Intellectica, 1999/1, 28, pp. 37-51

Mary A. PETERSON*

Organization, Segregation and Object Recognition

The question of whether or not past experience exerts an influence on perception, per se, or only on the outputs of a universal perceptual process has yet to be answered definitively. In the first part of this paper, I will present the Gestalt answer to this question as it arose, that is, as a critique of the empiristic approach[1]. The Gestalt critique was so successful that its tenets serve as basic assumptions in modern vision research, reviewed in the second part of the paper. One of these assumptions is that depth segregation must be accomplished in whole or in part before memories of objects can be accessed (i.e., before past experience can exert any influence). In the third part of the paper, I will summarize evidence indicating that this "depth segregation-first" assumption is incorrect. Some of this evidence dates to the early days of the Gestalt school; other evidence is more recent. Finally, I will consider whether or not this counter-evidence calls for a redux of empiristic psychology.

Key words: memory, perception, vision, depth segregation, empiricism, Gestalt, object recognition, figure-ground segregation

Organisation, séparation en profondeur et reconnaissance d'objet. Il n'existe pas à l'heure actuelle de réponse définitive à la question de savoir si l'expérience passée exerce une influence sur la perception en elle-même, ou si l'expérience ne fait que moduler les sorties d'un processus perceptif universel. La première partie de cet article présente la réponse gestaltiste à cette question telle qu'elle s'est élaborée à l'époque, c'est-à-dire comme une critique de l'approche empirique. La critique gestaltiste fut couronnée de succès au point que ses principes servent d'hypothèses de base dans les recherches contemporaines sur la vision, recensées dans la deuxième partie de cet article. L'une de ces hypothèses est qu'une séparation en profondeur doit être effectuée en tout ou en partie avant de pouvoir accéder aux souvenirs des objets (c.a.d. avant que l'expérience passée puisse exercer une quelconque influence). La troisième partie de cet article résume un certain nombre d'indications expérimentales qui tendent à montrer que cette hypothèse n'est pas correcte. Certaines de ces expériences datent des premiers travaux de l'école gestaltiste ; d'autres sont plus récentes. En conclusion, la question de savoir si ces expériences contradictoires justifient une reprise de la psychologie empiriste est discutée.

Mots-clés : mémoire, perception, vision, profondeur, empirisme, Gestalt, reconnaissance d'objet, séparation figure-fond.

* Department of Psychology - University of Arizona

[1] Following a distinction made by Koehler (1947), the term "empirist" will be used to refer to "...a psychologist who tends to explain a maximum of mental facts by previous learning." Koehler reserved the term "empiricist" for "...a philosopher who claims that all knowledge grows from outside experience."

I. The Empiristic Background and the Gestalt Critique

The empiristic psychologists (e.g., Mueller, Wundt) believed that independent points of retinal stimulation were united into the objects we perceive by the influence of past experience, acting through the laws of association. In other words, they believed that pointillistic sensations were organized into separate objects because they had been organized as separate objects in previous experience.

In criticism of the empiristic tradition, the Gestalt psychologists posed the following question: If the visual array were indeed pointillistic and unorganized prior to the operation of past experience (i.e., before access to, and outputs from, memory), then how could the appropriate memories be chosen from among a plenitude of memories? In order to limit which memories might be accessed by the visual array, the Gestaltists argued that some organization must first be imposed on the stimulation before memories of past experiences are accessed. According to the Gestalt psychologists, organization was composed of grouping and segregation processes (Koehler 1929/1947; Koffka 1935). This paper is concerned with segregation and its relationship to a particular form of past experience — object recognition. Accordingly, from here on in, the paper will focus on segregation; there will be little discussion of grouping.

Segregation. The Gestalt psychologists concentrated on figure-ground segregation — the differentiation of the visual field into figures and grounds. When two adjacent regions share a common border in a momentary view, the border between the two regions is often (although not always; see Kennedy 1973) assigned to one of the two regions. The region to which the border is assigned is the figure, the other region is the ground (Rubin 1915/1958). Two consequences follow from the assignment of the border to the figure and not to the ground:

(1) the figure appears to have a shape, whereas the ground, lacking a contour, appears shapeless, at least near the contours it shares with the figure;

(2) the ground appears to continue behind the figure; hence, the figure appears to be located in front of the ground (Rubin 1915/1958)[2].

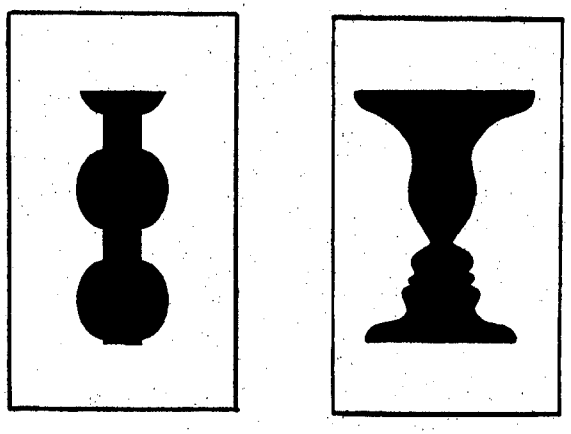

Figure 1.
Two figure-ground displays.
A. is a novel display. B. is Rubin's classic vase-faces display.

Both of these consequences can be seen clearly in Figures 1A and 1B. Figure 1A is an illustration of unambiguous display, which is typically organized as a black figure lying on a white ground. Figure 1B is Rubin's well-known ambiguous vase-faces display, where the black and white regions alternately appear to be the figure. In other words, figure-ground organization alternates, or reverses in Figure 1B, and the black and white regions are alternately seen to be shaped and unshaped, in front and behind, when they are seen as figures and grounds, respectively. Notice that the vase can be seen only when the black region appears to be figure, and the face profiles can be seen only when the white region appears to be figure. The fact that the vase and the faces are seen consciously in Figure 1B only when the region portraying them is seen as figure was taken as evidence

[2] Given the fact that the figure appears to lie in front of the ground, figure-ground differentiation is clearly a form of depth-segregation. The Gestalt psychologists intended their work on figure-ground differentiation using 2-D displays to generalize to the everyday 3-D world. More recently, some investigators have reserved the term "figure-ground segregation" for discussions of two-dimensional (2-D) displays, and have used the term "depth segregation" for discussions of three-dimensional (3-D) displays. In this paper, the general term "depth segregation" will often be used for 2-D as well as 3-D displays.

for the depth segregation-first hypothesis by many investigators (e.g., Hoffman and Richards 1985; Koehler 1929/1947; Rock 1975; Wallach 1949). But the Rubin vase-faces stimulus reveals only that apparent figural status and conscious recognition are coupled, not that one precedes the other.

Demonstrations like the one in Figure 1A were also taken incorrectly to indicate that depth segregation necessarily occurs before access to memory representations. This is because depth segregation occurs readily in Figure 1A, even though the center black region is novel. For novel displays, memories of specific objects clearly do not aid segregation; some other factors must be operating. Rubin, and psychologists after him, identified a number of low level factors that contribute to segregation (e.g., Harrower 1936; for review, see Hochberg 1971; Pomerantz and Kubovy 1986). As illustrated in Figure 1A, regions that are smaller in area, convex, reflectionally symmetric, closed, and surrounded are likely to be seen as figures; their adjacent, larger, enclosing regions are likely to be seen as grounds. All of these factors can potentially be assessed without access to representations of previously seen objects; thus, they form a means by which depth segregation can be accomplished without the aid of object memories. Of course, the fact that segregation can occur on the basis of low level cues without any influences from object memories does not imply that segregation always occurs without any influence from object memories. The latter claim requires systematic exploration of whether or not memory plays any role in the presence of other depth-relevant cues (e.g., the Gestalt cues and the monocular and binocular depth cues).

Despite the flawed arguments supporting the depth segregation-first assumption, the idea that depth segregation always occurs without any influence from past experience was readily accepted in the 1930's and continues to be accepted by most current vision scientists and perception theorists today.

II. Modern Theories of Vision Grounded on Gestalt Assumptions

Marr's (1982) general theory of vision is an example of a modern theory of vision that is grounded on the depth segregation-first assumption. Marr conceived of vision as organized into four sequential stages of processing. The first stage entails the construction of the primal sketch in which edges are made explicit; the second stage entails the construction of the $2^{1/2}$-D sketch in which surfaces and viewer-relative orientations emerge; the third stage is the construction of the 3-D model; and the fourth stage entails access to 3-D object models in memory. In Marr's view, these

stages are organized in a serial, hierarchical, manner: later stages are initiated after earlier stages are completed. Thus, in Marr's theory there is a clear sequence: depth-segregation first, access to object models (or representations) in memory later.

Marr was led to a serial hierarchical scheme in part by neuropsychological data showing that certain brain damaged patients can segregate depth accurately, and can correctly describe the shape of the figure in figure-ground displays; nevertheless, these patients are unable to recognize objects as familiar, to produce names for them, or to mime their function (Warrington 1982). For these patients, depth segregation is clearly possible when object recognition is not. In Marr's theoretical framework, these patients were understood to have preserved function of lower level processes (e.g., depth segregation) and deficits in higher level processes. Like the Gestaltists, Marr took evidence that depth segregation can be accomplished successfully even when object recognition cannot to imply that depth segregation is always accomplished without inputs from object recognition processes. (More recent hierarchical interactive processing models maintain the hierarchical arrangement between depth segregation and object recognition, but hold that depth segregation need only be attempted, and not necessarily accomplished, before object memories are accessed (e.g., McClelland 1985; McClelland and Rumelhart 1986; Vecera and O'Reilly 1998).

III. Evidence inconsistent with the depth-segregation-first assumption.

Despite the strong arguments made for the depth segregation-first assumption (e.g., Gottschaldt 1926, 1929; Koehler 1947; Marr 1982; Palmer and Rock 1994; Rock 1975; Vecera and O'Reilly, in press; Wallach 1949), a strand of evidence contradicting that assumption can be traced from Sander and Rubin to the present time. A brief review of the contradictory evidence follows.

Rubin (1915/1958). A little known fact is that Rubin (1915/1958) devoted a substantial part of his monograph on figure-ground organization to a finding that was inconsistent with the Gestalt idea that segregation must be completed before past experience could influence perception. Rubin reported careful experimental work indicating that when observers viewed a figure-ground display for a second time, they tended to perceive the same regions as figures as they had when they viewed it the first time; Rubin called this effect a figural after-effect. Thus, Rubin's monograph contained experimental evidence that some type of past experience can influence depth segregation. These data were consistent with Sander's proposals.

Because of the Gestalt coda that segregation must precede the influence of past experience, it became necessary for Gestalt psychologists to show that Rubin's figural after-effects were an artifact of experimental conditions. Consequently, a number of investigators attempted and failed to replicate Rubin's findings. Most of these investigators used experimental designs that differed substantially from Rubin's (e.g., Cornwell 1963; Dutton and Traill 1933; Gottschaldt 1926, 1929; Rock and Kremen 1957). Some investigators succeeded in replicating Rubin's figural after-effects (Gottschaldt 1929; Vetter 1965). However, much more attention was paid to the failures to replicate than to the successes. One reason is that most investigators accepted an argument made by Gottschaldt (1929) that the effects of past experience should be robust to changes in experimental design. This argument rendered the failed replications using procedures that were different from Rubin's more compelling than the successful replications using Rubin's own procedure. It should be noted that Gottschaldt's argument depends upon an implicit assumption that if past experience does influence segregation, it must dominate all other factors. But this assumption is neither necessary nor correct, a point to which I will return below.

Schafer and Murphy (1943). The proposal that past experience played a role in perceptual organization was taken up by the New Look psychologists who argued that personality and social factors such as motivation and need were effective determinants of perceptual organization (e.g., see Bruner and Goodman 1947; Murphy 1947; Postman, Bruner, and McGinnies 1948). By definition, motivation and need arose from previous experience.

Among the experiments carried out by the New Look psychologists, was one conducted on segregation by Schafer and Murphy (1943). Schafer and Murphy used stimuli such as the one shown in Figure 2A in which two adjacent regions shared a common border. Each of the regions on either side of the central border of Figure 2A depicts a distorted profile of a face, as can be seen clearly in Figures 2B and 2C. Thus, if figure-ground differentiation occurs at the central border of Figure 2A, only the face depicted by the region seen as figure should be perceived.

Before testing for figure-ground differentiation with the whole stimulus, Schafer and Murphy (1943) presented each half of the stimulus alone (as in Figures 2B and 2C) for many trials. On these trials, observers learned names for each of the two face profiles, and they were rewarded with a small sum of money when viewing one of the two faces, and penalized by a small sum of money when viewing the other face. In the test phase of their experiment, Schafer and Murphy (1943) showed observers Figure 2A, and asked them to

name the face they saw in the conjoined stimulus. They named the face that had been rewarded during training significantly more often than the face that had been penalized during training.

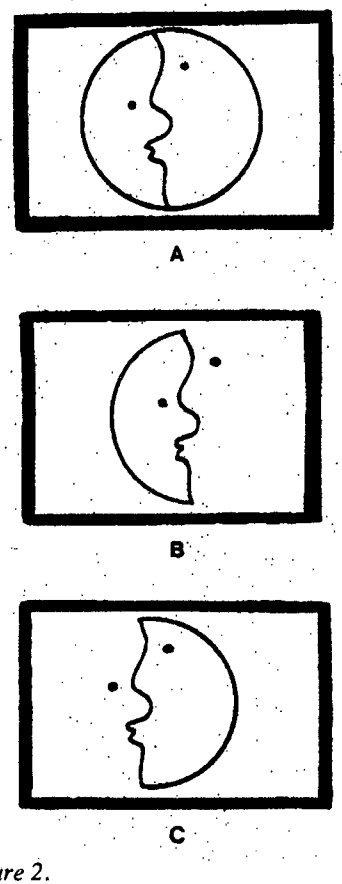

Figure 2.
Schafer and Murphy's display.
The full display is shown in A.
The two halves are shown separately in B and C.

If the observers were reporting their first perceived depth organization, Schafer and Murphy's results could be taken to imply that past experience can influence depth segregation, and hence, to cast doubt on the depth segregation-first assumption. However, Smith and Hochberg (1954) demonstrated that the exposure

durations Schafer and Murphy used at test were long enough to permit reversals of figure-ground organization. Therefore, it was argued that Schafer and Murphy's observers may not have been reporting the first perceived figure-ground organization. Instead, they may have had time to see both alternatives and to report whichever organization they pleased. Under such conditions, it would be plausible to argue that observers preferentially reported the rewarded alternative over the punished alternative. Because the depth segregation-first assumption was generally accepted, and because much of the other experimental work reported by the New Look psychologists was discredited (e.g., Pastore 1949), no one attempted to replicate Schafer and Murphy's experiment under conditions in which observers indicated their first perceived figure-ground organization.

Peterson et al, 1991-1994. The evidence I have taken to be inconsistent with the depth segregation-first assumption was obtained in a series of experiments using stimuli comprised of two adjacent regions sharing a common border, such as those in Figures 3 and 4. In these stimuli, the common border was designed to denote a known object along one side (the "high denotative" side) and a novel, or meaningless, object along the other side ("low denotative" side). The term "denotative" is used because unless a region appears to be figure, the object it denotes is not seen. (i.e., a region depicts an object only when that region appears to be figure; a point that is well demonstrated by the Rubin vase-faces stimulus in Figure 1B).

The basic design of the research in my laboratory involved showing observers such stimuli in both an upright and an inverted orientation, where upright was defined as the orientation in which the object on the high denotative side of the common border appeared in its typical orientation. It is known that when objects that have a typical upright are viewed in an inverted orientation, it takes longer to identify them than when they are viewed in an upright orientation (Jolicoeur 1985, 1988; Tarr and Pinker 1989).

Consequently, it is thought that changing the orientation of such stimuli from upright to inverted lengthens the duration of time before outputs from object representations are available to influence other processes. We reasoned that if outputs from object representations contribute to depth segregation, then their effects should be more evident for upright than for inverted displays. In particular, we proposed that if outputs from object representations favored the interpretation that the figure lies on the high denotative side of the shared contour, they would be revealed by a greater tendency to see the high denotative region as figure in upright stimuli rather than in inverted stimuli.

To test this hypothesis, we asked observers to report about which region appeared to be figure when viewing both upright and inverted versions of stimuli in which the border common to two adjacent regions was high in denotivity along only one side. The stimuli were designed such that a change in orientation from upright to inverted (and vice versa) left unchanged other cues known to be relevant to figure-ground and/or depth segregation (e.g., reflectional symmetry, relative convexity, relative area, interposition, binocular disparity).

Across a number of different experiments using both long (30 sec) and short (28-100 msec) exposures and measuring both reports about reversals of figure-ground relations and about the first perceived figure-ground organization, results indicated that the common border between two adjacent regions was more likely to be assigned to the high denotative side when the stimuli were upright rather than inverted (i.e., the figure appeared to lie on the high denotative side of the contour, whereas the low denotative side appeared to be ground).

Consider Figure 3, which is one of the stimuli used in the first demonstration of this effect (Peterson et al 1991). Figure 3 is a center-surround stimulus. The Gestalt configural cues of horizontal symmetry, enclosure, and smallness of relative area favor the interpretation that the black center region is the figure. Yet the border common to the black and white regions is high in denotivity on the white surround side (denoting a standing woman), and low in denotivity on the black center side. Thus, any object recognition influences on depth segregation were expected to favor the interpretation that the figure lay on the white (surround) side of the contour. Those influences were expected to be larger for upright than for inverted stimuli.

Observers viewed this figure (and another) for a number of 30-sec trials and reported about perceived figure-ground organization as it alternated during each trial. Two results were critical: First, once observers saw the white high denotative region as figure in upright displays, they maintained that organization significantly longer (before organization reversed such that the black low denotative center region appeared to be figure) than they did in inverted displays (13.8 sec vs. 7.2 sec). This result indicated that the high denotative side was maintained as figure longer when the stimuli were upright rather than inverted; but did not necessarily imply that any object recognition processes influenced the likelihood of obtaining the high denotative side of the contour as figure. The second result did indicate that the high denotative side was more likely to be obtained as figure: observers maintained the black low

denotative center region as figure for significantly shorter durations in upright than in inverted displays (6.8 sec vs. 18.8 sec). Thus, observers were more likely to reverse out of seeing the black low denotative center region as figure (and into seeing the white high denotative surround as figure) when viewing upright stimuli compared to inverted stimuli.

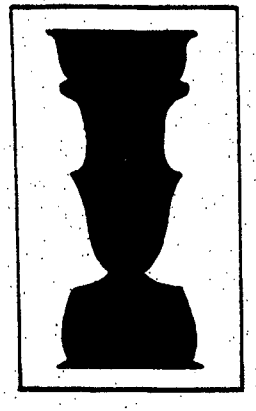

Figure 3.
The upright version of one of the stimuli used by Peterson et al (1991).

We took the increased likelihood of obtaining the high denotative region as figure in the upright orientation compared to the inverted orientation to indicate that outputs from object representations contribute to depth segregation. This is because the only variable known to change when the orientation of these displays was changed between upright and inverted was the quickness with which outputs from object representations matched by the high denotative side of the contour would be available to influence other processes, such as depth segregation.

Control experiments ruled out factors such as eye movements, motivation, or responses to demand characteristics as explanations for these effects. Other control experiments showed that these results did not depend on observers' knowledge about what object was depicted, or on the conscious recognition of the object depicted on the high denotative side of the contour. Even when observers knew that the surround in inverted displays denoted an inverted woman, and recognized it as such when the surround appeared to be figure,