

Original Contributions - Originalbeiträge

Shelia Guberman

Gestalt road to Necker cube perception

1. Introduction

The Necker cube is a 2D drawing, which we perceive as a 3D body that can be seen in two different spatial positions. It was first described by Swiss naturalist, L. A. Necker (Necker, 1832). This effect of switching is generally illustrated with the image of an isometric skeleton cube rather than Necker's original rhomboid. Since 1832, interest in the phenomenon of switching the percept during observation of the Necker cube has not faded.

According to Necker, the switch resulted from “an involuntary change in the adjustment of the eye for obtaining distinct vision.” Wheatstone (1838) claimed that the switch depended on mental activity. Until now debate centered on this topic remains very much alive (Pritchard, Heron, & Hebb, 1960; Kaufman & Richards, 1969; Gregory, 1970; Kawabata, Yamagami, & Noaki, 1978; Einhäuser, Martin, & König, 2004; Turvey, 2018). The questions are: do eye movements relate to perceptual switching or not, do eye movements precede the switch, or they are a consequence of the perceptual switch?

This paper was inspired by Feldman's works on organization of eye movements (Feldman, 2016), and Guberman's work on recognition of geometric figures (Guberman, 1980). Let us clarify the problem. We want to understand the mechanism of the switching phenomenon. The act of switching consists of three steps: (1) recognizing the image as the Necker cube in one of the two possible positions, (2) seeking for another interpretation of the same stimulus, and (3) recognizing the stimulus as the Necker cube in the second of the two possible positions. It is obvious that all three stages of switching deal with recognition of the Necker cube. If we understand what will happen in each of these three steps, we will be able to understand the role of eye movements in the Necker cube phenomenon.

There are no publications about the mechanism of Necker cube recognition in a 2D drawing. There are computer programs for polyhedron recognition in 2D drawings, but these are all based on point-to-point comparison of the net of vertexes in the presented image, and the net of vertexes for each type of polyhedrons

stored in the memory. It is clear that this is not a humanlike approach, and can't help in human perception research.

2. Recognition of 2D Geometric Figures

Before discussing the Necker cube recognition, let us discuss the existing models of perceiving 2D geometrical figures. The most powerful part of psychology of perception that deals with such kind of objects is Gestalt psychology. In 1923, Wertheimer presented a theory of perception in the domain of 2D geometrical figures (Wertheimer, 1923). He defined the notion of *Prägnanz* as the main characteristic of Gestalt: *short and meaningful description*, i.e., he moved the notion of Gestalt from the visual to the linguistic domain, which is natural: all examples of Gestalten in the literature are descriptions produced by human beings as a response to the question “What do you see?”.

Wertheimer defined the content of the description as well. He did it when explaining the “good continuation” principle: “one knows what a ‘good’ continuation is: **in designing a pattern one has a feeling how successive parts should follow one another**” (Wertheimer, 1923). It means that interpreting the linear drawing is redrawing the pattern the same way as you suppose it was drawn by the author, or, in other words, imitating the process of drawing. It is worth mentioning that the idea of looking on the line as the result of drawing is quite familiar to geometry. In Wikipedia the strict definition of the line has a remark: “The above definition in many ways allows us to convey our intuitive idea of the line as something drawn without retouching the pencil from the paper.”

As a result, the notion of Gestalt can be defined as follows: *the Gestalt is the description of how the object was created*. It proclaims the basic principle of Gestalt theory (restricted to domain of geometric figures)—**the imitation principle: perceiving the visual stimulus is describing the creation of the image**. It was shown that the imitation principle overarches all other Gestalt principles and gives the right answer to all examples of potential ambiguity in the perception of geometrical figures presented by Wertheimer (1923), Köhler (1975), Koffka (1935), Metzger (2006), and many followers up until the present day (Guberman, 2015, 2017).

In the domain of linear drawings, the minimal element of the drawing is the **stroke** (the segment of a line drawn without stops and interruptions). In geometrical figures, the objects are open or closed chains of a linear segment. In Figure 1A (Wertheimer's example), the redrawing of the image can be started from any end of any stroke. Starting with the rightmost corner one will inevitably trace all four sides of the quadrangle (four strokes) without jumping to any crossing line. The rest is the second quadrangle. So, following the

imitation principle gives only one Gestalt: two quadrangles. Nobody perceives bad Gestalten, and, consequently, don't need the mystical force that moves us in the sea of bad Gestalten to the "attractor." The small triangles that can be found between the two drawn polygons were never **drawn**—and, therefore, *never perceived*. In Figure 1C, the digit "4" is easily perceived: it consists of three strokes in particular arrangement, but despite the fact that an identical "4" can be cut out from Figure 1B, it is never perceived because it was not **drawn**, i.e., there are not three strokes in particular arrangement that denote the digit "4." In Figure 1D (Metzger, 2006), the "mantle" in the form of a bread roll and a 'core' in the form of an almond" were never **drawn**. It was shown that the imitation principle overarches all other Gestalt principles (Guberman, 2015, 2017)—the goal declared by Wertheimer in his 1923 paper. Our understanding of how the image was drawn dismisses all wrong perceptions (the bad Gestalten).

A couple of new Gestalt principles were presented not long ago demonstrating the power of the imitation principle (Pina, 2010). Figure 2 describes the corresponding images.

In each case the description (the Gestalt) tells us how the **object**, which is presented at the image, was created. It is important to note that in this case the perception is different from perception of the geometric figures when we describe how the **drawing** was created.

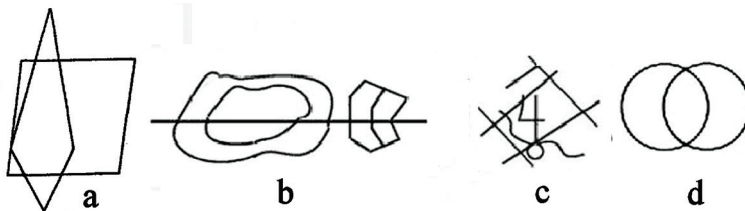


Fig. 1. The power of the imitation principle: a) when redrawing the image one guided by the "good continuation" principle follows each stroke without interruption and get only one percept; b) following the strokes one never draw the digit "4"; c) following the strokes one will inevitably draws the "4"; d) following the strokes one perceived only the good Gestalt – "two overlapping circles".

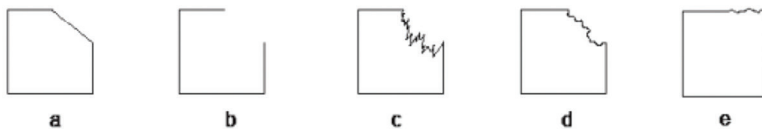


Fig. 2. (A and B)—a square perimeter whose top right corner is absent, missing, deleted, or cut; (C)—the glass square is broken in a corner; (D)—a gnawed or nibbled square; and (E)—the square is deformed by a scorch.

It was also shown that the imitation principle is applicable to not only linear drawings but also all kinds of communication signals (including handwriting and speech).

Eighty years after Wertheimer's paper was published, in which he referred to imitation as the fundamental principle of geometrical drawings' perception, the neural mechanism of the imitation principle was discovered (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992).

3. Recognition of the Prism

The nature of drawings of 3D geometric bodies is completely different. The 2D geometric figures are created from lines, which we see in the drawing. These drawn figures don't present something else—they themselves are the objects. When one is looking for the Gestalts of 2D geometrical figures, and asking the question: "How they were created?", the answer is: "Drawn by the pen." Unlike this, the 3D geometric bodies are not created from the drawn lines. If one wants to extend the Gestalt principle of imitation to a 3D geometrical body, one has to answer the question: *How was this body created?*

Let us start with a simple body—the prism. The question "how prisms are made" sounds a bit silly, but let us recall that this approach allowed the resolution of quite difficult problems pertaining to the recognition of other communicative signals—handwriting and speech recognition: the handwritten text is created by the movement of the hand, the 2D geometric figures are created by the moving pen, and speech sounds are created by the movement of articulatory organs. Quite surprisingly, the science of geometry contains the definition of a prism as a description of how prisms can be created: when a polygon is *translated* along a straight line (rigidly moved without rotation) it covers a volume of a prism. On the sketch of the prism, the ribs are traces of the movement of the polygon's vertices. Figure 3A presents the graphical description of the prism (not the view!), and at the same time it is identical to the drawing of the Necker cube—as can be ascertained based on the wire construction presented in Figure 3D. To present the view of the solid prism, the faces have to lose their transparency, and as a

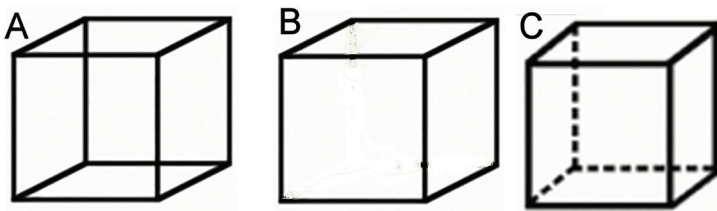


Fig. 3. Figures A, B, C.

result, some edges will disappear from the sketch. As a result appears the most adequate presentation of a prism (Figure 3B), where the existing but not visible edges are represented by dotted lines.

On this base was developed a computer program that recognizes prisms (Guberman, 1980). The algorithm is as follows: (1) find a polygon (the base), (2) find a set of parallel lines connected to the polygon vertexes (ribs), (3) check that no ribs cross the polygon (that means that the front face is opaque), and (4) check that the areas between the parallel lines are quadrangles (in other words—that the ends are connected by a chain of segments). This algorithm can also be used for recognizing pyramids. Only one minor change has to be made: lines originating from the vertices of the base should not be parallel, but should converge at one point. The algorithm also works for truncated prisms and pyramids.

When the experimenter wants to force the subject to perceive the particular view of the Necker cube, he presents a distorted image. At the image, the observer is forced to choose the desired front face, and thus the need to search for different candidates is avoided; and subsequently, the observer's line of vision follows the outgoing edges—precisely the two first steps of the algorithm described above. The last step (connection of the edges) is now unavoidable.

But how do people really draw a cube? The instruction can be found on the Internet (Figure 4). The sequence in the instruction is the same as that in the algorithm described above. This is one of the many illustrations demonstrating the role that drawings play in serving as instructions for creating 2D representations of 3D objects. The development of drawing skills in children demonstrates the evolution that takes place in the description of the object, in that there is a transition in this description from how the object was created to how it looks.

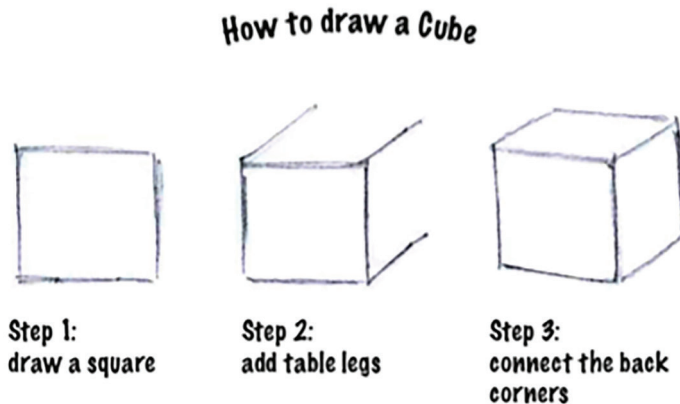


Fig. 4. Internet instruction for drawing a cube.

There is at least one person in the world who discovered the algorithm of cube recognition described above by introspection. Here is a quote from the book “What Makes a Cube a Cube?” by W. F. Ruth: “Participants “constructed” the particular cube or cubes that they see. move the eye along the edge leading from the “front” vertex toward the back left—then the cube-seen-from-below becomes instantly apparent. If the eyes do not make these movements, then the cubes do not appear and the lines remain on a flat surface” (Roth, 2011). It is amazing that this observation was made only by means of introspection.

4. Experimental Support 1

Do experiments support this model of perceiving the Necker cube? There is a small number of experiments that investigate the position of gaze fixations during observation of the Necker cube and switching of the percept. All data show that switches in perception of the Necker cube occur when the fixation point is located near the “Y” vertex. Figure 5 shows the results of three experiments (Ellis & Stark, 1978; Kawabata et al., 1978; Einhäuser, 1988) that show the position of the fixation points at the moment of switching. Keeping in mind the model presented above, we may now proceed to ask ourselves what pattern of fixation points needs to be expected. After identifying the drawing of the Necker’s cube as a 3D polyhedron, a polygon has to be found. It will be one of the quadrangles existing in the drawing. After that, the existence of parallel edges connected to the found quadrangles must be confirmed. After this the new percept is obtained. So, here are the predictions of this model:

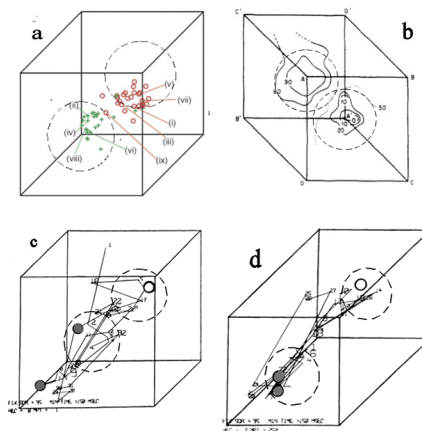


Fig. 5. Sequences of fixation points for a subject’s scans on two different cubes. The dotted circles show the central area of the front and back faces of the cubes. (A)—The small circles mark the type I events: the perception of the upper-right ambiguous corner was shifted from external to internal; the crosses mark the type II events: the shift for the lower-left ambiguous corner. (B)—Rates of fixation points’ occupation time while the type I event occurred; (C, D)—The big open circles mark the type I events, while the divided circles show the type II events.

- (1) The switching points occupy the central part of one of the quadrangles (but “not the exact center of the boundary” as it was found in the research dedicated to spontaneous fixation tendencies for visual forms (Kaufman & Richards, 1969)).
- (2) The switching point is located in the center of the quadrangle, which, after the switch, becomes the back face of the Necker’s cube.

One can see that these predictions are also in agreement with the experimental data—the switching points of gaze are located in the central parts of the squares. So, both theories satisfy the experimental data; however, the situation changes when the experiments are slightly changed. In all experiments, the stimulus is seen as a cube (to the degree that our vision can judge) except one case (Figure 5D). A series of “cubes” with increasing lengths (as a matter of fact prisms) was used in the experiments. The extreme prism is shown in Figure 5D. In this case, the switching fixation points are located quite far from the ambiguous “Y” corners, but still very close to the centers of quadrangles. This contradicts the hypothesis that fixation of the gaze at the “Y” corner moves this corner to the front, and supports our model that stated that the switching fixations are located in the central area of the quadrangle, which became the back face in the process of building the Necker’s cube.

Now one can see that negligence, in terms of usage of scientific terms, which took place two centuries ago (the substitution of the term “rhomboid” by the term “cube”) obscured the real situation for such a long time.

5. Experimental Support 2

Pritchard, in his research produced in collaborations with colleagues as well as his individually conducted research (Pritchard 1960, 1961), ran a number of experiments on perception of 2D and 3D geometric figures using the technique of stabilizing an image on the retina; and here are the results in the authors’ own words:

- (1) After a few seconds of viewing the stabilized image the subject discovers that he can still transfer his **attention** from point to point over the limited visual field.
- (2) It was found that the image of a simple figure, such as a single line, vanishes rapidly and then reappears as a complete image. A more complex target may similarly disappear and reappear as a whole; on the other hand, it may vanish in fragments, with one or more of its parts fading independently.
- (3) The complex figure invariably fades and regenerates in meaningful units. The several lines of a triangle, square, etc., act independently, with the exception that the activity of parallel lines in a figure is correlated (Figure 6).

- (4) The influence that the parallel lines exert on one another is evident (see Figure 6). This parallel-line effect is most striking with the Necker cube (Figure 7); in addition, when the cube is seen in three dimensions (as demonstrated in Figure 6, this is always the case with the stabilized image), surfaces that are separate but in parallel planes frequently act together.

The front and back of the cube may remain in sight, for example, while the other edges (the lines connecting the squares that constitute the front and back faces) have disappeared.

- (5) Figure 8, during fading, demonstrates the possible groupings of the parts of the image: H, P, and B.
- (6) Figure 9, however, shows that the effect can occur with a figure lacking both meaning and goodness of form: the closed loop also tends to act as a whole though it is quite irregular.
- (7) No support is provided for the idea that an angle or an intersection of lines (the "Y") is a perceptual element.
- (8) Figure 9 was divided in two parts. The division never happens inside the stroke.
- (9) During the fading of Figure 10, the subject sees new words made up of letters and parts of letters in the original.

These results are extremely important and support the authors' opinion that they allow renewal of the theory of perception. It is a big surprise that there were no followers of this work for 60 years.

How do these experimental results correlate with the model of perceiving the visual communicative signals (drawings and handwriting)?

1. The algorithm of recognizing polygons in the linear drawing stated the following: Start from any end point of any linear segment, follow the segment until the other end point, and continue with tracking the next segment. If the end of the current segment coincides with the starting point (i.e., the chain of segments is closed), the polygon is identified. In Figure 1, we see all consecutive stages of this process for a triangle and quadrangle (from right to left). The algorithm will present Figure 3 as consisting of two parts: the chain of linear segments and a closed curved line. Experiments correspond to the consequence of the imitation principle: when dividing a drawing into parts, not a single stroke can be cut.
2. The algorithm of recognizing text (cursive or printed) transforms the trajectory into a sequence of predefined elements, and after this the sequence is divided into parts, each of which is a letter (sometimes variants can appear)

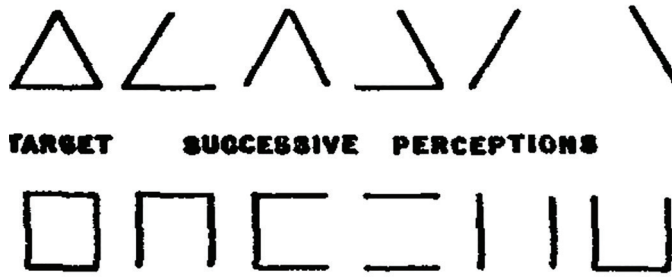


Fig. 6. Sequences of its parts that appear and disappear in the vision after a figure (triangle or quadrangle) were stabilized on the retina.



Fig. 7. Afterimages from stabilizing on the retina the image of the wired cube.



Fig. 8. This sign stabilized on the retina triggers during fading the letters H, B and P.

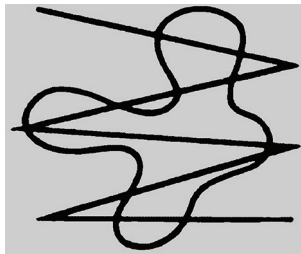


Fig. 9. This sign stabilized on the retina triggers during fading two parts – the curve and the zigzag.

PEER PEEP BEE BE

Fig. 10. The stabilized image (BEEP) and its afterimages: not a single stroke is broken and every piece is meaningful.

(Guberman, 1976). Figure 3 will generate the sequence {vertical line, dash, vertical line, upper ark, lower arc}. Different divisions of this sequence generate the characters H, P, and B, but every division gives superfluous elements, and so, a meaningful word can't be recognized. This result repeats the results obtained with the stabilized Figure 3 (see point 9 above). Additionally, in point 10, one can see that the operation of searching different divisions of the sequence of elements for meaningful interpretations of the handwriting is echoing the results presented in the fading of Figure 5.

3. The algorithm of cube recognition starts with finding a square (that will be defined as the front face). Any square can be chosen as the front face (see Figure 2). The second step is to find the parallel ribs connected to this face. In the experiments described above it was emphasized that parallel lines were always noticed by the eye (see point 5). The last step is defining the back face and concisely sending the signal that the particular view on the cube is perceived.

Pritchard's results impart some confidence that the imitation principle underlying the perception of geometric figures represents some real processes in the brain. Still, the fundamental question of "how we follow the lines in the image" is not answered. At the same time, one can't neglect the conclusion that follows from Pritchard's experiments with the Necker cube stabilized on the retina: the stabilized Necker phenomenon exists even when the gaze is fixed. Does it dismiss the results discussed above? No, because the relations between eye movement and sites of gaze fixation are real. It only means that the brain has to generate mechanisms to deal with the problem, and this is not an exception in neurophysiology and psychology. One very close example is handwriting recognition. The history of medicine knows a number of cases wherein the patient lost the ability to read handwriting (although, for the most part, the patient's vision remained largely intact). Psychologists found that the lost ability can be compensated by different means: (1) the patient can read the text by following the trajectory of the writing with the finger (and even with the nose); or (2) if this method also failed to manifest results, the patient nonetheless was able to read the text if somebody else followed the trajectory of the script using a pen.

6. Summary

The proposed model answers the question raised by Turvey (2018): how we select one of two justified orientations. It is defined by choosing the square that will be the front face of the cube. It is worth mentioning that this was the fuzzy conjecture of Gregory: "The change of appearance cannot be due to bottom-up processing. It must be set downwards by the prevailing perceptual hypothesis of what is near and what is far" (Gregory, 1970). The model also stated that the fixation point, in which the switch of percept occurs, is not the cause of the switch—it

is only a signal to other parts of the brain that the procedure of recognizing the percept of the cube is successfully finished.

The cube can be recognized by starting with any other quadrangle. However, we prefer the square because the right angle exercises a remarkable effect in drawing our attention. Wertheimer found that the right angle is exceptional because it has *prägnanz* (which means that its description is short and meaningful) (Wertheimer, 1923). The peculiarity of a rectangle is also due the fact that it indicates that it belongs to the surface, which is normal to the line of view and is not distorted; and so, it reflects the real shape of one of the faces.

We still didn't know the details of what happens between the switches. How does our brain follow the strokes without moving the gaze along the line? This question is highly relevant and vital for the processing of 2D linear images and handwriting as well. It is obvious that this mechanism is part of the mirror neurons phenomena but significantly differs from instances investigated until now: the imitation of gestures occurs synchronically with the gestures themselves, while perception of linear drawing demands animation of the dead images.

This result was achieved within the framework of the imitation principle of Gestalt theory (restricted to the perception of communicative signals). It is difficult to accept that we decipher handwriting and geometrical sketches by means of our hand, speech by means of our mouth, music by means of our vocal fold, and dances by means of our legs. But we don't have to be upset: our older solid brothers—the physicists—with great difficulty accepted that matter and energy can't be cut infinitely, that an electron can behave as a wave and as a particle, that the time is relative, and mass can be transformed into energy.

Acknowledgments

The author thanks Prof. A. Feldman and Prof. M. Latash for devoting their attention to the problems discussed in this study, and participating in the discussions of the results. The author acknowledges the Chief Editors of GTA journal Dr. G. Stemberger, Dr. F. Tocafoondi, and Dr. A. Tiziano who supported me during last 20 years in my work.

Conflict of interest statement

The author declares no competing financial interests.

Abstract

The study of cases of illusory or unstable perception of some visual stimuli allows exploration of the psychology of perception of the surrounding world. The wired construction known as “Necker cube” is one such stimulus: it can be perceived as a cube whose front face is seen higher than the back face or vice versa. The switch can occur intentionally or spontaneously. The investigations were focused on switching parameters, relation of

the switching to eye position, pre-history, and environment. Here we define that the kernel of the problem is recognizing the 2D drawing as a 3D Necker cube. To this end, we have expanded Gestalt's psychology methods that allow us to recognize 2D figures in drawings for recognizing 3D figure in a flat drawing (including the Necker cube). The presented algorithm for recognizing the cube based on the imitation principle allowed the development of the model of switching between two possible perceptions of the Necker cube. The paper shows that the predictions are in conformity with previously available experimental data. The results confirm the imitation principle of perception, and suggest expanding our research on perception to a wider class of 3D figures, opening a window into the internal processes of perception.

Keywords: Gestalt theory, Necker cube, ambiguity, imitation principle, communicative signals.

References

- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91(1), 176–180. doi:10.1007/BF00230027
- Einhäuser, W., Martin, K., & König, P. (2004). Are switches in perception of the Necker cube related to eye position? *European Journal of Neuroscience*, 20(10), 2811–2818.
- Ellis, S. R., & Stark, L. (1978). Eye movements during the viewing of Necker cubes. *Perception*, 7(5), 575–581. doi:10.1068/p070575
- Feldman, A. G., (2016). Active sensing without efference copy: Referent control of perception. *Journal of Neurophysiology*, 116(3), 960–976. doi:10.1152/jn.00016.2016
- Gregory, R. (1970). *The Intelligent Eye*. London, England: Weidenfeld and Nicolson.
- Guberman, S. (1980). *Recognition as Imitation 2. Recognition of Polyhedrons*. Preprint # 108. Moscow, Russia: Keldysh Institute of Applied Mathematics.(in Russian).
- Guberman, S. (2015). On gestalt theory principles. *Gestalt Theory*, 37(1), 25–44.
- Guberman, S. (2017). Gestalt theory rearranged back to Wertheimer. *Frontiers in Psychology*, 11(8), 1182.
- Guberman Sh. A., & Rozentsveig V. V., (1976). Algorithms for reading of handwritten texts. In: Avtomat. i Telemekh., 5, 122–129.
- Kaufman, L., & Richards, W. (1969). Spontaneous fixation tendencies for visual forms. *Perception & Psychophysics*, 5, 85–88.
- Kawabata, N., Yamagami, K., & Noaki, M. (1978). Visual fixation points and depth perception. *Visual Research*, 18(7), 853–854.
- Koffka K. (1935). *Principles of Gestalt Psychology*. Lund Humphries, London.
- Köhler, W. (1975). *Gestalt Psychology*. New York, Scarborough, Ontario: A Mentor Book.
- Metzger, W. (2006). *Laws of Seeing*: Cambridge, Mass.
- Necker, L. A. (1832). Observations on some remarkable optical phenomena seen in Switzerland; and on an optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid. *Philosophical Magazine*, 329–337.
- Pinna, B. (2010): New gestalt principles of perceptual organization. *Gestalt Theory*, 32(1), 11–78.
- Pritchard, R. M. (1961). Stabilized Images on the Retina. *Scientific American* 204:72–8. doi: 10.1038/scientificamerican0661-72. PMID: 13738093.
- Pritchard, R. M., Heron, W., & Hebb, D. O. (1960). Visual perception approached by the method of stabilized images. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 14(2), 67–77. doi:10.1037/h0083168
- Roth, W. M. (2011). *What Makes a Cube a Cube?* T&F e-books. doi:10.4324/9780203817872
- Turvey, M., (2018). *Lectures on Perception*. London: Tailor & Frensis.
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II, in *Psychologische Forschung*, 4, 301–350. English translation. <http://psychclassics.yorku.ca/Wertheimer/forms/forms.htm>

Author's Biography

Shelia Guberman got Master degree in electronics, Master Degree in mathematics, PhD in nuclear physics, and PhD in Artificial Intelligence. He published 5 books, 200+ scientific papers in nuclear physics, geophysics, seismology, medical diagnostics, image processing, perception psychology, neurophysiology, Artificial Intelligence. Retired as the Chief Scientist in Keldysh's Institute of Applied Mathematics (Russian Academy of Sciences).

Shelia Guberman is the author of core technologies of four American companies. His main achievements are: application of theory of similarity to nuclear processes, development of the technology of use of pattern recognition for applied problems, computer recognition of cursive handwriting, computer analysis of X-ray photos, efficient method for oil exploration, the theory of global D-waves – supposed triggers of strong earthquakes.

Address: Keldysh Institute of Applied Mathematics, (Retired), P.O. Box 2411, Cupertino, CA 95014, USA.

E-mail: guboil@hotmail.com

Orcid: 0000-0003-1420-0225

