	4	44	#b<<ba>>	3	b.ba.ba.ba	4.3	4.8148	2.059	2.1111	3
bbababb	4	49	S'[2*bab]	5	bba.b.abb	1.95	2.9259	2.0493	1.8056	1.2857
bbabbaa	5	36	#<<2ba>>a	4	bba.bba.a	3.85	3.7037	1.4804	1.6389	1.7857
bbabbab	2	45	#<<2*ba>>	3	bba.bba.b	3.8	3.8929	2.1968	1.7222	3.2143
bbabbba	4	51	<<a3*b>>	3	bba.b.bba	3.9	4.1481	1.7508	1.8333	2.2143
bbabbbb	6	90	<<4*ba>>	3	bb.a.bb.bb	3.7	3.037	1.9233	1.3333	1.7143
bbbaaaa	5	65	#<<b>>4*a	3	b.b.b.aa.aa	2.15	2.2222	1.2169	0.5833	1.6429
bbbaaab	7	58	#<<<3>*<ba>>>	3	b.b.ba.a.ab	3.85	3.6296	1.0894	1.2778	2.4286
bbbaaba	7	39	#<<b>>\<ab>\<a>	4	b.b.ba.ab.a	4.95	4.4444	1.6378	1.3333	2.7143
bbbaabb	2	63	#<<b>>\<2>*<ab>	4	b.bba.abb	3.15	3.6667	1.9481	1.5833	2.0714
bbbabaa	7	36	#<2>*(b)(ba)>a	5	b.b.ba.ba.a	5.1	4.5556	2.0069	1.4444	1.7857
bbbabab	5	59	#3*b<<ab>>	4	b.b.ba.ba.b	4.35	4.5185	2.0973	1.8611	2.2857
bbbabba	5	45	#b<<2*ba>>	4	b.b.ba.b.ba	4.15	4.0357	2.1822	1.1944	1.5714
bbbabbb	2	100	#<<<3*ba>>	3	bbb.a.bbb	1.3	1.7037	1.0691	1.2778	1
bbbbaaa	5	65	#<<b>>3*a	3	bb.bb.a.a.a	1.8	2.2143	1.3318	1.1944	0.7143
bbbbaab	6	67	#<<<4*b2*a>>	4	b.b.b.ba.ab	4.15	3.8148	1.661	1.3611	1.3571
bbbbaaba	5	60	#<<<b>>b>\<a>	3	b.b.b.ba.ba	4.3	4.037	1.885	1.75	0.9286
bbbabb	7	90	#<<b>>a2*b	4	bb.bb.a.bb	3.65	3.4074	1.5058	1.4444	2
bbbbaaa	4	72	#<<b>>2*a	3	bb.bb.b.a.a	1.9	2.3333	1.0597	1.1667	1.0714
bbbbbab	4	99	#<<b>>ab	3	bb.bb.b.a.b	2.95	2.8148	1.6215	1.4722	1.2857
bbbbbba	2	110	#<<b>>a	2	bbb.bbb.a	2.25	2.5926	1.2851	1.3056	0.9286
bbbbbbb	0	337	#<<b>>	1	bbb.bbb.b	1	1.4444	1.2179	0.8056	0.5

Note—In the patterns, “a” indicates an open circle and “b” indicates a closed one. Stab = stability; Res = resonance; I = SIL; Unpr = unpracticed subjects; Pr = practiced subjects; imm = immediate; interm = intermediate.