Perception & Psychophysics
1993, 53 (3), 350-352

The notion of elements in the visual field in a theory of visual attention: A reply to van der Velde and van der Heijden (1993)

CLAUS BUNDESEN
Copenhagen University, Copenhagen, Denmark

Van der Velde and van der Heijden (1993) criticized the way in which the notion of elements in the visual field is used in my theory of visual attention (Bundesen, 1990) and in my reinterpretation of Nissen's (1985) findings (Bundesen, 1991). The critique is repudiated here. I defend the critical assumption that perceptual units (elements in the visual field), formed on the basis of precategorical, sensory information, can be addressed by later processes before any perceptual categorizations have been made. I also point out that, despite the claims of van der Velde and van der Heijden, my reinterpretation of Nissen's findings is testable.

Van der Velde and van der Heijden (1993) claimed to show that the theory of visual attention presented by Bundesen (1990) "contains a serious flaw in the form of the abstract identity of an element, separate from its attributes, that is used in the process of perceptual categorization" (p. 349). They also claimed to show that "the attempt made by Bundesen (1991) to explain the results reported by Nissen (1985) on the basis of a model in which selection by location is treated on a par with selection by color or by shape has failed" (p. 349). In response to this critique, I will explain the notion of an element in the visual field in some detail. Subsequently, I will repudiate the claims made by van der Velde and van der Heijden.

The Notion of an Element in the Visual Field

In the theory of Bundesen (1990), the notion of an element in the visual field is treated almost as a primitive. Let me indicate how the theory fits in with a fairly broad and conventional conceptualization of visual processing (cf. Bundesen, 1992). The conceptualization makes the notion of an element more concrete.

Sensory processing. Visual processing of stimulus objects begins with registration of retinal images at the level of photoreceptors and proceeds through a number of stages. Broadly speaking, the first major stage of processing (the *sensory* stage) consists in extraction of information about the visible surfaces (cf. Gibson, 1950, on "literal perception," and Marr, 1982, on "early vision"). Presumably it produces a representation (a *visual impression*) in which properties such as local surface color and depth are made explicit for each position in the visual field (cf. Marr, 1982, on "the 2½-D sketch"). Sensory processing is parallel across the visual field, and to a first approximation, it is automatic and "bottom up" or "data driven" (see, e.g., Cornsweet, 1970; Marr, 1982; but see also Gregory, 1966, 1970).

Unit formation. Following Neisser (1967), Kahneman (1973), and others, I suppose that the next stage of processing generates a segmentation of the visual input by Gestalt grouping operations. At this stage, parts of the scene represented in the total visual impression are defined as separate *perceptual units* by criteria based on proximity, similarity, and continuity (for more detailed discussions, see, e.g., Beck, Prazdny, & Rosenfeld, 1983; Julesz, 1981; Olson & Attneave, 1970). By defining groups of perceptual units as higher level perceptual units, a hierarchical part-whole organization of perceptual units can be created (cf. Palmer, 1977). The set of perceptual units is the set of *elements in the visual field*. The process of defining elements in the visual field (i.e., unit formation) is mainly data driven (cf. Pomerantz, 1981), although "top-down" effects of familiarity and perceptual set have been recognized (see, e.g., Wertheimer, 1923).

Perceptual testing. For some categories F, a measure of the strength of the sensory evidence that an element x in the visual field belongs to category F is computed by comparing element x (actually, a structured visual impression of element x, formed at the previous stage of processing) with a memory representation of visual characteristics (such as color, shape, or location) of members of category F. A category F with this property is said to be a *perceptual category*. Examples of perceptual categories are the class of *red* elements (a color category), the class of *circular* elements (a shape category), and the class of elements located *above fixation* (a location category).

Attentional selection. In the theory of Bundesen (1990), visual recognition and attentional selection consist in making perceptual categorizations of the form "x belongs to F," where x is an element in the visual field and F is a perceptual category. The computed strength of the sensory evidence that x belongs to F is one of the factors that determine the likelihood that the perceptual categorization "x belongs to F" is made (selected). Other factors are pertinence values of perceptual categories (i.e., measures of the importance of attending to particular types of elements) and perceptual decision bias parameters. The way in which the factors interact is explained in the theory.

Three Points Made by van der Velde and van der Heijden

Van der Velde and van der Heijden (1993) summarized their critique in three points. Their third point is the most general one, and I will treat it first.

Point 3. Point 3 can be stated as follows. Bundesen's (1990, 1991) theory "assumes that each element x is la-

This research was supported by a grant from the International Human Frontier Science Program Organization. Correspondence should be addressed to C. Bundesen, Psychological Laboratory, Copenhagen University, Njalsgade 90, DK-2300 Copenhagen S., Denmark.

Copyright 1993 Psychonomic Society, Inc.

beled. . . . Labeling an element . . . requires a process of discrimination. . . . The only way to discriminate elements in the visual field is to refer to at least one attribute" (van der Velde & van der Heijden, 1993, p. 348). Therefore, (1) at least one attribute of an element must be known before the element can be labeled. That is, (2) at least one correct perceptual categorization of an element must have been made before the element can be labeled. Bundesen's analyses are untenable because they violate this principle.

In my opinion, Point 3 is untenable. Statement 1 may be warranted, but Statement 2 is not. It clearly makes sense to assume that in the visual system, color, shape, and location information about stimuli is available ("known") at a precategorical, sensory stage of processing and that such information can be used for defining ("labeling") elements at a stage of unit formation, before any perceptual categorizations are made. It also makes sense to assume that, once a unit has been formed, the unit can be used (referred to) by later processes. The way in which perceptual units are represented and addressed in the visual system is an open question, but several possibilities have been suggested.

An ordinary way of forming a unit in a database is to attach representations of parts or properties of the unit to be formed to a common superordinate node. A particular unit (element) can then be referred to as, say, "the unit represented by node *i*," and a statement of the form "element x_i belongs to category *F*" can be reformulated as "the element represented by node *i* belongs to category *F*." Suppose the node is stored at a particular location in memory, say, address *i*. Then the unit can also be referred to as "the unit represented at address *i*," and a statement of the form "element x_i belongs to category *F*" can be reformulated as "the element represented at address *i* belongs to category *F*."

Another way of forming perceptual units has recently aroused considerable interest. It is based on the idea that temporal synchrony of neuronal discharges may be the glue for binding together representations of different parts or properties of an element (Milner, 1974). Cross-correlation studies suggest that cortical neurons in the visual system can synchronize their responses, depending on coherence of features in the visual field (see, e.g., Engel, König, Kreiter, & Singer, 1991; Gray, König, Engel, & Singer, 1989). Specifically, cells representing parts or properties of a given element may discharge in phase with each other and out of phase with cells representing parts or properties of other elements. If so, a particular perceptual unit can be referred to as "the unit represented by discharges in time slice *i*," and a statement of the form "element x_i belongs to category *F*" can be reformulated as "the element represented in time slice *i* belongs to category *F*."

Sufficient data for determining how perceptual units are represented and addressed in the visual system are not yet available. However, as indicated by the above considerations, it makes sense to assume that perceptual units, formed on the basis of precategorical, sensory information, can be addressed by later processes before any perceptual categorizations have been made. By the same token, despite the claims of van der Velde and van der Heijden (1993), it makes sense to suppose that sometimes only incorrect perceptual categorizations of an element are made.

Points 1 and 2. To explain the findings of Nissen (1985), I assumed that in Nissen's experiments, the probability p_L of making a correct perceptual categorization with respect to location ("above," "below," "left," or "right") was 1. Points 1 and 2 of the critique concern this assumption.

Point 1 is that p_L "does not equal 1 because of some (variable) external circumstance (e.g., because location categorization is very fast, as Bundesen suggests), but out of necessity in order for the analysis to hold. This means that there is a necessary one-to-one correspondence between an element and its location in Bundesen's theory: Whenever the location is known, so is the identity of the element, and vice versa. In other words, location and element are indistinguishable" (van der Velde & van der Heijden, 1993, p. 348).

It is true that I had to assume that $p_L = 1$ "in order for the analysis to hold," that is, in order to explain the main result of Nissen (1985; i.e., Equation 2 of van der Velde & van der Heijden, 1993). In general, however, p_L can be different from 1. According to my analysis, the value of p_L should depend on the experimental conditions, and the relationship found by Nissen (1985; Equation 2 of van der Velde & van der Heijden, 1993) should break down in conditions in which $p_L < 1$. Thus, if my reinterpretation of Nissen's findings is correct, the relationship found by Nissen should break down if the location discrimination required by the subjects were made difficult.[1]

Point 2 is that my analysis "can be seen as an attempt to explain a form of asymmetry in the data on the basis of a . . . symmetry. . . . Such an attempt is destined to fail because asymmetry cannot be explained solely on the basis of symmetry" (van der Velde & van der Heijden, 1993, p. 348).

Point 2 is surprising. To explain the asymmetry in Nissen's (1985) data, I explicitly hypothesized an asymmetry by assuming that in Nissen's experiments, p_L was 1. Indeed, I showed that the hypothesized asymmetry was sufficient to account for the asymmetry in the data by use of my theory of attention.

REFERENCES

Beck, J., Prazdny, K., & Rosenfeld, A. (1983). A theory of textural segmentation. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 1-38). New York: Academic Press.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, **97**, 523-547.

Bundesen, C. (1991). Visual selection of features and objects: Is location special? A reinterpretation of Nissen's (1985) findings. *Perception & Psychophysics*, **50**, 87-89.

Bundesen, C. (1992). Concept of visual sensation in a theory of visual attention: A theoretical note. *Perceptual & Motor Skills*, **74**, 874.

CORNSWEET, T. N. (1970). *Visual perception*. New York: Academic Press.

ENGEL, A. K., KÖNIG, P., KREITER, A. K., & SINGER, W. (1991). Interhemispheric synchronization of oscillatory neuronal responses in cat visual cortex. *Science*, **252**, 1177-1179.

GIBSON, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.

GRAY, C. M., KÖNIG, P., ENGEL, A. K., & SINGER, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, **338**, 334-337.

GREGORY, R. L. (1966). *Eye and brain*. London: Weidenfeld & Nicolson.

GREGORY, R. L. (1970). *The intelligent eye*. New York: McGraw-Hill.

JULESZ, B. (1981). A theory of preattentive texture discrimination based on first-order statistics of textons. *Biological Cybernetics*, **41**, 131-138.

KAHNEMAN, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

MARR, D. (1982). *Vision*. San Francisco: Freeman.

MILNER, P. M. (1974). A model for visual shape recognition. *Psychological Review*, **81**, 521-535.

NEISSER, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

NISSEN, M. J. (1985). Accessing features and objects: Is location special? In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance* (Vol. 11, pp. 205-219). Hillsdale, NJ: Erlbaum.

OLSON, R. K., & ATTNEAVE, F. (1970). What variables produce similarity grouping? *American Journal of Psychology*, **83**, 1-21.

PALMER, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, **9**, 441-474.

POMERANTZ, J. R. (1981). Perceptual organization in information processing. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 141-180). Hillsdale, NJ: Erlbaum.

VAN DER VELDE, F., & VAN DER HEIJDEN, A. H. C. (1993). An element in the visual field is just a conjunction of attributes: A critique of Bundesen (1991). *Perception & Psychophysics*, 53, 345-349.

WERTHEIMER, M. (1923). Untersuchungen zur Lehre von der Gestalt II. *Psychologische Forschung*, **4**, 301-350.

NOTE

1. In an instructive review of the first draft of Bundesen (1991), Mary Jo Nissen argued that this implication of my interpretation ought to be tested.

(Manuscript received September 14, 1992;
revision accepted for publication September 17, 1992.)