What is the Origin of the Gestalt Principles?*

Dejan Todorovic**
dtodorov@f.bg.ac.rs

ABSTRACT

Gestalt principles (regularities of figure–ground articulation and grouping laws) account for the organization of the visual field, but what is the origin of those principles themselves? Three answers to the question where the Gestalt principles come from are discussed. The first answer, that one should not look for explanations of phenomenal facts outside of phenomenal facts, is criticized. Two other, non–exclusive answers, namely that past perceptual experience and the structure of the visual nervous system may underlie the Gestalt principles, are elaborated. Arguments of the classical Gestalt authors against the relevance of these factors are examined. It is suggested that the biological importance of the Gestalt principles is that they may function as heuristic cues for the presence of physical objects.

An image can be described as a spatial distribution of tiny colored dots. This is literally true for displays on computer screens: such an image with, say, a million dots, is physically completely described when the color of each dot at each position is determined; this requires five pieces of information for each dot, two for its spatial co–ordinates and three for the specification of its color. However, our conscious perception of such an image is not equivalent to a union of one million spatially distributed punctuate color sensations. One difference is that we are aware not only of local features, such as position and color, but also of properties of more extended regions of the image. For example, when a set of, say, 100 000 contiguous red dots is surrounded by a set of 900 000 yellow dots, the red portion is seen as a unit, a ‘figure’, and the yellow surround as another unit, the ‘ground’, the two units being separated by a contour.

* This research was funded by the Serbian Ministry of Science and Education, Grant No. OI179033.
** Laboratory of Experimental Psychology, Department of Psychology, University of Belgrade.
of some shape. Although there are countless ways in which that set of million dots could be thought of as being partitioned into two or more subsets, in this case it is only the articulation into these particular two sets of 100 000 and 900 000 dots which is actually seen. In our conscious apprehension of the stimulus, the dots within each of these two portions are perceptually integrated into a whole, and each of the two wholes is perceptually segregated from the rest of the field.

Another supra-local phenomenal aspect of the visual field is the fact that several figures, say a set of small shapes arranged along a circle, can be perceived as belonging together as constituents of a group, a hierarchically higher–order perceptual unit (in this case the circle). Conversely, an individual figure, say the letter ‘T’, can be perceived as being partitioned or subdivided into natural parts, that is, hierarchically lower–order units (in this case the vertical line and the horizontal line, which can itself be seen as subdivided into two halves, joined at the point where the horizontal touches the vertical). Grouping refers to the way a hierarchy is built up starting from lowest–order constituents, whereas partitioning refers to how it is broken down, starting from the highest–order whole.

Figure and ground, integration and segregation, groups and parts, articulation and hierarchy, are all aspects of the phenomenal organization of the visual field, which are not explicitly contained in the point–by–point description of the stimulus input. These notions apply both for 2D images, whether or not they are physically made up of tiny dots, as well as for the perception our 3D environment. The question that arises is how to account for this organization. Given a concrete visual field, how can we predict which particular portions, out of the huge number of conceivable alternatives, will be perceived to belong together as visual units or figures, which figures will be perceived as forming groups, which portions of a figure will be perceived as its natural parts, etc? Such questions were initially posed and also answered, at least in part, in two seminal publications early in the 20th century. One was a book by Edgar Rubin (1915/1921), dealing with figure–ground articulation, and the other was a paper by Max Wertheimer (1923), introducing the Gestalt laws of grouping. These issues were further discussed and developed by other classical Gestalt psychologists (Köhler, 1947; Koffka, 1935; Metzger, 1936/2008, 1966, 1975a, 1975b) as well as other researchers; for a more recent account see Palmer (2003), and for brief reviews see Peterson & Salvagio (2010) and Todorović (2008). In this research a number of stimulus conditions were iden-
tified that are conducive to figure–ground articulation (such as that those portions will tend to be seen as figures which are small, surrounded, convex etc) and grouping (such as proximity, similarity, closure etc). I will refer to these two related issues (features of figure–ground articulation and grouping laws) together as to the 'Gestalt principles'. They belong to the best known contributions of Gestalt psychology, and are reported in most contemporary perception textbooks, and also in the perception chapters of many textbooks of general psychology.

However, despite of their relative prominence, there are a number of basic aspects of these notions that are in need of more study. For example, there is no definitive ‘official’ list of the Gestalt principles (but see Metzger, 1966), nor are there clear rules that would predict what happens when more than one grouping law applies in a given display but different laws favor different organizations, although such situations have been considered early on (Wertheimer, 1923). However, the issue that I will be dealing with here is the question where the Gestalt principles come from. They account for the organization of the visual field - but how are they themselves to be accounted for? Why do we see figures on grounds? Why is it that when some elements of the visual field comply with the grouping laws, we tend to see them as belonging to perceptual wholes? For example, according to the law of proximity, we tend to see elements that are near each other as belonging together in a group - but why is that the case? What are the principles behind the principles? Such questions were occasionally discussed in the classical Gestalt literature, though not in much detail.

One potential account of their origin is that all Gestalt principles are special instances of an overarching general rule, the principle of Prägnanz or Good Gestalt. This rule states that the phenomenal organization that is actually perceived is singled out from all other possible organizations in that it is as ‘pregnant’ or good as stimulus conditions permit; here ‘maximal possible goodness’ carries the connotations of simplicity, unification, regularity, balance, orderliness and the like. However, one problem with this account is that although the notion of goodness or simplicity makes sense intuitively, it is not easy to apply it generally to concrete cases, or to compare levels of simplicity or goodness of different configurations. In this sense the particular Gestalt principles, such as proximity, continuity, closure etc, are notions that are better defined and easier applicable than their supposed generalization. For example, given a configuration such as ●● ●● ●●, the principle of proximity correctly predicts that its
six elements will be perceived as being subdivided into three pairs rather than, say, into a pair and a quadruple, or into two triples etc. But why would three pairs be simpler, better, more regular etc than two triples? One could argue that the stimulus conditions are such that in this case maximal goodness is manifested in such manner that organization according to proximity is the best possible; however, it is not clear that such a formulation is a conceptual advance over and above just stating that the organization is in accord with proximity. On the other hand, if Prägnanz is accepted as a valid generalization, there still remains the problem where it itself comes from? Why would phenomenal organization tend to be maximally good? Is the Prägnanz principle like an axiom, so that it simply has to be acknowledged as a foundational rule that needs no further justification? Or is it more like a theorem, to be derived from some more basic principles? I will argue below that the Gestalt principles have two roots, one based on the learned structure of the external physical world, and the other on the innate structure of the internal neural system.

Closely related to the Prägnanz principle are formulations in terms economy of visual processing, maximal simplicity, minimal complexity, least information load etc. An advanced formalization of these notions is provided by the structural information theory (van der Helm, 2007). However, although this approach has dealt with several issues in visual perception studied by the Gestalt psychologists, it has not yet been directly applied to grouping laws or figure-ground appearance.

Here I will discuss two types of accounts of origins of Gestalt principles, involving past experience and neural structure. Although neither of them was deemed particularly satisfactory by the classical authors, I will argue that they are relevant, non-exclusive origin candidates. However, I will first discuss the idea that attempting to seek a non-phenomenological account of these principles is fundamentally ill-advised.

Gestalt Principles and Experimental Phenomenology

According to the approach of experimental phenomenology «phenomenal facts have to be explained only with other phenomenal facts» (Vicario, 1993, p. 209). Therefore, if we try to explain phenomenal facts through non-phenomenal means, we fall in the error called «violation of the rules of categorical analysis» by Lorenz (1973), that is, the explanation of facts at a certain level of complexity (e.g., mental facts) with facts at a lower level of complexity
(e.g., physical and physiological facts) (Vicario, 1993, p. 201). From such a viewpoint, searching for the origins of Gestalt principles in terms of past experience or neurophysiology, as sketched below, is an enterprise doomed from the start because it is inappropriate, for general methodological reasons.

However, there are several difficulties for such an attitude. For example, it is not clear why one should be bound by Lorenz’s ‘rules of categorical analysis’, which forbid between–levels of explanations. To the contrary, in the history of science facts on one level have repeatedly been explained by facts at another level, such as biological facts that were explained by chemistry, or chemical facts that were explained by physics. Thus there is no compelling reason to base the decision whether an across–level account is feasible or not feasible on a general rule that simply excludes any explanations of this type as reductionist; rather such questions should be decided on a case by case basis, with independent arguments. Note also that the classical Gestalt authors have not been bound by such a rule in their theorizing. As noted below, Wertheimer himself introduced a grouping rule based on past experience; incidentally, he also favored a neurally based account of apparent motion. Furthermore, both Köhler (e.g. 1938, 1947) and Koffka (1935) discussed the neural counterparts of consciousness and the notion of their isomorphism. Also, Metzger (1936/2006) wrote that «we will have to understand the nature of this domain [cerebral cortex] if we want to have any prospect of approaching the laws of vision from the outside by the physiological route» (p. 191). Finally, it would be interesting to learn how the question of the origin of Gestalt principles would be treated within experimental phenomenology. This seems to be a legitimate scientific question, but it is not clear how it would be answered using only phenomenological means, because it asks about the reasons for the very existence of phenomenological facts. The answers sketched below in terms of past experience and neural structure may, of course, be wrong, but at least they are genuine attempts to answer the origin question.

The Role of Past Perceptual Experience

Classical Gestalt authors had a rather negative attitude with respect to explanations based on past experience and learning (see Köhler, 1947). Metzger (1966, p. 741) even claimed that such accounts are not regular scientific assumptions but belong more properly to the realm of the psychology of prejudice or group pressure! This attitude was a reaction to the predominant role
that introspectionists and behaviorists attributed to learning for accounting for most psychological functions. Nevertheless, the notion that some aspects of the organization of the visual field could be acquired through learning was not foreign to classical authors. Wertheimer (1923) acknowledged the existence of a Gestalt grouping law based on past experience or habit, manifested in reading (p. 331). For example, he noted that we are likely to perceptually subdivide a sequence of characters such as ‘314cm’ into two parts, ‘314’ and ‘cm’, rather than, say, into ‘31’ and ‘4cm’, presumably because we have learned to differentiate letters and numbers, and how they are usually combined. Similarly, when reading script we are likely to perceive a continuous string such as ‘mum’ as consisting of three ‘natural’ parts, ‘m’, ‘u’, and ‘m’, rather than break it down into some other constituents, because of our acquired knowledge of the forms of the letters of the Roman alphabet (see also related examples in Todorović, 2008). However, Wertheimer (1923) contended that this experience principle was just one among several Gestalt grouping laws, and not a particularly strong one. He argued for this by constructing a number of ingenious displays in which our experience with letters is put into competition with other Gestalt principles, such as continuity and closure. What one predominantly sees in such configurations are closed and continuous but unfamiliar forms, rather than the otherwise very familiar letters from which they are composed (p. 334–335).

Nevertheless, it can be reasonably argued that not just one but several other Gestalt principles may have been acquired or at least affected by experience to some extent (Brunswik & Kamiya, 1953; Rock, 1975). To motivate this claim, let me start with some observations involving the comparison of the standard Gestalt displays with situations in everyday perception. The visual demonstrations provided by Rubin, Wertheimer and others were typically very simple, schematic drawings, such as sequences of filled and unfilled circles, straight and curved line strokes, basic geometric figures, and the like. However, with respect to the organization of the visual field, such displays share a number of features with more complex configurations, likely to be encountered in usual environments.

To illustrate this, consider first some features of a simple figure–ground display, such as in the example in the first paragraph of this text. Rubin (1915/1921) noted some conditions under which the percept of a figure on a ground arises. For example, if a small portion of the visual field is fully surrounded by a larger portion, the small one will tend to be seen as figure and the
What is the Origin of the Gestalt Principles?

large one as ground. He also described several differential features of figures and grounds. For example, the figure appears as solid and thing–like and it attracts attention, whereas the ground appears as less definite and more like amorphous ‘stuff’; although both are flat, the figure appears as if located more in front, whereas the ground appears as if lying more to the back; there is a sense of the ground continuing behind the figure, unseen but amodally present; the contour that separates the figure and the ground is perceived to belong to the figure and to give it shape, whereas it does not belong to or shape the ground.

Note that all these features are closely analogous to certain physical and visual properties of real 3D objects on real backgrounds, for example an apple lying on a table. When we look at the apple, it covers a relatively small extent of our visual field, and is fully surrounded by the rest of it; we attend to the apple and see it in sharper focus than the table and the remainder of the visual field; the apple is physically closer to us and the table is further away; the table is visually partly occluded by the apple but physically continues behind it; finally, the projected contour of the apple delimits its own shape and has nothing to do with the shape of the table.

Consider now the Gestalt grouping laws of proximity / similarity / closure / common fate, which claim that elements that are near to each other/ are similar to each other/ form closed contours / move together, tend to be perceived as a group or unit. Note that real objects tend to exhibit analogous features. The surfaces of objects are rarely completely uniform but often involve some patterns or textures, whose elements are generally near each other and tend to be mutually similar, or at least more near and similar to each other than to the textures of other, surrounding objects; the projected outer contours of physical objects are generally closed; finally, when objects move, their surface elements all move in a related manner.

Still another Gestalt principle is the principle of good continuation, according to which elements that form smooth continuations tend to be perceptually grouped together as units. For example, in a configuration shaped like the letter ‘T’ we tend to perceive the two adjoining horizontal line segments as belonging together and thus being grouped into a whole, the top horizontal line (rather than, say, grouping one horizontal line segment with the vertical line, as in a Greek letter ‘T’ configuration). Such features of simple displays have counterparts in frequent everyday situations in which an object partly occludes another, located further away from the observer. This circumstance often gene-
rates a ‘T–junction’, in which the boundary of the nearer object (corresponding to the horizontal line in the ‘T’) interrupts the boundary of the one further away (corresponding to the vertical line). The point of this example is that what appears as a unit in a simple display (the horizontal line of the letter T) corresponds to a unit in a more complex everyday display (boundary of nearer object). As another example, a configuration in the form of an ‘X’ tends to be seen as two crossing straight lines rather than two touching ‘V’ forms. The preferred percept is in accord with situations involving a thin elongated object in front and the boundary of another object (or another thin object with different slant) in the back, whereas the non-preferred percept would correspond to the rare situation of precise alignment of two V–formed shapes.

These examples should suffice to show that there is a rather close correspondence between some features of real 3D objects and analogous features of simple 2D displays often used to illustrate the Gestalt principles. What is the explanation of this correspondence? A pre-established harmony between the functioning of the visual system and the physical structure of the external world? Purely innate mechanisms that ensure veridical experience? Perhaps a more promising approach is to assume that the Gestalt principles may have been, at least in part, affected by experience with the corresponding properties of objects in environmental scenes.

Early on children are exposed to various static and dynamic scenes and observe the environment from different viewpoints when carried around, and later through own locomotion and manipulation of graspable objects. Based on such constant exposure to the structure of the physical world surrounding them, coupled with active exploration, they are in a position to discover various general features of objects, such as that objects tend to have smooth outlines, that they tend to move as wholes, that behind them there are generally no holes but other objects further away and only accidentally occluded from view, whose shape has nothing to do with the shape of the occluding objects, etc. Such and related properties of the external world could in this way be internalized through prolonged experience. According to this idea, the reason for the correspondence noted above is, at least in part, that we apply statistical regularities acquired in everyday life when viewing simple displays. For example, the reason that when we look at a display such as a red patch on a yellow background (as described in the first paragraph above) we have the impression that the figure is in front and owns the border and that the background continues behind the figure, is, in part, due to its resemblance, in basic features, to real
objects on real backgrounds, about which we have learned through overwhelming experience that they are indeed as a rule positioned in front of the background and own their border.

It is important to note that the claim that learning and past experience may affect perception can have two distinctly different meanings, which can be labeled as recognition-based and feature-based. A popular meaning, but not one I have in mind here, refers to familiarity and recognition of certain objects or their classes. This recognition-based meaning of experience is exemplified in Wertheimer’s grouping law based on experience, which refers to our acquaintance with particular letters or classes of letters. The possibility that figure–ground articulation can be based on this type of experience with objects was also acknowledged by Koffka (1935, p. 210), but no empirical data were available at that time. Such data were later provided in studies such as by Peterson and collaborators involving displays with two adjoining homogeneous regions, one of which resembles a familiar object (such as a silhouette of a human, animal or a thing) and the other which is an unfamiliar, abstract shape. In such cases the former region is generally perceived as figure and the latter as ground (Peterson & Skow–Grant, 2003). However, this is not the type of effect of learning that I am discussing here. As Köhler (1947) noted, if perception of wholes would depend only on such experience then «specific entities would be segregated in the field only to the extent to which they represent known objects» (p. 82). Similarly, Koffka (1935) observed that «patterns for which we had no experience should be absolutely ambiguous with regard to the figure–ground articulation» (p. 209).

My stress here is on an effect of past experience which is not recognition-based. Rather, the proposal sketched above is feature-based, that is, it attempts to relate the effectiveness of figure–ground articulation and grouping laws to experience with certain general perceptual features of objects as visual entities. Thus regardless of category (whether an object is an organism or an inanimate thing of a particular recognizable kind), and familiarity (whether it is well known or completely novel), an object in our world will tend to subtend a relatively small portion of the visual field, have closed and continuous outlines, move as a whole, exhibit surface patterns or textures with micro–components that are near and similar to each other, etc.

In the preceding considerations I have tried to make plausible the claim that Gestalt principles may in part be based on learning. Whether this claim is true, however, is an empirical question concerning which there are not many data,
and the existing data are not unanimous. On the one hand, Quinn, Burke & Rush (1993) showed that 3 month old infants were able to use lightness similarity as a grouping principle, and Quinn, Bhatt & Hayden (2008) found that 3–4 months old infants can use the principle of proximity. Such findings indicate rather early manifestation of some Gestalt principles, although they do not necessarily rule out the possibility of previous learning. On the other hand, Spelke et al (1993) found that although adults could reliably use the principles of color and texture similarity, good continuation and good form, 5–month old and 9–month old infants could use them only weakly, and 3–month old infants could not use them at all, and Quinn et al (2002) found that although 6–7 month old infants could use the principle of form similarity, 3–4 month old infants could not. Such data indicate gradual acquisition of some Gestalt principles which may be based on learning, although it may also depend on the maturation level of the visual system. I will continue the discussion of the role of past experience after first considering the potential neuro–physiological foundations of the Gestalt principles.

The Role of the Visual System

Experience with features of objects, as sketched above, may be a condition for the effectiveness of Gestalt principles, but such an account is silent with respect to the implementation of such principles in the brain. The classical Gestalt authors, although acknowledging the role of neurophysiology in general, felt that the contemporary knowledge of the structure and function of the nervous system was of little help to understand the perceptual issues they were concerned with. Whereas nerves were thought of as machine–like conductors imposing rigid structure, they favored domains that allow free, dynamic interactions of self–organizing forces. Thus Köhler (1971) wondered «how can the segregation of visual objects as circumscribed units be brought about by action currents in connecting neurons?» (p. 255), and preferred accounts in terms of current flow through cortical tissues. Koffka (1935) discussed the effects of «forces in the physiological field» (p. 117) on perceptual organization, but did not specify their neural correlates in any detail. Metzger (1936/2006) wrote that «with our perceptual theory we do not bow to physiology but present challenges to it» (p. 197). He not only felt that the physiology of his time was unable to meet such challenges, but also that it has «again and again obstructed and diverted the discovery and recognition of the actual laws of seeing» (p.
What is the Origin of the Gestalt Principles?

188); he even suggested that it may be more fruitful to regard the brain not as a set of nerves but as a system of finely distributed fluids, because some features of such a system would be analogous to some organizational aspects of the visual field.

However, since those days our knowledge of the actual structure and functioning of the visual system has achieved tremendous advances. A general overview of the relation of various Gestalt notions and nervous activity was provided by Spillmann (2009). Here I will only briefly sketch some aspects that are directly relevant for Gestalt principles. For example, there are data concerning the possible neural basis of figure–ground articulation and the phenomenon that the contour belongs to the figure rather than to the ground: some cortical visual neurons respond differently to the same local luminance step, depending on whether the figure is located at one side or at the other side of the step (Zhang & von der Heydt, 2010). Other studies are relevant for the law of good continuation: they suggest that its effectiveness may be based on long–range interconnections between cortical neurons responsive to lines and edges of similar orientations (Hess & Field, 1999).

Furthermore, one can speculate that the principle of proximity may be related to reactions of neurons with different receptive field sizes. Recall the example presented earlier, in which six dots are perceived as sub–grouped into three pairs of dots. The perception of six individual dots may be subserved by neurons with relatively small receptive fields, which would be able to resolve the individual dots. On the other hand, neurons with large receptive fields may not be able to resolve the two nearby dots within a pair and therefore would react to the pair in a similar way as to a single elongated object, but may well be able resolve two neighboring pairs because of the greater distance between them. Thus these neurons may signal the presence of three rather than six units, which may be the neural basis of perceiving three groups in this stimulus.

The similarity principle may have some physiological basis as well. For example, according to this principle the sequence of twelve elements such as \|\| | | |/// is perceived as being divided into three sub–wholes consisting of four identical elements each. The neural basis of this perceptual achievement may be the fact that sets of neurons tuned to different orientations would preferentially respond to elements within each group, and that neurons within each set would mutually facilitate each other through short– and long–range interconnections (Ko et al., 2011). In slightly more complex examples of simi-
licity principles at work, such as ●●●●●, several sets of neurons might respond better to the elements in the first group, whereas other sets of neurons would respond better to the elements in the second group; in this particular case, one set of neurons could be the off–center neurons and edge detectors and the other set could be the on–center neurons and line detectors. In general, one could posit that for any case of perceptual similarity/difference there might be a corresponding similarity/difference in the pattern of neural reactions. The simultaneous firing and mutual interconnections between neurons of the same class could provide the neural basis for perception of belonging and grouping.

Discussion and Criticisms

The considerations in the previous two sections are not backed by much experimentation and amount more to suggestions and speculations. Concerning the role of past experience, little is still known to what extent young organisms actually pick up those features of objects that are critical for the Gestalt principles, and utilize them in visual field organization. Concerning the role of neurophysiology, although there are some relevant and promising data, currently there is still not enough support to decide on the details of any of the postulated relations of Gestalt principles and neural activity. Furthermore, there is the difficulty of how exactly to conceive of the relation between perceptions and corresponding neural states; however, this so–called «hard problem of consciousness» is a very general, partly philosophical topic, not confined to the issues dealt with here. Nevertheless, these two avenues of research seem to me well worth exploring. In the following I will indicate how this approach may meet some criticisms that were already voiced, in one form or another, by the founders of the Gestalt movement.

The classical Gestalt authors did consider claims for the role of learning but, except for Wertheimer’s experience principle, they generally rejected them. For example, Metzger (1936/2006) wrote that «the fundamental laws of perception were present before ... experience», that they «are not fundamentally changed by experience», and that without them and their stability «the store of past experience could neither be collected nor utilized» (p. 180). Similarly, Koffka (1935) claimed that «experience with things and figures can be had only after things or figures have been established as parts of the behavioral environment» (p. 210). Köhler (1947) wrote that «sensory organization
appears as a primary fact which arises from the elementary dynamics of the nervous system» (p. 118). From this perspective, one could argue that rather than the Gestalt principles being based on experience, it is experience itself that is based on these principles.

In reply one can argue as follows. Certain basic aspects of some Gestalt principles could indeed be provided by Köhler’s «elementary dynamics» of innate neurophysiological structure, such as sketched above in the section on the visual system, thus making it possible to start the processes of visual field organization going. Furthermore, some aspects of the functioning of the visual system could be based on the experience, not of the individuals but of the species, in the form of evolutionary forces sculpting the structure of the nervous apparatus. Nevertheless, it is not necessary to suppose that all Gestalt principles are present in full form right from the beginning. Rather, their fine-tuning and breadth of application could proceed through a phase involving learning early in life, as sketched above.

Recall that the rejection of any substantial role of past experience was thought to be supported by Wertheimer’s demonstrations, noted above, that our knowledge of letters can easily be overcome by other grouping laws, such as continuity and closure. However, this does not necessarily mean that past experience is generally a poor determinant of visual organization, but rather that experience with particular classes of letters (recognition-based experience) is less frequent and therefore has a weaker effect on perceptual organization than experience with features of objects such as continuity or closure, to which we are constantly exposed whenever we open our eyes (feature-based experience).

Metzger claimed that the following example argues against the idea that the principle of common fate derives from experience: «If three flies sit still on a window pane and three others crawl around on it, the three that are moving seem to belong together, even if they are moving in different directions. For this reason it is wrong to believe that the law of common fate involves only a matter of memory of the known behavior of solid bodies» (1936/2006, p. 35; 1975b, p. 93); Köhler (1947) used the same example. However, this example of common fate could still be based on experience. Note that it is not true that points on moving bodies must necessarily move in the same direction. This only applies for translations but not for rotations. Complex motions of rigid bodies, such as that of a falling leaf or a hurled rock, can be analyzed into translational and rotational components, which can change their translational direc-
tion and rotational center from moment to moment. Furthermore, motions of bodies with flexible connections, such as a human walking or a bird in flight, involve combinations of different motions of several component parts. These examples show that in daily life we have been exposed to many cases of complex patterns of synchronous element motions other than pure parallel translations, induced by motions of unitary objects.

Role of Gestalt Principles

One of the key contributions of the Gestaltists was the insight that the phenomenal fact that we see objects as segregated wholes is not a simple consequence of the physical fact that objects are segregated wholes. Namely, their property of being physical units, that is, relatively coherent chunks of shaped matter that usually move more or less independently of other objects and have characteristic surface features, is not inherited by their optical projection upon the retina, which is just a bundle of rays striking a collection of quasi-punctate sensors. Köhler (1947) pointed out that «each element of a physical surface reflects light independently; and in this respect two elements of the surface of an object, such as, for instance, a sheep, are no more related to each other than one of them is to a surface element in the animal’s environment. Thus in the reflected light no trace is left of the units which actually exist in the physical world ... so far as retinal stimulation is concerned, there is no organization, no segregation of specific units or groups» (p. 95). The conclusion from this is that the fact that our phenomenal visual world does exhibit organization is something that needs to be explained. Accounting for the form of this organization is a job for the Gestalt principles.

One reason that the classical Gestaltists questioned explanations based on experience with real objects was the existence of discrepancies between physical objects and perceived wholes (Köhler, 1947, p. 83; Koffka, 1935, p. 77). One instance of perceived wholes not corresponding to real objects but being grouped in accord with the proximity principle is the example provided above, in which we perceive each of the three pairs of dots as a unit although they are not real unitary objects, but simply pairs of dots that happen to be located near each other; another instance is the case of star constellations, which are seen as grouped although the individual stars have nothing to do with each other physically and may be located at hugely different radial distances from the observer. Also, in the above example of the similarity principle, we experience the four
What is the Origin of the Gestalt Principles?

disks and the four squares as being grouped, although they are just independent marks on paper. However, although these are good demonstrations of grouping principles at work, they are not necessarily representative. Köhler (1947) pointed out that «it is quite true that organization often forms continuous wholes and groups of separate members where no corresponding physical units exist. But when contrasted with the large number of cases in which organization gives a picture of objective facts, this disadvantage will rightly be regarded as negligible» (p. 96). Generally, phenomenal organization «tends to have results which agree with the physical world ... in other words, that “belonging together” in sensory experience tends to agree with “being a unit” in the physical sense» (p. 95).

There is a certain analogy between these considerations concerning the role of the Gestalt principles in the organization of the visual field and the role of depth cues in the perception of the third dimension. Namely, similarly to the case of perceptual organization, the phenomenal fact that we see the world in 3D is not a simple consequence of the physical fact that the world indeed has three dimensions. The reason is that in the projection of the world upon the retina the third dimension is lost, and thus our perception of depth is something that needs to be accounted for. This is the job of depth cues, such as ocular convergence, disparate images, perspective, occlusion, shading etc. Generally, these depth indicators are consistent with each other and enable reliable depth perception in everyday situations. However, they are not fail–proof indicators of real depth, but only a set of heuristics that exploit some of the features usually accompanying real depth. Thus in some situations they may be inconsistent with each other and also lead to discrepancies between perceived depth and physical depth, for example when they trigger depth percepts in flat images.

Analogously to convergence, disparity, occlusion etc being cues to depth, one can think of proximity, similarity, closure etc as being cues for “object-ness”. Objects of many different kinds are always present in our environment, and being able to register and differentiate them efficiently is obviously of great biological importance. The problem is that in the projection upon the retina their separateness and independence is lost, and the corresponding optical field is a single plenum, that is, a wide–angle spatially continuous two–dimensional array of light. Nevertheless, usually a portion of the field that exhibits a characteristic texture, has closed smooth contours, moves as a whole etc., corresponds to a physical object. Therefore, it is very useful for the visual sys-
tem to incorporate heuristics, such as the various Gestalt principles, in order to detect objects. These principles help ascertain which parts of the optical array correspond to different physical objects. Generally there is a remarkable correspondence of physical units and perceived units, in that a segment of the visual field that belongs together perceptually has a counterpart in a portion of the outer world that “hangs together” physically. However, such heuristics for identifying objects are not fail–proof: from this functionalistic perspective, the object–detecting system can err in two ways, like any signal detection device. One error is a “false alarm” type of failure: this happens when a portion of the visual field is perceived to belong together, although it does not correspond to a physical object. As noted above, a bunch of dots drawn near each other on a piece of paper are perceived to belong together, although there is nothing physical that unites them; however, a similar bunch that constitutes the surface texture of an object would correctly signal its presence to perceivers whose perceptual mechanisms incorporate the principle of proximity. The other error is a ‘miss’ type of failure: it happens in cases of camouflage, in which objects are physically present but are not perceived to belong together, and thus do not exist as phenomenal units. In such cases the Gestalt principles favor a visual field organization that does not segregate object correctly and precludes their recognition, which, of course, is the purpose of camouflage. Note that these two types of failures do not support the claim that the Gestalt principles are unaffected by experience, but rather that they are not perfect deterministic indicators of presence of objects, but fallible probabilistic cues.

In sum, although the question of origin of Gestalt principles is currently more a matter of speculation than empirical tests, it can be argued that both learning and neural action may form their basis, and that their purpose in vision is to help detect environmental objects.

REFERENCES


What is the Origin of the Gestalt Principles?


Peterson, M. A. & Salvagio, E. (2010), Figure–ground perception. *Scholarpedia, 5*(4): 4320. http://www.scholarpedia.org/article/Figure–ground_perception


