From Fragments to Objects – Segmentation and Grouping in Vision
T.F. Shipley and P.J. Kellman (Editors)

11

Simplicity, Regularity, and Perceptual Interpretations: A Structural Information Approach

Rob Van Lier, NICI, University of Nijmegen, PO Box 9104, 6500 HE Nijmegen The Netherlands

Upon opening our eyes we start interpreting our visual surrounding. Within a fraction of a second, the light falling on the retinae is transformed into an understandable world. This output of the visual system is so convincing that it can easily be confused with the thing that is perceived, but in fact it is nothing more than a reflection of it. Notwithstanding this basic uncertainty, we mostly agree on what we perceive. On the whole it appears that -- although each visual pattern can be interpreted in many different ways -- just one specific interpretation is preferred. In the history of visual science, students of perception have searched for organizational grouping principles that determine how an arbitrary visual pattern is interpreted. The notion of simplicity advocated in this chapter attempts to combine various aspects of perceptual grouping within one framework and is a central issue within the Structural Information Theory (SIT), initiated by Leeuwenberg (1969, 1971) and further developed since then. In this chapter, a brief introduction to the notion of simplicity and SIT's regularity based quantifications is given. Furthermore, applications of SIT to various stimulus domains (series, surfaces, objects) are illustrated.

FROM PRÄGNANZ TO SIMPLICITY

In perception research, consideration of simplicity can be traced back to the early Gestalt psychologists (e.g. Wertheimer, 1912; Köhler, 1920; Koffka, 1935). To predict perceived interpretations of visual patterns, these psychologists formulated various so-called Gestalt laws, such as good continuation, closure, and proximity. It appeared, however, to be virtually impossible to predict which rule applies to what pattern. Moreover, there was no a priori hierarchy between the separate laws: in principle, each law could overrule -- or could be overruled by -- another law. In order to cope with these problems, Koffka (1935) formulated the law of Prägnanz. This law can be regarded as an attempt to integrate the Gestalt laws, and stresses an overall tendency toward "good" forms. Nevertheless, even on the basis of this general law, predictions of perceived interpretations could hardly be made, as no clear specifications were provided. An important step forward was the formulation of the global-minimum principle by Hochberg and McAlister (1953). The global-minimum principle states that for a given pattern its simplest possible interpretation will be perceived. Hatfield and Epstein (1985) made a distinction between three types of simplicity: procedural simplicity, phenomenal simplicity, and descriptive simplicity. Procedural simplicity requires that the preferred interpretation is attained by the most economical process. The notion of phenomenal simplicity endorses the idea that the preferred interpretation contains a maximum of regularity. Finally, the notion of descriptive simplicity states that the description of the preferred interpretation is as compact as possible. The above notions do not exclude each other. Actually, the latter two are more or less integrated in the claim within SIT that the simplest description reflects the interpretation with the highest degree of regularity. Hatfield and Epstein (1985) further remarked that the global-minimum principle is an empty concept if it does not have an operationalization, that is, it needs a measure of simplicity. Indeed, in order to test the global-minimum principle, a quantifiable measure of simplicity is required (cf. Boselie & Leeuwenberg, 1986).

Quantification of simplicity

Formalization of regularity. Several descriptive systems, or coding systems, have been proposed by various scientists (e.g. Simon & Kotovsky, 1963; Vitz, 1966; Vitz & Todd, 1969; Leeuwenberg, 1969, 1971; Restle, 1979, 1982; Jones & Zamostny, 1975; Deutsch & Feroe, 1981). Although not all coding systems were designed for visual patterns, each of them transposes a pattern into a symbolic representation. Within SIT, a visual pattern is represented by a symbol series referring to, for example, lengths and angles in the contour of a visual pattern. Then, by means of coding rules, the initial series can be reduced to a shorter series or code. This reduction proceeds by extracting as much regularity as possible from an initial series. These regularities are based on identities between the elements in the symbol series.

Theoretical considerations (Van der Helm, 1988; Van der Helm & Leeuwenberg, 1991; Van der Helm & Leeuwenberg, 1996) have led to a formalization of the concept of regularity, which, in turn, affected the quantification of simplicity. The starting point of this formalization is the concept of accessibility, stating that appropriate coding rules are characterized by so-called holographic regularity and transparent hierarchy. Holographic regularities are regularities constituted by identity structures in which all substructures express the same kind of regularity. As a consequence, the structure of the regularity is invariant under growth (e.g., a repetition is a holographic regularity because it can be expanded by adding additional elements, preserving its repetition character). The concept of transparent hierarchy holds that different regularities can be combined only if their groupings are hierarchically compatible, i.e., the grouping by one regularity builds on the grouping by the other regularity. These concepts are extensively discussed and clarified in Van der Helm and Leeuwenberg (1991) and Van der Helm, Van Lier, and Leeuwenberg (1992), and Van der Helm (2000). Considering all possible regularities in a pattern, it appears that only a few regularities possess the qualities of holography and transparency (Van der Helm & Leeuwenberg, 1991). These regularities form the kernel of the set of coding rules that have been used since SIT's conception (Leeuwenberg, 1969; 1971), namely Iteration, Symmetry, and Alternation (e.g., the symbol series aaaa can be encoded into 4*(a) by means of the Iteration rule; the series abba into S[(a)(b)] by the Symmetry rule; and the series akal into <(a)>/<(k)(l)> by the Alternation rule). In addition, Van der Helm and Leeuwenberg (1986, 1991) showed that, by means of the concept of accessibility plus a mathematical technique called the shortest path method, the simplest code can be found without generating each and every possible code separately. SIT's basic concepts of holographic regularity and transparent hierarchy provide a coherent approach to the problem of defining regularity. Until the development of the concept of accessibility, the transformational approach, based on mathematical Group theory, was the only formalism available to specify kinds of regularity. The latter approach has been applied to phenomena of perceptual organization by various researchers (cf., Garner, 1974; Palmer, 1991). By that approach, regularity is assessed by groups of transformations that leave a visual pattern invariant. Van der Helm and Leeuwenberg (1996) compared the transformational approach and the formalization of regularity as specified within SIT. The authors concluded that with respect to repetition the two approaches hardly differ from each other, whereas they differ essentially with respect to symmetry. For example, within the transformational approach, mirror symmetry is an all-or-nothing property, whereas within SIT it is a graded property as it is represented in a point-by-point fashion (a difference which can be related to various phenomena such as noise insensitivity in symmetric dot patterns). Furthermore, the authors related the difference between repetition and symmetry to so-called 'objectness', comprising a more dominant role of repetition in the case of multiple objects and a more dominant role of symmetry in the case of a single object (cf. Baylis & Driver, 1995). Another example on which the holographic and the SIT approach differ is the perceived goodness of symmetric patterns. Whereas according to the transformational approaches, a 3-fold symmetry would result in a better pattern than a 2-fold symmetry, the SIT approach does not predict a simple increase in perceived goodness (and would actually predict the opposite in this particular case, see Van der Helm and Leeuwenberg, 1996). Van der Helm and Leeuwenberg (1996) found some scattered evidence that supported their predictions. Since SIT and the transformational approach make diverging predictions about a number of perceptual phenomena, the relative merits of the two theories should soon be evident.

Information load. Before the accessibility criterion was incorporated into SIT, several measures of Structural Information were considered. The measure that was most commonly used was meant to reflect the amount of memory space needed for the storage of a given code. Roughly all parameters and operations in a structural code were counted to determine the amount of Structural Information. Van der Helm et al. (1992) argued, however, that this measure was conceptually troublesome as incomparable entities contributed to the complexity of that code (see also Hatfield & Epstein, 1985). In addition, our view is that the semantic implication of a code (i.e., an interpretation) should not be judged on the basis of the required memory space but, in psychologically more meaningful way, on the basis of the semantic content of the code, i.e., the description of regularity. The accessibility criterion, mentioned above, paved the way to a better-defined information load. This new information load considers all parameters at all hierarchical levels in a structural code. To exemplify such hierarchical codes consider the minimum codes of the symbol series 'ababab' and 'abcab', being '3*(ab)' and 'S[(ab),(c)]', respectively. The information load of the first code equals 3: 'a' and 'b', and the group 'ab', which is considered as a parameter at a higher code level. The second code yields an information load of 4: 'a', 'b', 'c', and the group 'ab' (by means of contrast, the latter encoding can be compared with the encoding of the series 'abcba' into S[(a)(b),(c)], with the parameters 'a', 'b', and 'c', having an information load of 3). An algorithm has been developed that can achieve the code with the lowest information load, given an initial series (PISA; Van der Helm, 1988).

As an example of how the information load can be used to predict perceived organization, let us consider a case of perceptual grouping in serial patterns. In Van der Helm et al. (1992), subjects had to indicate their preferred organization of a symbol series, by way of a paired-comparison (see Figure 1A). Two possible organizations of the series were given by means of a grouping of symbols (Figures 1B and 1C). While the black and white dots in A can be represented by the series 'ababb', the groupings shown in B and C reveal specific encodings of that series, i.e. '2*(ab) b', and 'S[(a),(b)] 2*(b)', respectively. According to the information measure outlined above, the I-load of the first code equals 4 (based on the parameters 'a', 'b', another 'b' and the group 'ab'), whereas the I-load of the second code equals 3 (based on the parameters 'a', 'b', and another 'b'). On the basis of this difference, the latter

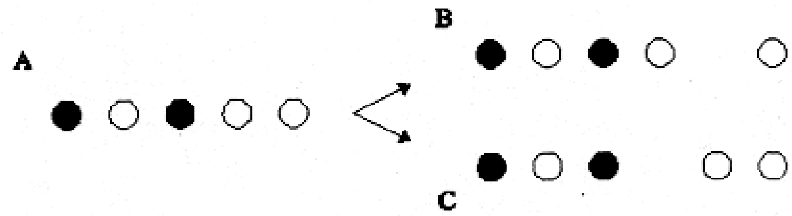

Figure 1. Two possible groupings within a symbol series (A). Generally C is preferred to B, which agrees with SIT.

organization is predicted to be preferred (which, in fact, was in line with the actual preferences). Notice that if all parameters and operators contributed to the information load, the prediction would be reversed; in this case the I-loads would be 4 and 5, respectively. Notice further that the new information load, by way of accounting for descriptive hierarchy, generally favors clustering in smaller units.

Obviously, the minimum code does not necessarily express all identities in a pattern. This is a consequence of the specific coding language, which allows only specific combinations of regularities within one code. This 'underspecification' might be considered (at first glance) as a drawback of the coding system. However, it should be realized that demanding the description of all identities within one code would result in mere template codes without any classification implications. As a matter of fact, the description of subsets of identities is essential for the classification of patterns (cf. Collard & Buffart, 1983). Yet, nondescribed identities may have a perceptual role as well, as their description allows alternative interpretations of the same pattern. For example, the series 'aaba' can be encoded into 'aS[a,b]', expressing the identity of the second and fourth element, and into '2*(a)ba', expressing the identity of the first and second element. In neither case are all possible identities accounted for. Because of this "incompleteness", Collard and Buffart (1983) suggested that identities which remain unexpressed in the minimum code are expressed in complementary codes revealing different interpretations of the same pattern. The perceptual relevance of such alternative codes has been demonstrated in various experiments on the salience of pattern interpretations (Mens & Leeuwenberg, 1988). The rivalry between different interpretations leads us to the following topic.

Diversity and integration by simplicity

Local versus global completions. In line with the minimum principle, it is assumed that the perceptual system generates multiple interpretations and that the simplest one is selected. Herbart (1850) already suggested such a competition between different interpretations. According to Herbart the dominant percept may vary in strength, depending on the attractiveness of the competing interpretations. We will illustrate this rivalry by means of some research in the amodal completion domain. When looking around we notice that objects usually occlude parts of themselves and parts of other objects. In those cases, the perceptual system seems to complete the occluded part. Two completion tendencies are often distinguished, namely local and global tendencies. Local completions merely depend on local shape characteristics at regions of occlusion (e.g., Kellman & Shipley, 1991; Wouterlood & Boselie, 1992). Global completions take into account global figural regularities, such as symmetry or repetition (e.g., Buffart, Leeuwenberg & Restle, 1981). In Van Lier, Van der Helm, and Leeuwenberg (1995) and Van Lier, Leeuwenberg and Van der Helm (1995), we focussed on both local and global completion tendencies. In these studies, patterns were examined for which the local completion tendency proceeded by means of linear continuation of the occluding contours (i.e., the completion itself was as simple as possible), whereas the global completion tendency resulted in the most regular shape (i.e., the simplest completed shape). Patterns can be constructed for which, on the one hand, local and global strategies converge to the same shape and, on the other hand, local and global strategies diverge to different shapes. In Figure 2, an example is given of both type of patterns. Clearly, the latter type of patterns seem to be much more ambiguous than the former type of patterns.

In Van Lier, Van der Helm, and Leeuwenberg (1995), participants were asked to draw their spontaneous pattern interpretation, (like in, e.g., Buffart, Leeuwenberg, & Restle, 1981;

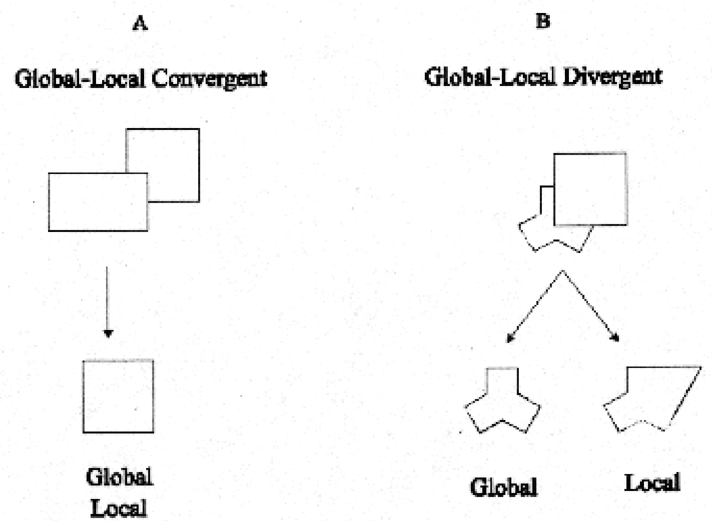

Figure 2. Two types of completion patterns. In A, Global and local strategies converge to the same completion. In B, Global and local strategies diverge to the different completions.

Boselie, 1988; Boselie & Wouterlood, 1989) whereas a second experiment concerned the simultaneous-matching task, partly adopted from Gerbino and Salmaso (1987). The results revealed an interdependency of the strengths of local and global completions and supported the notion that the preference for a specific completion is the outcome of a competition between these completions. Additionally, in Van Lier, Leeuwenberg, and Van der Helm (1995), converging evidence was found for the generation of multiple completions (using the so-called primed-matching paradigm; an experimental method which is based on the facilitating effect of a prime on the matching of a pair of representationally similar successive test items; see Beller, 1971; Sekuler & Palmer, 1992; Bruno, Bertamini, & Domini, 1997). Comparable results confirming the special status of local and global completions were obtained by Sekuler (1994) and Sekuler, Palmer and Flynn (1994).

On the one hand, these results demonstrated that a strict local approach to visual occlusion does not always correctly predict the perceived interpretations. On the other hand, some cases of local completion were evident, so the observed tendencies challenged SIT's simplicity approach. In the next section, I review SIT's integrative approach dealing with both local and global aspects.

SIT's integrative approach. In Van Lier, Van der Helm and Leeuwenberg (1994) we have argued that both the simplicity of the shapes and their relation with the pattern must be considered in determining the complexity of an interpretation. In our view, this relational aspect is of decisive importance in achieving an interpretation. To that end, a distinction was made between the memory complexity and the perceptual complexity of an interpretation. Whereas the memory complexity specifies the amount of information that is required to store an interpretation in memory, the perceptual complexity is based on the relationship between the pattern and its interpretation. The perceptual complexity can be considered as a selection criterion related to the difficulty of achieving a specific interpretation for a given visual pattern (as proposed in Van Lier et al., 1994). SIT's integrative approach is based on the observation that both local and global aspects jointly determine the perceptual complexity. These aspects are embedded in three different types of structure, indicated by the internal structure, the external structure, and the virtual structure, respectively. As outlined below, for each pattern interpretation the complexities of each of these structures can be determined in terms of structural information.

The internal structure of an interpretation deals with the perceived shapes themselves. To determine the complexity of the internal structure, the shape is mapped to a symbol series. Heuristically, this can be done by tracing the contour of the shape after having labeled all edges and angles such that equal edges and equal angles obtain equal symbols. Second, the symbol series is reduced by extracting the maximum amount of regularity from such a series. For example, a square can be represented by the symbol series, or primitive code, 'kakakaka' (with 'k' referring to, e.g., the edges and 'a' to the angles) which can be reduced by applying the Iteration rule (revealing '4*(ka)', with $I_{internal}$=3). In general it holds that the more regular the shapes are, the simpler is its description, and the lower the complexity of the internal structure (compare, for example, the primitive code of a rectangle which could be 'kalakala', revealing the minimal code: 2*(<(k)(l)>/<(a)>), with $I_{internal}$=4).

The external structure refers to the positional relation between the perceived shapes. The account of the external structure can be considered as related to Rock's avoidance-of-coincidence principle. Rock (1983), argued that the visual system seeks interpretations in which positional coincidences are avoided. As an example, consider Figure 2A once more. The preferred occluded-square interpretation can be explained on the basis of a tendency toward the simplest shape but one may also argue that the preferred interpretation avoids a coincidental meshing of borders as would, for example, be the case if there was no completion (i.e., the so-called 'mosaic' interpretation, in which an L-shape is adjacent to a rectangle). Notice that applying the avoidance-of-coincidence principle to the pattern in Figure 2B, does not reveal which completion (global or local) would be predicted to be preferred. The simplest-shape tendency and the avoidance-of-coincidence principle do not always lead to the same predictions. In Figure 3 (inspired by Kanizsa, 1985), for example, patterns A-C are all perceived as a black square occluding a second shape. The latter could be a regular octagon (D) or a more complex shape (E). Now, the preference for interpretation D appears to be weakest in pattern A and strongest in pattern C, leaving pattern B in an intermediate position. This phenomenological difference in completion strength seems to vary

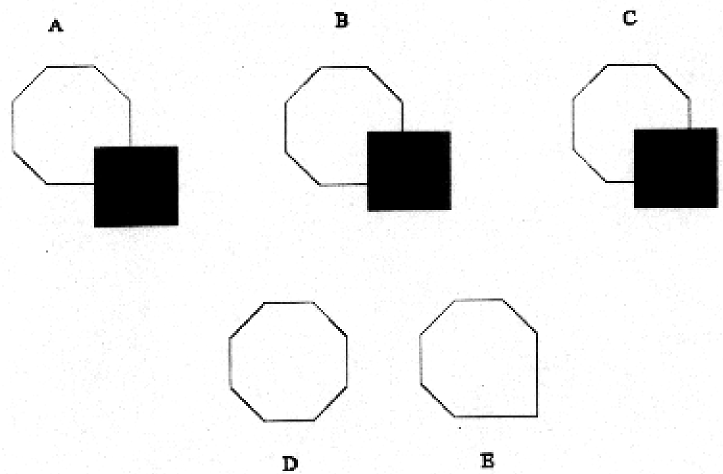

Figure 3. Patterns A-C can all be interpreted as a square occluding a second shape (as in D and E). Apparently, interpretation D is less prevalent in A and most prevalent in C, whereas for interpretation E the reverse holds. This agrees with the tendency to avoid coincidences between the contours of the shapes.

with the degree of coincidence between the location of the edges of the square and the location of the corners of the octagon (notice that the numbers of times the edges and corners are at the same location decreases going from A to C, see Van Lier, Van der Helm, & Leeuwenberg, 1995).

To illustrate the relationship between inter-object regularity and perceived (im)probability further, consider the following example. In Figure 4, four configurations of two matches are depicted as they may show up after a random throw. Many observers would rate the configurations on the left to be far more probable the result of a random throw, than the configurations on the right. In fact there is no such an unequivocal difference in probability of occurrence between these configurations. What differs, however, is that there are more configurations like A than there are like D. That is, the class of all 'A-like' configurations is much larger than the class of all 'D-like' configurations. The number of configurations within a class is determined by the regularities within that class. The more regular the positions, the smaller the class, the less likely a specific configuration seems to be the result of a random throw. In this example, the positional regularity between the matches can be conceived as a perceptual obstacle with respect to the two-matches interpretation. In Van Lier et al. (1994), $I_{external}$ was determined for various types of junctions and their dissociations (with respect to the configurations A-D of Figure 4, for example, $I_{external}$ would be 0, 1, 2 and 3, respectively). In general, it holds that the more coincidental is a specific junction between elements belonging to different objects, the higher is the complexity of the external structure (notice that this analysis is in line with the notion that connection between surfaces is a very powerful grouping factor, e.g., Saiki & Hummel, 1998; Van Lier & Wagemans, 1998).

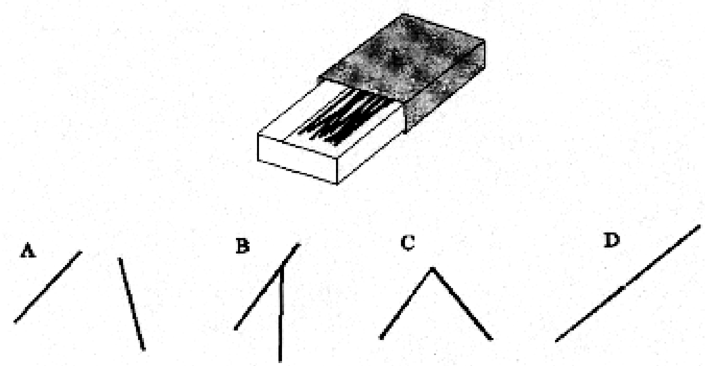

Figure 4. Four configurations of matches (e.g., after an arbitrary throw). The fact that there are more configurations like A than there are configurations like D explains the difference in perceived probability of occurrence. The perceived probability can be related directly to the regularity between the matches.

Finally, the virtual structure deals with the occluded (non-visible) parts of the perceived shapes. More specifically, it concerns the elements that are not present in the proximal stimulus, but are present in the interpretation of that stimulus. For the pattern in Figure 2B, the global completion has the simplest shape, yet seems to be disfavored because of the larger number of occluded structural elements (with regard to Figure 2B, for example, $I_{virtual}$ would be 5 for the global completion and 1 for the local completion). In Van Lier et al. (1994) and Van Lier, Van der Helm, and Leeuwenberg (1995) we have argued that the more structural elements of a specific surface interpretation are occluded, the less likely it is that the interpretation will be preferred.

So, resuming, regularities in the internal structure support a specific interpretation, whereas regularities in the external structure weaken that interpretation. Additionally, a specific completion is weakened by an increase in the complexity of the virtual structure. To a certain extent, $I_{internal}$, $I_{external}$, and $I_{virtual}$ correspond to three well-known tendencies in the domain of visual occlusion, namely the simplicity of shape, the avoidance of coincidence, and the good-continuation principle (the latter as far as completion is concerned). In Van Lier et al. (1994), we have tested the hypothesis that the sum of the perceptual complexities of the three structures, i.e. the total perceptual complexity (I_{total}) of the most preferred interpretation of a pattern is lower than that of any other interpretation of that same pattern. This was done on a set of 144 patterns stemming from different studies (Boselie, 1988; Boselie & Wouterlood, 1989; Buffart et al., 1981). It appeared that of these patterns the most preferred interpretation had the lowest $I_{internal}$ in 53% of all cases, the lowest $I_{external}$ in 65% of all cases, and the lowest $I_{virtual}$ in 49% of all cases. Only in 3% of all cases all three structures had the lowest complexity. However, I_{total} was the lowest for the most preferred interpretation in 95% of all cases. In Figure 5, various examples are given. The regions in the Venn-diagram diagram provide a way to classify completion patterns according to their most preferred completion tendencies in terms of the three separate structures. For example, for pattern A, only $I_{internal}$ would predict the most preferred interpretation correctly, whereas for pattern B, both $I_{internal}$ and $I_{external}$ would predict correctly, etc. It should be noted, however, that for each and every pattern the inter-pattern comparison on I_{total} determines the most preferred interpretation. Thus, for example, a high value for $I_{virtual}$ for a specific interpretation does not imply that this specific interpretation is unlikely (see also Van Lier, 2000). One may argue that the summation of the complexities is rather arbitrary and that possibly weighing factors could have been added. In Van Lier et al. (1994), however, we have shown that this simple summation of terms revealed good predictions (revealing an Ochamian parsimony-justification for this specific choice). Further on in this chapter (section 'simplicity versus likelihood'), an additional argument will be given in favor of the summation of the terms.

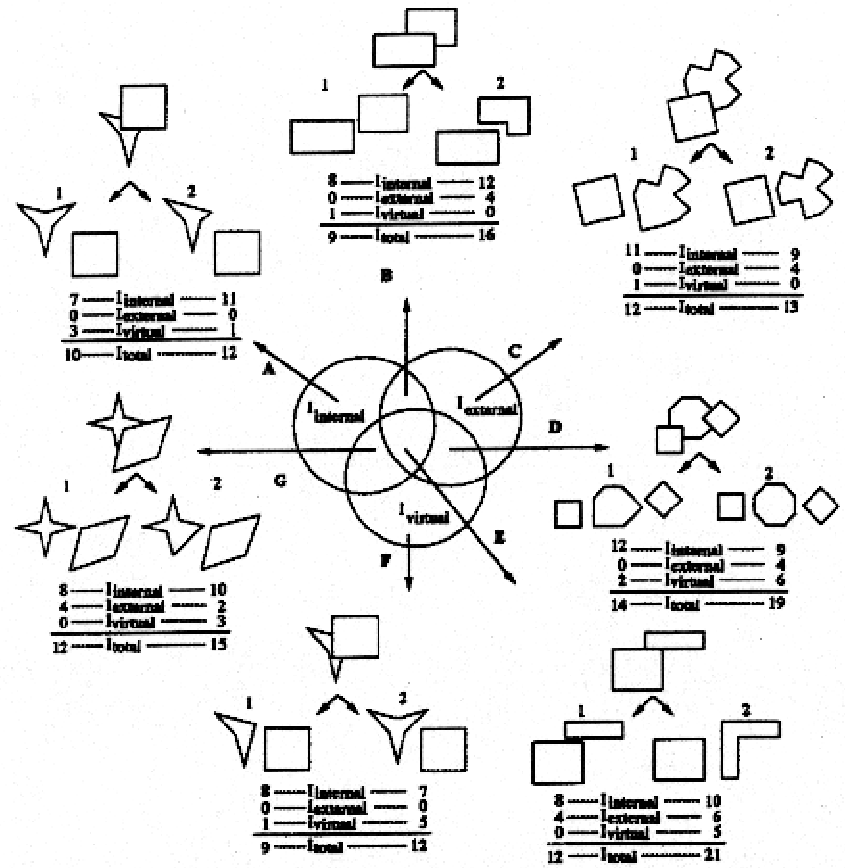

Figure 5. Various patterns with possible interpretations. The circles in the Venn-diagram indicate whether the complexity (I) for a specific structure (internal, external, virtual) is lowest for the most preferred interpretation (labeled with number 1). For each region an example is given. In each of these cases Itotal is lowest for the most preferred interpretation (adapted from Van Lier et al., 1994).

From 2D surfaces to 3D objects

SIT's structural descriptions are not restricted to series or surfaces but apply to 3D objects as well (e.g. Leeuwenberg 1971; Leeuwenberg & Van der Helm, 1991; Leeuwenberg, Van der Helm, and Van Lier, 1994; Van Lier, 1999a, Van Lier, Leeuwenberg, & Van der Helm, 1997; Van Lier & Wagemans, 1999). Again, in line with the notion of simplicity in visual shape, and in analogy with the implications on the previous domains, the simplest object description is predicted to be selected by the visual system. To illustrate SIT's object descriptions, consider Figure 6 (for a more detailed explanation see e.g. Leeuwenberg & Van der Helm, 1991).

The vase-like object (Figure 6A) can be described by means of two components: a circular component (O-shape) and an S-shaped component (S-shape). Now, moving the S-shape along the O-shape (Figure 6-A1), under a specific angle of connection between the two shapes, would mimic the contours of the vase. If, the other way around, the O-shape is moved along the S-shape, an S-shaped tube would evolve (Figure 6-B1). Notice that the vase could, in principle, also be reconstructed with, for example, the S-shape as path (Figure 6-A2). In such a case, however, the O-shape would have to vary its size while moving along the S-shape. Moreover, the angle of connection between the shapes would vary. Therefore, the latter description is much more complex and less preferable. In a similar way, the description depicted in Figure 6-B2 would be much less preferable. Within SIT, the component that serves as path in the simplest reconstruction is referred to as the superstructure (O-shape in Figure 6A, S-shape in Figure 6B), whereas the components that have to be moved along that path (S-shape in Figure 6A, O-shape in Figure 6B) are referred to as subordinate structures. This hierarchical relationship between descriptive components is a crucial aspect of SIT's

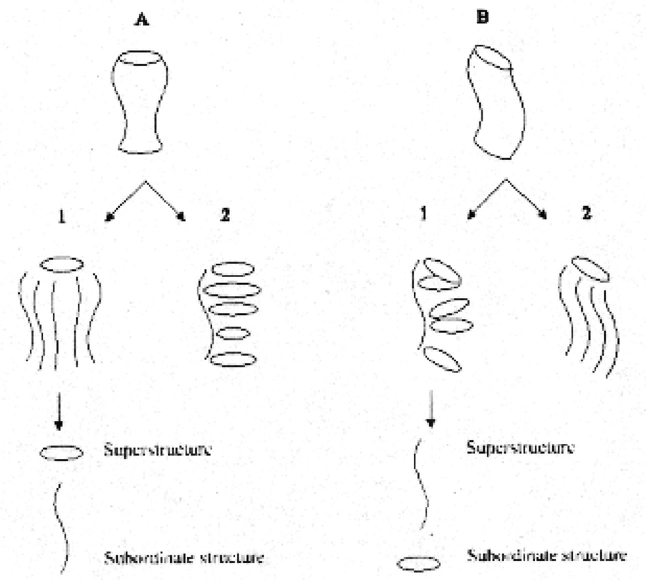

Figure 6. According to SIT, objects A and B can be described by the same components (an S-shape and an O-shape). For A, the simplest description is achieved when taking the O-shape as superstructure and the S-shape as subordinate structure (as in A1), whereas for B the situation is reversed (as in B1, see text).

object representations.

It is important to notice that the descriptive components in principle may have any possible shape, as long as they describe the object with a minimum of structural variation -- or stating it otherwise, with a maximum of structural constancy -- as possible. The descriptive components are therefore not given a priori but in fact stem from the simplest representation. Notice further that the super-subordinate assignment proceeds irrespective of the actual relative size of the components. These characteristics contrast with other structural description models, perhaps most noticeably with the Recognition By Components model (RBC; Biederman, 1987). In the latter approach, descriptive components of certain simple objects, called geons, are defined a priory by means of so-called nonaccidental properties (i.e., properties of the 2D projection that generally do not arise from projections of arbitrary oriented 3D objects). In Leeuwenberg et al. (1994), object classifications according to RBC and SIT were compared with each other. Whereas within RBC the selection of the internal axis of an object depends on metrical considerations (e.g., the direction of the longest elongation), within SIT such considerations do not play a role. Leeuwenberg et al. (1994) demonstrated that that judged similarities between objects largely rely on structural aspects, leaving a secondary, non-dominant, role for metrical aspects (more specifically it was shown that such judgements followed classifications based on constancy and symmetry of the descriptive parts rather than their size). Additionally, Van Lier et al. (1997) investigated whether the descriptive components at the superstructural level perceptually dominate the components at the subordinate level (i.e. the superstructure-dominance hypothesis, see also Leeuwenberg & Van der Helm, 1991). From the experiments, in which the primed matching paradigm (as in, e.g., Sekuler & Palmer, 1992) was used, it appeared that, amongst others, the matching of objects was facilitated much more if their superstructure was primed than if their subordinate structure was primed. We concluded that this supported the notion that there is an internal superstructure dominance, which can be assessed during the construction of an object representation. Other experiments (e.g., Van Bakel, 1989; see also Leeuwenberg & Van der Helm, 1991) have supplied converging evidence on the differential status of superstructures and subordinate structures. The results suggested that the perceived unity in visual stimuli primarily depends on the unity of the superstructure (more specifically, it appeared that a subdivision or segmentation in the superstructure affected unity judgements more strongly than a subdivision or segmentation in the subordinate structure did). Ongoing research aims at a further investigation of the generalizability of these findings.

Additionally, the results of SIT's integrative approach on surfaces as outlined previously, support future investigations on the applicability of the integrative model on 3D object perception. Recent research on preferred completions of self-occluded 3D object parts (Van Lier & Wagemans, 1999) has already shown the impact of intrinsic object regularities on a set of 3D objects in which the differential effect of the external structure and the virtual structure were minimized. The combined impact of the three structures can be illustrated by

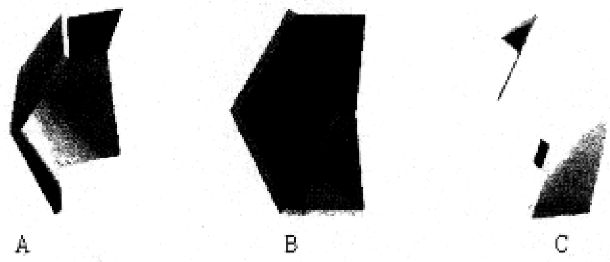

Figure 7. Three possible views of the same object. The internal structure (a viewpoint independent aspect) is the same for all views, whereas the external and the virtual structure (being viewpoint dependent aspects) vary across the images.

means of the three images in Figure 7.

The images A to C (Figure 7) in fact comprise one and the same object, viewed from three different orientations. Therefore, the internal structure (dealing with the object itself) does not vary between the images. This is not the case for the external structure and the virtual structure. In Figure 7B, for example, the external structure (dealing with the spatial positioning of the object) disfavors the 3D interpretation, due to the very specific object orientation, inducing coincidental alignments of object parts. Furthermore, the virtual structure (dealing with the object parts that are not visible) varies between images. An interesting aspect of such a 2D-to-3D extension is the fact that SIT's integrative approach actually accounts for viewpoint independent aspects and viewpoint dependent aspects. That is, the internal structure can be considered as viewpoint independent as it merely describes intrinsic object characteristics, whereas both the external structure and the virtual structure are of a viewpoint dependent nature as they evidently depend on the positional relationship between the observer and the objects. Recent discussions in the literature have made clear that both viewpoint independent aspects and viewpoint dependent aspects need to be accounted for in explanatory models of object perception. For example, Biederman and Gerhardstein (1993) suggested that, within the RBC model on object perception, viewpoint independence does not hold in case of accidental views or in case of occlusion of crucial object parts. In terms of the integrative approach, these two restrictions on viewpoint independency conceptually fit with the account of the external structure and the virtual structure, respectively. More specifically, accidental viewpoints would lead to high $I_{external}$ values (because of the coincidental image regularities) and the occlusion of complex parts would lead to high $I_{virtual}$ values. In general, the approach implies that intrinsic object regularities increase object identification, whereas accidental views or views in which large parts are (self-)occluded decrease object identification. Current research (Van Lier, 1999b) aims at the further development of SIT's integrative approach, and the underlying regularity-based

simplicity metric, as an account of both viewpoint independent and viewpoint dependent aspects. In this way, the proposed extension of the integrative approach links some well-known opposing tendencies in 2D surface occlusion and completion research with those of 3D object recognition, based on the notion of global simplicity.

Global Simplicity: Further Observations

Simplicity versus Likelihood. A theoretical problem which deserves some attention here and which links up with the previous issue is whether perception is simplicity-based or likelihood-based. The likelihood principle, originally formulated by Von Helmholtz (1867/1962), states that the preferred perceptual organization reflects the most likely object or event. Several researchers (Mach, 1886; Sober, 1975; see also Hatfield & Epstein, 1985; Chater, 1996) have argued that the likelihood principle and the minimum principle could be two sides of the same coin, as in many cases the most likely interpretation is the simplest interpretation and vice versa. Although we do not agree with this exchangeability of concepts (e.g., Leeuwenberg and Boselie, 1988; Leeuwenberg et al., 1994; Van der Helm, 2000; Van Lier, 1996) there is an interesting parallel between SIT's integrative approach (Van Lier et al. 1994, Van Lier, 2000) and Bayesian likelihood inferences as recently shown by Van der Helm (2000). According to the latter, the probability of a certain event can be considered as being based on an a priori probability and a conditional probability. An analogous distinction can be made in SIT's integrative approach (Van Lier et al., 1994; Van Lier, 1999b). That is, the viewpoint independent component (i.e., the internal structure) is to be considered as the a priori argument whereas the viewpoint dependent components (i.e., the external structure and the virtual structure) are to be considered as conditional arguments. More specifically, Van der Helm (2000) argues that whereas the likelihood principle can be considered to be based on the maximization of the probability (*p*) for a certain hypothesis (*H*) and data (*D*) according to the equation $p(H|D) = p(H) \times p(D|H)$ (focusing on the numerator in Bayes rule), the simplicity principle is based on the minimization of the complexity or information load (*I*) according to the equation $I(H|D) = I(H) + I(D|H)$, where *H* can be read as 'the interpretation' and *D* as 'the proximal stimulus'. These equations merge into each other under the substitution $I = -\log p$ (e.g., Pomerantz & Kubovy, 1986) or $p = 2^{-I}$ (Leeuwenberg & Boselie, 1988; although care should be taken as both *p* and *I* are expressed in different ways in the latter equations, see also Van der Helm, 2000). Notice that the foregoing provides supplementary support (beside mere parsimony) --from a rather unexpected side-- for taking the mere sum of the complexity values in the determination of the I_{total}. As indicated, this does not mean that there would be a conceptual equivalence between the approaches. The intuitive (im)probability of a specific positional relation between objects (or between observer and objects) actually depends on regularity-based categorization but not the other way around (see also Leeuwenberg & Boselie, 1988; Van Lier et al. 1994; Van der Helm, 2000). In everyday life, however, the probability of occurrence is such a compelling subjective experience that it is often taken as a cause rather than as an effect. Regarding the nonaccidental properties, we do not deny their effects, but, again, we do deny that probability is the appropriate concept for explaining them.

Local prevalence, global simplicity. Before concluding this chapter it seems expedient to briefly return to the concepts of local and global. Although the terms local and global are frequently used in the literature, some care should be taken with these concepts as there is no unique definition of them. For example, Navon (1977) refers with the term global to the low resolution structure of a pattern in contrast to the fine-grained local details at high resolution levels (cf. Robertson, 1996). Somewhat related to this, the local-global distinction often refers to the span of certain pattern properties. In this way, the occurrence of a specific junction type can be considered as a local cue, whereas a bilateral symmetry can be considered as a more global property, as it covers a greater part of a pattern. As argued in the previous sections, both local and global figural aspects influence the preferred percept, whereas predictions of preferred interpretations were always based on a global measure of simplicity. These predictions concern, for example, grouping of serial symbols, completion of surfaces, and the dominance of object components. In Figure 8, from each of these stimulus domains, an example is presented once more. Below each pattern, two plausible 'solutions' are given. The qualifications "global" and "local" are given on the basis of merely metric stimulus properties. For pattern A, the grouping in A' has the largest component and has been qualified as global, whereas A" is qualified as local. For pattern B, completion B' is qualified as global, as it is the more symmetric one, whereas B" is qualified as local. For pattern C, the dominance

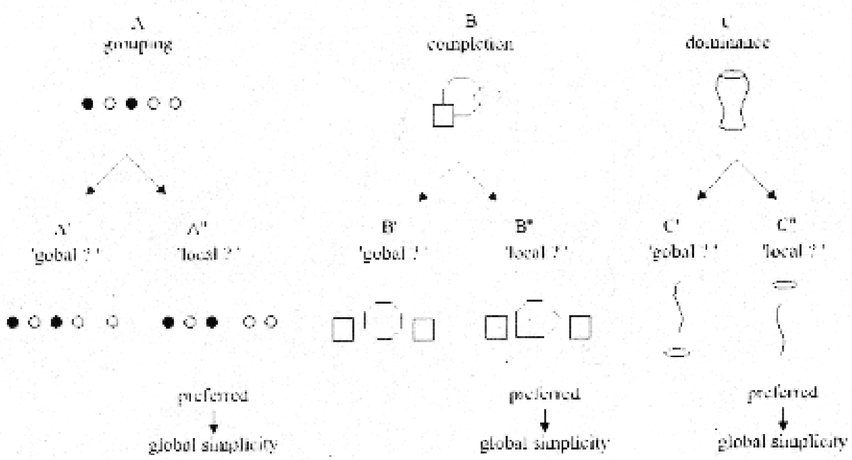

Figure 8. Three examples to illustrate that preferences may be labeled 'local' on an intuitive basis but are, in fact, predicted by a global measure of simplicity.

relationship in C' is qualified as global, as it considers the largest contour component as a superstructure, whereas C" is referred to as local. Notice, however, that for each pattern the 'local solution' is predicted on the basis of global simplicity. So, at an intuitive level there seems to be a discrepancy between the local prevalence and global simplicity whereas, in fact, the 'local solutions' are in line with the global predictions. The notions of local and global as indications of figural aspects are, instead of a priori explanatory concepts, best to be considered as a-posteriori assignments.

FUTURE RESEARCH AND CONCLUDING REMARKS

The SIT approach is still in a state of flux and will be further developed in the coming years. As mentioned, the extension of SIT's integrative approach to 3D objects, incorporating viewpoint independent and viewpoint dependent aspects, is one of the current research lines (Van Lier, 1999b). To explicate this extension a little further, it can be said that, in addition to the simplicity of the perceived shape, the goodness of the 2D image for a specific 3D object -- i.e., the goodness of view-- will be taken into consideration. Whereas both accidental views and occlusions tend to produce relatively simple 2D images (imagine, for example, a bucket seen from above), good views generally reveal rather complex images. One of the current issues concerns the relation between, on the one hand, the goodness of view, and, on the other hand, the difference between the complexity of the actual image of an object and the complexity of a general viewpoint image of that object (in which most of its structure is visible in a nonaccidental way, i.e., no coincidental 2D image regularities).

The above extension from 2D to 3D implies a gradual incorporation of more natural stimuli. A further step in that direction concerns the extension of the domain to quasi-regular shapes (as in Van Lier, 1999a). The patterns in Figure 9, for example, clearly reveal a limited class of plausible interpretations. Although the exact metrical values of edges and angles are all different, it appears that preferred completions reveal the same 'fuzzy' characteristics as the visible part. Alternative descriptions that capture such characteristics for both 2D and 3D shapes (as proposed in Van Lier 1999a) will be a topic of further research.

SIT is first of all a quantitative, representational, approach to perception. Its formalization of regularity, combined with the notion of simplicity, allows classification and prediction of perceptual interpretations. An advantage of SIT is that it is applicable to a wide range of stimuli, treating various perceptual phenomena from the same point of view. It will be clear, however, that SIT is not a process theory. That is, within SIT, perceptual organization is primarily attributed to the simplicity of representation and not to certain process assumptions (e.g., Leeuwenberg & Van der Helm, 1991). It's approach to perceptual organization phenomena provides a way of thinking about perceptual issues that is firmly rooted in the Gestalt tradition and until the present day builds further on those initial ideas

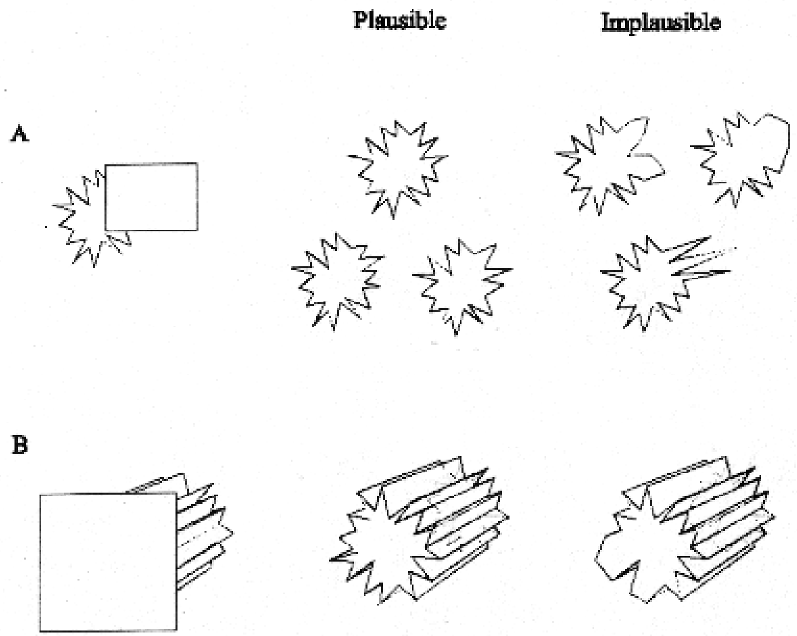

Figure 9. Plausible interpretations of A and B are clearly not restricted to a few exemplars but may actually extent to a whole class of interpretations.

(see also Palmer, 1999). Hopefully, the open questions mentioned in this chapter, will lead to further investigations into SIT.

ACKNOWLEDGMENTS

The text of this chapter is partly based on the author's PHD-thesis. Part of the research reviewed in this chapter has been made possible by the Netherlands Organization for Scientific Research (NWO). The writing of this chapter has been made possible by a grant from The Royal Netherlands Academy of Arts and Sciences (KNAW). RVL thanks Emanuel Leeuwenberg and Peter Van der Helm for their support.

REFERENCES

Baylis, G. & Driver, J. (1995). Obligatory edge assignment in vision: The role of figure and part segmentation in symmetry detection. *Journal of Experimental Psychology:*

Human Perception and Performance, 21, 1323-1342.

Beller, H. (1971). Priming: Effects of advance information on matching. *Journal of Experimental Psychology, 87,* 176-182.

Biederman, I. (1987). Recognition by components: A theory of human image understanding. *Psychological Review, 94,* 115-147.

Biederman, I., & Gerhardstein, P. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1162-1182.

Boselie, F, & Leeuwenberg, E. (1986) A test of the minimum principle requires a perceptual coding system. *Perception, 15,* 331-354.

Boselie, F. (1988). Local versus global minima in visual pattern completion. *Perception & Psychophysics, 43,* 431-445.

Boselie, F., & Wouterlood, D. (1989). The minimum principle and visual pattern completion. *Psychological Research, 51,* 93-101.

Bruno, N., Bertamini, M., & Domini, F. (1997). Amodal completion of partly occluded surfaces: Is there a mosaic stage? *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1412-1426.

Buffart, H., Leeuwenberg, E., & Restle, F. (1981). Coding theory of visual pattern completion. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 241-274.

Chater, N. (1996). Reconciling simplicity and likelihood principles in perceptual organization. *Psychological Review 103,* 566-581.

Collard, R, & Buffart, H. (1983) Minimization of structural information: A set-theoretical Approach. *Pattern Recognition, 16,* 231-242.

Deutsch, D, & Feroe, J. (1981) The internal representation of pitch sequences in tonal music. *Psychological Review, 88,* 503-522.

Garner, W. (1974). *The processing of information and structure.* Potomac, MD: Erlbaum.

Gerbino, W., & Salmaso, D. (1987). The effect of amodal completion on visual matching. *Acta Psychologica, 65,* 25-46.

Hatfield, G., & Epstein, W. (1985). The status of the minimum principle in the theoretical analysis of visual perception. *Psychological Bulletin, 97,* 155-186.

Herbart, J. (1850) *Lehrbuch zur Psychologie* (Leipzig: Leopold Voss Verlag).

Hochberg, J., & McAlister, E. (1953). A quantitative approach to figural 'goodness'. *Journal of Experimental Psychology, 46,* 361-364.

Jones, M, & Zamostny, K. (1975). Memory and rule structure in the prediction of serial patterns. *Journal of Experimental Psychology: Human Learning and Memory, 104,* 295-306.

Kanizsa, G. (1985). Seeing and thinking. *Acta Psychologica, 59,* 23-33.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception.

Cognitive Psychology, 23, 141-221.

Köhler, W. (1920) *Die Physischen Gestalten in Ruhe und im stationären Zustand.* Braunschweig: Vieweg.

Koffka, K. (1935). *Principles of Gestalt psychology.* New York: Harcourt, Brace & World.

Leeuwenberg, E. (1969). Quantitative specification of information in sequential patterns *Psychological Review, 76,* 216-220.

Leeuwenberg, E. (1971). A perceptual coding language for visual and auditory patterns. *American Journal of Psychology, 84,* 307-349.

Leeuwenberg, E. & Boselie, F. (1988). Against the likelihood principle in visual form Perception. *Psychological Review, 95,* 485-491.

Leeuwenberg, E., & Van der Helm, P. (1991). Unity and variety in visual form. *Perception, 20,* 595-622.

Leeuwenberg, E., Van der Helm, P., & Van Lier, R. (1994). From geons to structure. A note on object representation. *Perception, 23,* 505-515.

Mach, E. (1886) *Beiträge zur Analyse der Empfindungen.* Jena: Gustav Fisher.

Mens, L., & Leeuwenberg, E. (1988). Hidden figures are ever present. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 561-571.

Navon, D. (1977). Forest before trees: the precedence of global features in visual perception. *Cognitive Psychology, 9,* 353-383.

Palmer, S. (1991). Goodness, Gestalt, Groups, and Garner: Local symmetry subgroups as a theory of figural goodness. In G. Lockhead, & J. Pomerantz (Eds.), *The Perception of Structure.* (pp. 23-39). WashingtonDC: APA.

Palmer, S. (1999). *Vision Science: Photons to Phenomenology.* Cambridge, MA: MIT Press.

Pomerantz, J. & Kubovy, M. (1986). Theoretical approaches to perceptual organization. In K. Boff, L. Kaufman, J. Thomas (Eds), *Handbook of Perception and Human Performance.* (pp.1-46). New York: John Wiley.

Restle, F. (1979). Coding theory of the perception of motion configurations. *Psychological Review, 86,* 1-24.

Restle, F. (1982). Coding theory as an integration of gestalt psychology and information processing theory. In J. Beck (Ed), *Organization and representation in perception* (pp. 31-56). Hillsdale, NJ: Lawrence Erlbaum.

Robertson, L. (1996). Attentional Persistence for features of hierarchical patterns. *Journal of Experimental Psychology: General, 125,* 227-249

Rock, I. (1983). *The logic of perception.* Cambridge MA: MIT Press.

Sekuler, A. (1994). Local and global minima in visual completion: Effects of symmetry and orientation. *Perception, 23,* 529-545.

Sekuler, A., & Palmer, S. (1992). Perception of partly occluded objects: A microgenetic analysis. *Journal of Experimental Psychology: General, 121,* 95-111.

Sekuler, A., Palmer, S., & Flynn, C. (1994). Local and global processes in visual completion.

Psychological Science, 5, 260-267.

Simon, H., & Kotovsky, K. (1963) Human acquisition of concepts for sequential patterns. *Psychological Review, 70,* 534-546.

Sober, E. (1975) *Simplicity.* London: Oxford University Press.

Van Bakel, A. (1989) *Perceived unity and duality as determined by unity and duality of superstructure-components of pattern codes.* Masters Thesis, University of Nijmegen.

Van der Helm, P. (1988). *Accessibility and simplicity of visual structures.* PhD Thesis. University of Nijmegen.

Van der Helm, P. (2000). Simplicity versus Likelihood: From Surprisals to Precisals. *Psychological Bulletin, 126,* 770-800.

Van der Helm, P., & Leeuwenberg, E. (1986). Avoiding explosive search in automatic selection of simplest pattern codes. *Pattern Recognition, 19,* 181-191.

Van der Helm, P., & Leeuwenberg, E. (1991). Accessibility, a criterion for regularity and hierarchy in visual pattern codes. *Journal of Mathematical Psychology, 35,* 151-213.

Van der Helm, P., & Leeuwenberg, E. (1996). Goodness of visual regularities: A non-transformational approach. *Psychological Review, 103,* 429-456.

Van der Helm, P., Van Lier R., & Leeuwenberg, E. (1992). Serial pattern complexity: Irregularity and hierarchy. *Perception, 21,* 517-544.

Van Lier, R. (1996) *Simplicity of visual shape: A structural information approach.* PhD-Thesis, University of Nijmegen.

Van Lier, R. (1999a). Investigating global effects in visual occlusion: From a partly occluded square to the back of a tree trunk. *Acta Psychologica* (Special issue: 'Object Perception and memory') *102,* 203-220.

Van Lier, R. (1999b). From image to object: An integration of global and local aspects. Internal report (Project proposal), University of Nijmegen.

Van Lier, R. (2000) Separate features versus one principle: A comment on Shimaya (1997). *Journal of Experimental Psychology: Human Perception and Performance, 26,* 412-417.

Van Lier, R., Van der Helm, P., & Leeuwenberg, E. (1994). Integrating global and local aspects of visual occlusion. *Perception, 23,* 883-903.

Van Lier, R., Van der Helm, P., & Leeuwenberg, E., (1995). Competing global and local completions in visual occlusion. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 571-583.

Van Lier, R., Leeuwenberg, E., & Van der Helm, P. (1995). Multiple completions primed by occlusion patterns. *Perception, 24,* 727-740.

Van Lier, R., Leeuwenberg, E., & Van der Helm, P. (1997). In support of hierarchy in object representations. *Psychological Research, 60,* 134-143.

Van Lier, R., & Wagemans, J. (1998). Effects of physical connectivity on the representational unity of multi-part configurations. *Cognition, 69,* B1-B9.

Van Lier, R., & Wagemans, J. (1999). From images to objects: Global and local completions of self-occluded parts. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1721-1741.

Vitz, P. (1966) Preference for different amounts of visual complexity. *Behavioral Science, 11,* 105-114.

Vitz, P., & Todd, R. (1969) A coded element model of the perceptual processing of sequential stimuli. *Psychological Review, 76,* 433-449.

Von Helmholtz, H. (1867) *Handbook on Physiological Optics* Vol. III. Translation (1962) from the 3rd German edition (New York: Dover Publications)

Wertheimer, M. (1912) Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie, 12,* 161-265.

Wouterlood D., & Boselie, F. (1992). A good-continuation model of some occlusion phenomena. *Psychological Research*, 54, 267-277.