Overview of the History of Visual Theory

Speaker: Lothar Spillmann **October 6, 2009**

To be able to navigate, find food and find mates, the subjective visual world must correspond to the physical world. Yet, from the introductory lecture (Sept. 22.2009) we remember that any correspondence is limited because of the properties of the visual sense. The following examples illustrate where the correspondence breaks down:

1. We sometimes perceive something for which there is no direct correlate in the physical stimulus (e.g. illusory contours).
2. We sometimes do not see objects, although they are clearly present before our eyes (e.g. hidden objects).
3. We sometimes see objects, but differently from the way they are (e.g. geometric-optical illusions). Check the illusion battery by Michael Bach and Viperlib on the web.

These examples show that *naïve realism* in the strict sense (i.e., the perceived world is identical to the physical world) cannot be correct. However, despite such limitations, we get around in the world sufficiently well, at least most of the time. Metzger's book shows where our vision is fooled (e.g. camouflage), but also where it is creative in the sense of restoring incomplete information in the stimulus (e.g. filling-in, completion); or by providing us with *perceptual constancies* that render a perceived object invariant despite changes in illumination, distance, and viewpoint (e.g. brightness, size, and shape constancies). Without such invariances we would be unable to form experiences. The neuronal mechanisms underlying such accomplishments are therefore assumed to be innate.

The fundamental question to be asked is how are visual sensations (e.g. brightness, color, movement) tied together to represent an object? How is the information that impinges onto the retinal receptor mosaic converted to a coherent percept (i.e., unity)? In Gestalt terminology: What becomes figure, what ground? Furthermore, which stimuli are grouped together to become objects in perception? This latter question is called the *binding problem*.

During the last 250 years, there have been three major approaches to the question of perceptual unity.

(1) English Empiricism

The English philosophers Locke, Berkeley and Hume stated that the mind is a blank

slate (Latin: tabula rasa) and that all information is imprinted on it comparable to writing on a blackboard. Binding was assumed to occur by association as follows:

Impressions that often occur together tend to be associated with one another. Such association leads to ever-larger assemblies and finally percepts. The basis of form perception during development thus was *spatio-temporal contiguity*; shapes were learnt from combining individual elements according to the frequency with which they occurred together. The psychologist Donald Hebb (McGill U, Montreal) took this approach to formulate a neurophysiological hypothesis according to which synapses are formed between cells activated together (look up the Hebb rules on the Internet). If "A" typically occurred at the same time and place as "B", but not "C", the linkage between the two excitations would become reinforced and evolve into an assembly. Metzger asserts that there are so many percepts for which we could not possibly have formed associations that the empirical approach to visual perception cannot be right. Yet, we experience these percepts effortlessly, even if we see them for the first time.

(2) Judgment Theory

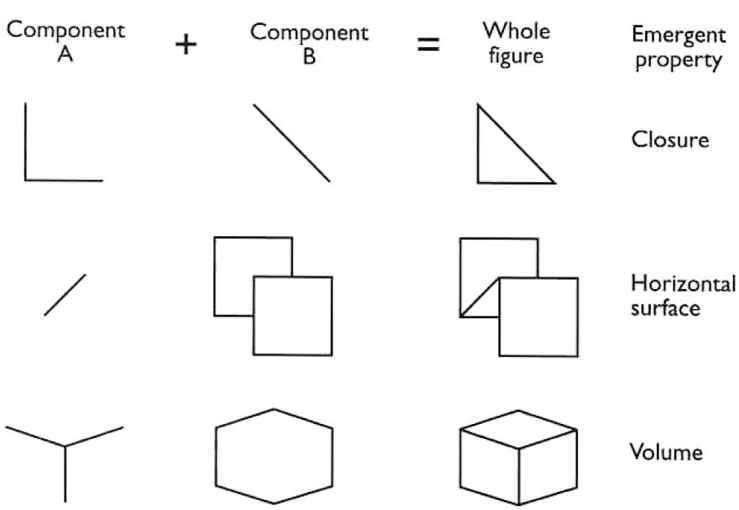
The physicist Hermann von Helmholz suggested that we have percepts by virtue of unconscious inferences. He claimed that in many instances we do not actually see objects; we only "think" them or make inferences about them. According to Gregory who adopted Helmholtz's view, perception is *problem-solving*. When confronted with incomplete object representations, we make hypotheses and postulates about them to reconcile the stimulus with our expectation. Take the Kanizsa triangle. By seeing an illusory triangle partially occluding the three "pacmen", we make sense of an unlikely stimulus pattern (the three bites) in the most plausible manner. Thus, the illusory percept constitutes the solution to a problem.

(3) Gestalt Theory

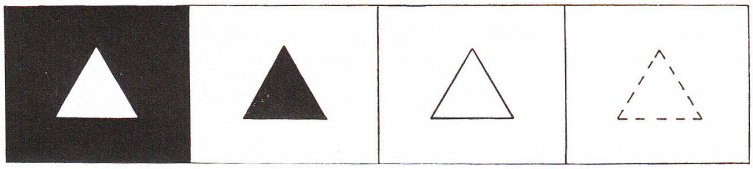
Max Wertheimer, Wolfgang Koehler und Kurt Koffka proposed that perceptual unity originates in the observer as an emergent property; it is created due to factors inherent in the brain. These factors are called *Gestalt factors* and comprise the factors of good continuation, closure, proximity, symmetry, similarity, and common fate.

According to this view a stimulus becomes perceptually organized in the simplest, most symmetrical and balanced manner, consistent with the principle of *Prägnanz* or good Gestalt. Gestalt factors are descriptive, or phenomenological, categories and are difficult to quantify. Yet, they are self-evident, persuasive and allow predictions. Also, they are present in early infancy and in the animal kingdom and

cannot be overridden by better knowledge.



The figure on the left demonstrates how something new (the whole) emerges from its parts. This is called *super-summativity* (from Palmer).



The figure below shows that the percept of a triangle is independent of the way it has been drawn: This is called *transposition*.

Super-summativity and transposition define a Gestalt.

Metzger says that we find the visual world ready-made before our eyes without our doing. For example, we can distinguish a trapezoid presented in the frontal plane from a square presented on a tilted plane, and a small nearby square from a large distant square, although the stimuli projected onto the retina in each case are the same. Associationism alone cannot explain this; depth cues need to be taken into account.

Despite long and hard battles between the proponents of the above three approaches, the truth probably lies in the middle. Developmental studies have shown that we are born with a certain "brainware" and need to fine-tune the software in an ongoing "discourse" with the environment, to acquire the exquisite functions needed for perception. For this we need an inventory of appropriate stimuli at the right time. This is called the *critical period*, within which the visual system has plasticity for such adjustments. This window usually is open only in childhood. Examples (see the previous lecture on Sept. 29): A squinter will irreversibly lose the ability to see stereo-depth, if his/her strabismus (the deviant optical axes) is not corrected before the age of – say – 4 years. An astigmatic person will have sharp vision only for a narrow range of line orientations, if he/she is not optically corrected

early in life. People born with a milky cornea or lens will not have acute form vision, unless they are operated shortly after birth. A corneal or lens transplant later in life will not return their vision (however, see the recent paper by Held on Molyneux's question). These observations point to a fine interplay between innate visual functions and learning (in the sense of Hebb).

The neurophysiological mechanisms underlying these developmental deficits have to a large extent been clarified by studies in cat and monkey and are well understood. In contrast, the neuronal processes and mechanisms responsible for figure-ground organization are still mysterious. Neurophysiologists have suggested that a potential explanation of Gestalt vision is *binding by synchrony*. (Read the article by Wolf Singer on Scholarpedia.) For example, neurons stimulated by collinear stimuli (factor of good continuation) or coherently moving stimuli (factor of common fate) will tend to "fire" precisely at the same time (in the millisecond range). More work is needed.