Neuroscience and Behavioral Physiology, Vol. 39, No. 7, 2009

Effects of Amplitude-Frequency Characteristics of a Noise-Masked Test Stimulus on the Shapes of Visual Evoked Potentials

A. V. Merkul'ev[1] and N. S. Merkul'eva[2]

Translated from Rossiiskii Fiziologicheskii Zhurnal imeni I. M. Sechenova, Vol. 94, No. 8, pp. 928–944, August, 2008. Original article submitted March 9, 2006. Revised version received May 21, 2008.

Visual evoked potentials produced in response to a reversive checkerboard pattern presented in conditions of additive noise were recorded. Changes induced by noise in both the shapes of evoked potentials and the structure of the test stimulus were compared. The nature of changes in the shapes of evoked potentials was found to correlate with the nature of changes in the amplitude-frequency spectrum of the stimulus. These results support the gestalt psychology point of view that the visual system uses spatial frequency rather than discrete means for describing information.

KEY WORDS: visual evoked potentials, spatial frequency spectrum, gestalt psychology.

Gestalt perception is a poorly studied phenomenon. In engineering psychology, there are different views as to the extraction of figures from the background (signal from noise). The literature contains a comparative review of existing approaches and the concepts based on them [19]. More detailed information on individual approaches have been presented in [7, 15, 21]. Within the framework of these approaches, numerous models of perception have been developed, differing in terms of their underlying algorithms. However, an algorithm is a sequence of actions, and it is known that a given result can be obtained by different means. The sequence and number of steps used by the visual system for information processing are also important. It is vital to know which class of mathematical descriptions of images is used in gestalt perception.

The whole multitude of possible mathematical descriptions of images can be divided into two groups – arithmetical and geometrical. The latter can in turn be divided into local geometrical and global geometrical descriptions. Thus, three classes of possible mathematical descriptions are obtained. In accordance with these classes, all existing models of visual perception are also divided into three groups: "sample comparison" models, local features models, and spatial frequency models [7].

"Sample comparison" models. A typical example of a mathematical description is the representation of images in computers. An image is divided into points (pixels), each of which is characterized by coordinates (*XY*) and brightness (*L*). This is a graphical image form, which a person sees on a computer screen. The same image in the computer memory consists of a two-dimensional matrix of numbers. Each number in the matrix corresponds to brightness, while the row and column numbers indexing the number correspond to the coordinates of the pixel.

In "sample comparison" models, the brightnesses of image points are taken as features [7]. The working principle of these models consists of a point-by-point comparison of a noise-masked figure with reference images stored in memory, followed by decision taking with respect to identifying which reference image generates the smallest difference [13, 22]. The sample comparison procedure is very simple: the matrix of the second image is subtracted from the two-dimensional number matrix of the firest image. The extent of the difference between the images can be

[1] 15th Central Normative Research Laboratory of the Navy, 25 Konstantinovskaya, Petrodvorets, 198510 St. Petersburg, Russia; e-mail: gbyjxtn@yandex.ru.
[2] Neuromorphology Laboratory, I. P. Pavlov Institute of Physiology, Russian Academy of Sciences, 6 Makarov Bank, 199034 St. Petersburg, Russia.

0097-0549/09/3907-0683

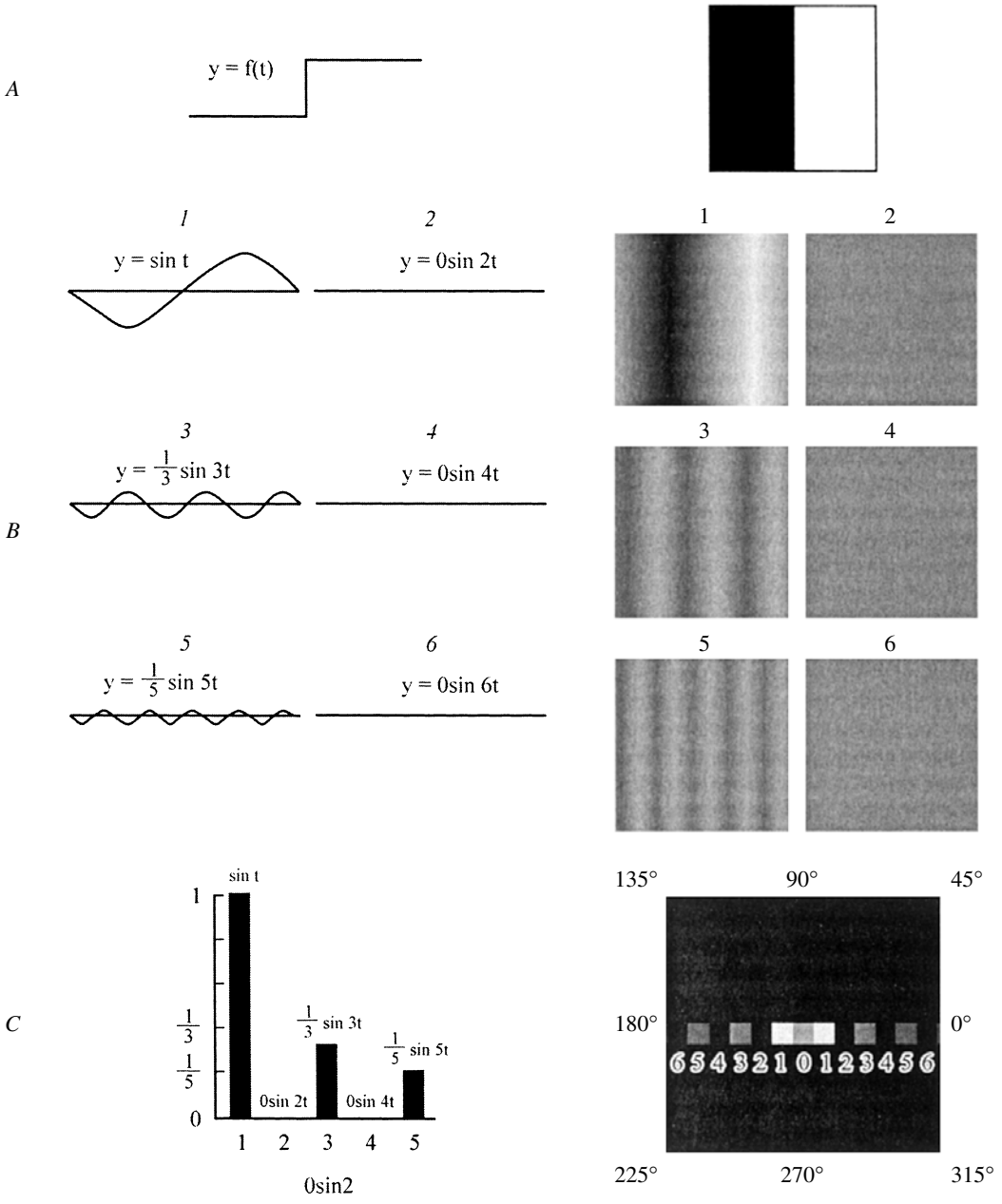

Fig. 1. Extraction of Fourier spectra. *A*) Examples of one-dimensional (left) and two-dimensional (right) signals; *B*) incomplete series of components of these signals consisting of sinusoidal Fourier components. Numbers indicate the frequencies of the components (number of periods in signal); *C*) graphical representation of the incomplete set of Fourier components of both signals in the form of amplitude-frequency Fourier spectra. For explanation see text. *A* and *B* are adapted from [6] with alterations.

expressed, for example, as the sum of the squares of the differences between all correlating comparison points in the matrix. A simulation has demonstrated that the average recognition results using "sample comparison" models for images of noise-masked figures is consistent with the average results of recognition of these same figures by trained observers [13].

Local features models. A characteristic of local geometrical descriptions is that the image is divided not into pixels, but into consistent geometrical outline elements. The figure outline is usually described using geometrical features of different views, for example, the convexity/concavity of the curve of a part, the straightness/curviness of an interval; the smoothness/sinuosity of the curve of a part, etc.

Each geometrical feature is assigned a mathematical code. That is, a local geometrical description of an image is mathematically a matrix of codes. Examples of matrixes of this type are presented in [8].

This approach to description has advantages over the point-by-point approach in terms of compactness, as the number of local geometrical features in the image is much smaller than the number of pixels [2, 8]. However, local geometrical descriptions are entirely unsuitable for object percepts. Firstly, there are no universal features suitable for images of any object. For example, convexity/concavity is not applicable to objects with a predominance of rectilinear parts in the outline. Secondly, most geometrical features are not subject to formalization [7]. Furthermore, there is no agreement in selecting a set of features and rules for combining them [19]. As a result, none of the models of perception based on local geometrical descriptions is used in practice.

Spatial frequency models. An example of global geometrical image descriptions is provided by the spatial frequency Fourier spectrum. Unfortunately, the Russian literature contains virtually no descriptions of the structures of Fourier spectra (the exception is [11]), so further understanding requires detailed consideration of this means of image description.

Any image can be regarded as resulting from the summation of sinusoidal components called gratings. Each grating is characterized by a spatial frequency (the number of periods per image or the width of the observer's field of view), the amplitude (contrast), and the phase. Resolution of the function into its sinusoidal components is called spectral analysis. Spectral analysis of one-dimensional and two-dimensional signals (mages) does not differ fundamentally.

Figure 1, *A* shows a one-dimensional rectilinear signal (left) and a black-and-white pattern (right). This pattern has a rectilinear brightness profile analogous to the profile of the signal shown at left. The only difference in the images is that one is two-dimensional. The Fourier components of both signals are shown in Fig. 1, *B* in order of increasing frequency. In both cases, there are no components with even frequencies, which is a characteristic feature of these signals, but is not a general rule (most images have complex spectra). Gratings, unlike the Fourier components of one-dimensional signals, are, firstly, two-dimensional, and secondly, are characterized by an additional parameter – orientation. This pattern is described by gratings of a single orientation: 0° (180°).

The set of Fourier components of the signal is shown graphically as a spectrum. In the case of one-dimensional signals, the abscissa shows frequency and the ordinate shows amplitude (Fig. 1, *C*, left). This type of spectrum is called an amplitude-frequency spectrum. The corresponding image spectrum is two-dimensional, as, apart from frequency and amplitude, it reflects the orientations of Fourier components. Thus, each Fourier component is represented on the spectrum as a point (Fig. 1, *C*, right).

The Fourier spectrum of an image is presented in polar coordinates. The origin of the coordinates is located in the center of the spectrum. The higher the frequency of the Fourier component, the further from the center the point corresponding to this component is plotted. The brightness of the point is proportional to the contrast of the grating. The angle formed by this point, the coordinate origin, and the horizontal depends on the orientation of the grating. As all the gratings in this example have the same orientation, all the points on the spectrum are located along a single line, corresponding to an orientation of 0° (180°). Phase Fourier spectra are also plotted, where the amplitude is replaced by the phase values of the Fourier components.

The aim of the present work was to identify which description – point-by-point or spatial frequency – is reflected by visual evoked potentials in response to stimuli presented in conditions of additive halftone noise. Noise introduces changes, which are not seen by "sample comparison" models but induce significant alterations to the amplitude-frequency spectra of the images presented. When visual evoked potentials are reflected by point-to-point stimulus descriptions, these changes will not lead to any significant changes in the shapes of potentials. If the evoked potentials are reflected the spatial frequency description, the shapes of potentials will change significantly when the parameters of noise change, and the nature of these changes will correlate with the nature of the changes in the amplitude-frequency spectrum of the stimulus. Thus, the task before us included comparison of the changes caused by noise in the shapes of visual evoked potentials and in the structure of the pattern presented to the subjects.

The complete Fourier spectrum (including the amplitude and phase parts) was not addressed in this study, as it is equivalent to a point-by-point description of an image [7]. Comparison of results obtained in experiments with complete Fourier spectra [10] and data obtained by testing an "sample comparison" model [13] indicate that both descriptions have identical advantages and disadvantages. Phase Fourier spectra, unlike amplitude spectra, carry information relating to the concrete spatial position of each element in an image [27]. Thus, the appearance in an image of even a small element of noise leads to significant changes in the whole of the phase spectrum. That is, the phase part of the spectrum is fundamentally unsuitable for use in models for the extraction of a signal from noise. This is not the case with amplitude-frequency spectra. The amplitude-frequency characteristics of test stimuli correlate significantly with the thresholds of the overall perception of fragmented images. In formal mathematical terms, fragmentation is equivalent to superimposing a multiplicative mask with apertures onto a figure [18].

As regards models operating with local geometrical descriptions, methodological considerations make them unsuitable in situations in which the object is partially obscured by noise. Comparison of matrixes of object fea-

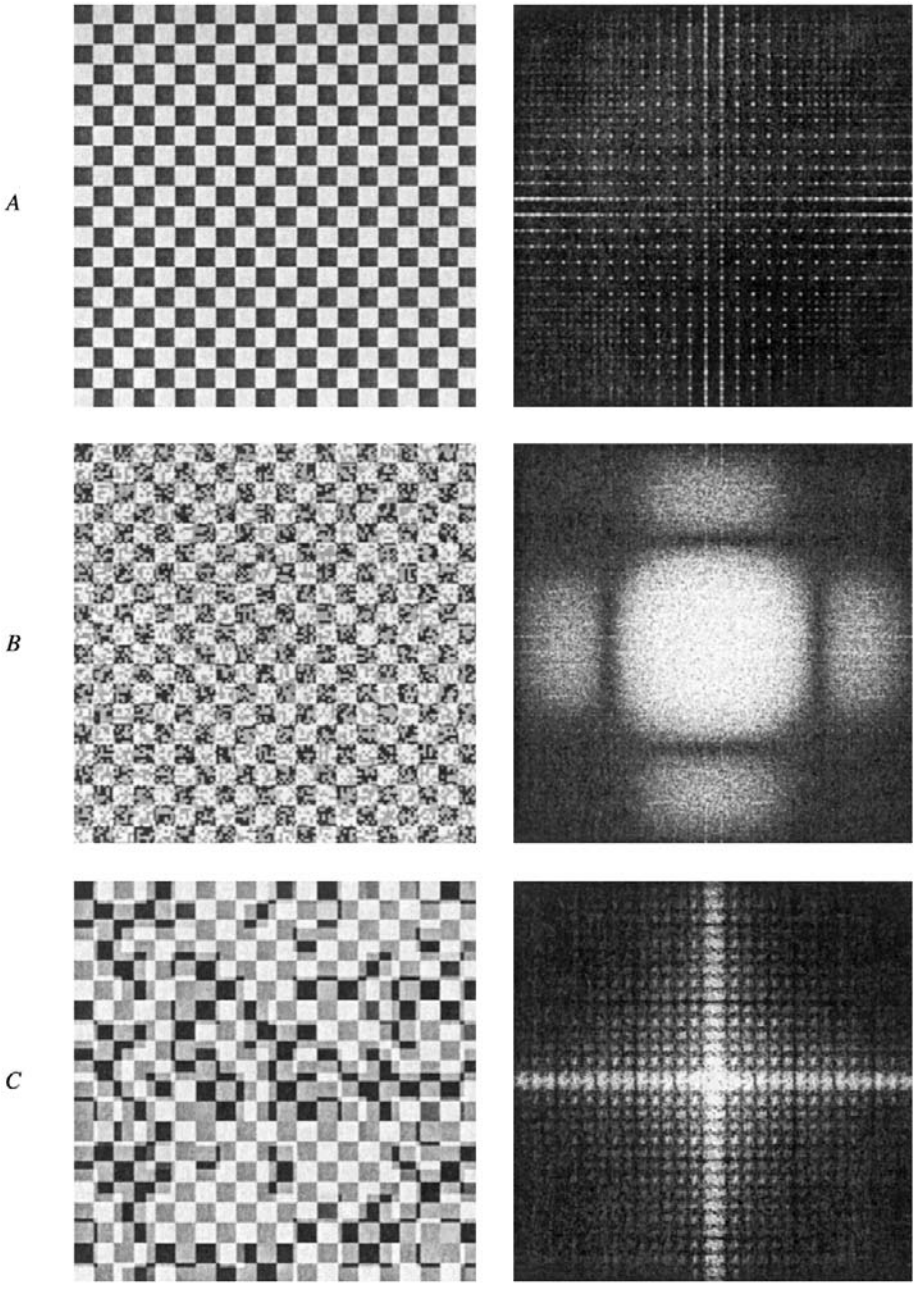

Fig. 2. Test stimuli and their two-dimensional amplitude-frequency spectra. *A*) Checkerboard pattern without noise; *B*, *C*) checkerboard patterns on the background of noise with element sizes of 3 and 24 pixels respectively. Spectra are shown at right. For explanation see text.

ture codes are performed by working round the outline, which must be a closed line [2, 8]. Noise not only obscures part of the object, but also breaks up its outline. Thus, adherents to models using local geometrical features take the view that the visual system performs preliminary spatial frequency processing of the image to extract the object's outline. Only then are algorithms operating with local geometrical features triggered [17, 26].

METHODS

A total of nine subjects aged 19–25 years took part in the study, and all had normal visual acuity.

Recording of visual evoked potentials. Visual evoked potentials were recorded using a Telepat 01 encephalograph with programs by Yu. D. Kropotov and V. A. Ponomarev, Institute of the Human Brain, and S. V. Pronin, I. P. Pavlov

Institute of Physiology. The signal was passed from the output of the amplifier (bandpass 0.5–30 Hz) via an analog-to-digital converter with a sampling frequency of 250 Hz to a computer. Recording was performed from occipital lead Oz relative to combined ear electrodes. Stimulation was performed using a 17" Sony Trinitron monitor located 1.6 m from the subject.

Stimuli consisted of a black-and-white checkerboard pattern (20×20 cells) of size 8.57°. The cell size in the pattern corresponded to a spatial frequency of 1.22 cycles/° (Fig. 2, *A*). The contrast of the pattern was determined using the Michaelson equation – $(L_{max} - L_{min})/(L_{max} + L_{min})$, where L_{max} and L_{min} are the maximum and minimum brightness values – was 0.64. The subject was told to fix the gaze on a black spot located at the center of the monitor. The checkerboard pattern was presented binocularly 50 times in a reversion regime (at a frequency of 1 Hz) on a uniform background and on the background of additive noise. The noise was stationary, i.e., did not change on reversion of the test stimulus.

Five types of noise were used, with different element sizes: 3, 6, 12, 24, and 48 pixels (9.30, 4.65, 2.33, 1.16, and 0.58 cycles/°, respectively). Checkerboard patterns with addition of two of these noises are shown in Fig. 2, *B*, *C*. Noises were synthesized using a computer simulation program. The only controllable noise parameters (apart from element size) were contrast (0.3 in all cases) and probability (50% in all cases). Constant mean image brightness was produced by the program at different noise parameters.

Thus, with different noise element sizes, 50% of the pattern area was always noise-free and the mean brightness of the noise-masked surface was always identical. In these cases, differences between noise-masked and non-noise-masked images were always constant for "sample comparison" models regardless of noise element size.

Visual evoked potentials to checkerboard patterns without noise and evoked potentials to noise-masked checkerboard patterns (individually for each noise) were averaged separately for each subject. Differences in evoked potentials were assessed using correlation coefficients and mean square differences at each point. Significant differences in the main components of evoked potentials were compared using the Wilcoxon *T* test. Critical values for tests were taken from biometric tables [14].

Construction of Fourier spectra of images. Spatial frequency spectra of images were obtained using a program developed by V. B. Makulov and V. N. Pauk at the S. I. Vavilov State Optics Institute. Images were transformed into the two-dimensional spatial frequency spectrum by the program using a fast Fourier transformation algorithm. Drops in the amplitude of the Fourier components at low and high frequencies were by two orders of magnitude, and decreases in amplitude from low to high spatial frequencies occurred very sharply. As a result, the overall spectrum appeared black with the exception of a small white spot at the center. Spectra were therefore presented by applying an equalization procedure (logarithmic smoothing) to the amplitude. As a result, the spectral components of the spectra shown in Fig. 2, right appear white.

One-dimensional sections were extracted from one-dimensional spectra at the main orientations: 0°, 45°, 90°, and 135°. Sections corresponding to the diagonals of the spectrum (45° and 135°) were averaged separately, as were sections with the horizontal and vertical components of the spectrum (0° and 90°). Thus, the two-dimensional amplitude-frequency spectra of test images were presented in one-dimensional form as two sections corresponding to the main orientations.

RESULTS

Electrophysiological studies. Averaged (for all subjects) evoked potentials recorded in response to the checkerboard pattern presented in conditions of noise and a uniform background are shown in Fig. 3, along with results obtained from comparisons. Visual evoked potentials to the non-noise-masked pattern consisted of five main components: two positive components with latent periods of 124 (P124) and 252 msec (P252) and three negative components with latent periods of 92 (N92), 172 (N172), and 352 msec (N352). These evoked potentials were comparable with the evoked potentials obtained in response to presentation of the same checkerboard pattern in conditions of additive noise (Fig. 3, *A–D*, left).

Significant differences in visual evoked potentials to noise-masked and non-noise-masked patterns were seen for negative components N92 and N172 with all noises used in the present study (Table 1). The amplitudes of the negative components were found to change when the size of the noise elements changed, which is clearly evident in the example of component N92. This component completely disappeared in the presence of noise with an element size of three pixels (Fig. 3, *A*). It appeared in the presence of six-pixel noise, though its amplitude was very low. Further increases in noise element size were accompanied by gradual increases in the amplitude of the N92 component, though it remained significantly smaller with a noise element size of 48 pixels than in the absence of noise (Fig. 3, *D*).

This time pattern of differences in response to changes in the characteristics of noise was also demonstrated by other the other components of evoked potentials which were compared. Figure 3, *A–D* shows plots of the squares of the differences at each point in the visual evoked potentials being compared. The plots clearly show four peaks, the first three of which have latencies corresponding to those of the negative components of the evoked potentials to the non-noise-masked pattern, N92, N172, and N352, respectively, while the fourth peak did not correspond to any of its components (it was an unidentified component). The plots

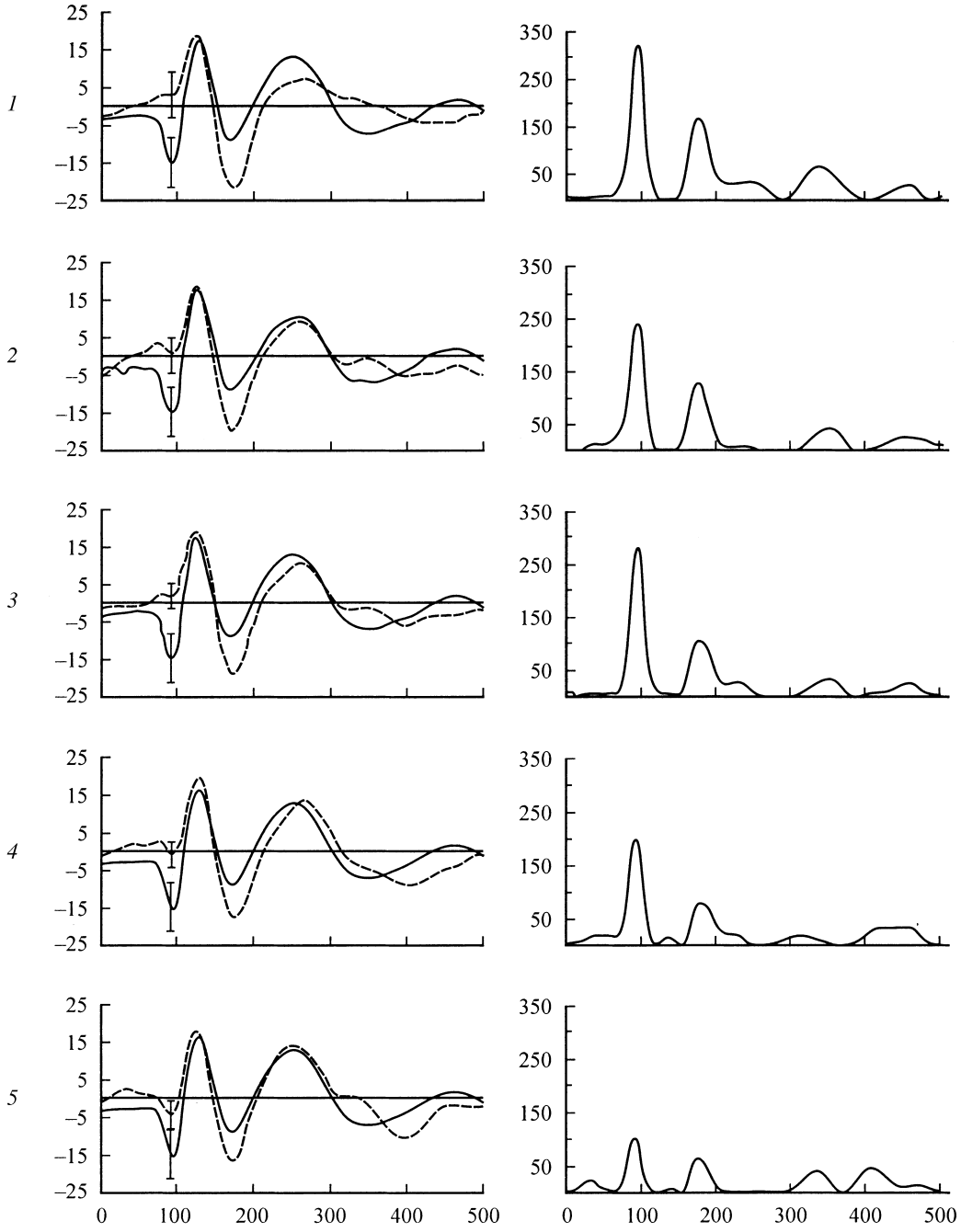

Fig. 3. Differences in evoked potentials to patterns with and without noise. Noise element sizes: *1*) 3 × 3; *2*) 6 × 6; *3*) 12 × 12; *4*) 24 × 24; *5*) 48 × 48 pixels. The abscissas of all plots show time, msec; the ordinates of the plots to left show evoked potentials to noise-masked (dotted plots) and non-noise-masked (continuous plots) checkerboard patterns (μV); the ordinates of the plots at right show the squares of the differences at each point of the evoked potentials being compared. Significant intervals at a significance level of 5% are shown for the N92 component.

show that the greatest differences were characteristic of the early components. However, in the presence of noise with the biggest element size used here (48 pixels), the magnitudes of the difference in the amplitudes for all visual evoked potential components became essentially identical. The mean square difference characterizing the total difference in the shapes of the evoked potentials being compared decreased monotonously as noise element size increased (Table 2). Correlation analysis also demonstrated a monotonous increase in the similarity of evoked potentials

TABLE 1. Wilcoxon *T* Test Values for the N92 and N172 Components of Visual Evoked Potentials Under Comparison

Noise element size, pixels		3 × 3	6 × 6	12 × 12	24 × 24	48 × 48
VEP component	N92	0	2	0	1	4
	N172	0	3	3	4	5

Note. For a sample set of $n = 9$, TsT values were 8 and 9 at significance levels of 1% and 5%, respectively.

TABLE 2. Statistical Similarities and Differences in Visual Evoked Potentials to Noise-Masked and Non-Noise-Masked Checkerboard Patterns for Different Noises

Noise element size, pixels	3 × 3	6 × 6	12 × 12	24 × 24	48 × 48
Correlation coefficient, *R*	0.59	0.67	0.70	0.77	0.83
Mean square difference	42.87	30.12	28.93	27.37	19.51

to noise-masked patterns with evoked potentials to the non-noise-masked as noise element size increased (Table 2).

Image studies. One-dimensional profiles of amplitude-frequency spectra of noise-masked checkerboard patterns are shown in Fig. 4, *A*. This shows that spectrum shapes were not identical in different orientations. The spectrum of the checkerboard pattern itself had no horizontal or vertical components, so the horizontals and verticals of the spectrum of the noise-masked pattern only contain noise. The diagonal components of the spectrum of the noise-masked pattern contained spatial-frequency components of both the pattern and the noise. The checkerboard pattern was characterized by a periodic brightness profile, so the energy distribution across its spectrum was banded. The first component (harmonic) of the checkerboard pattern had a spatial frequency of 1.22 cycles/°, which was determined by the size of the cells in the pattern. The other components of the pattern were odd harmonics (3.66, 6.10, ..., 18.30 cycles/°). The brightness distribution by area occupied by noise was aperiodic, so the spectrum was not banded. However, because of the predominance of elements of a single size in the noise, its spectrum contained elevations, reaching peaks corresponding to the odd harmonics.

Differences in the spectra of the noise-masked and non-noise-masked patterns were assessed using only the mean square differences presented in Table 3. These data show that increases in noise element size were accompanied by a monotonous increase in the difference between the spectra of the noise-masked and the non-noise-masked patterns.

A further characteristic of the noises used here is the non-uniform distribution of their Fourier coefficients for different orientations of the spectrum: it is evident from the plots in Fig. 4, *A* that in the diagonal orientations, noise is much less apparent than in the horizontal and vertical orientations. The width of the noise-masked range in the diagonals of the spectra decreased with increases in noise element size. In Table 4, mean noise amplitudes in the ranges located between the harmonics of the checkerboard pattern are given for the diagonals of the spectrum (averaging allows reliable determination of the upper limit of the spectral range occupied by each of the noises).

Comparison of electrophysiological and image results. As noise element size increased, the evoked potentials to noise-masked checkerboard patterns became more similar to evoked potentials obtained in response to presentation of the non-noise-masked pattern. However, the mean square difference of the spectra of the noise-masked and the non-noise-masked patterns increased. The widths of the noise-masked range in the horizontal and vertical orientations was constant, and the noise occupied the whole range of the spectrum of the pattern (Fig. 4, *a*, left).

However, the width of the diagonals of the spectrum occupied by noise was very consistent with the electrophysiological results. Figure 4, *B* compares plots demonstrating changes in this characteristic and the magnitudes of differences between evoked potentials with increases in noise element size. It is clear that the nature of changes in the structure of the spectrum of the pattern corresponds to the nature of the changes in the differences between the shapes of the evoked potentials as a whole. Thus, additive noise leads to significant changes in the shapes of evoked potentials, the extent of which corresponds to the magnitude of the range of the spectrum of the pattern occupied by noise.

DISCUSSION

The data obtained here provide evidence that visual evoked potentials reflect a geometrical rather than an arithmetical description of the stimulus. This is also indicated by data from other electrophysiological investigations. Visual evoked potentials to inverted and non-inverted patterns are known not to show significant differences. A similar conclusion was drawn for visual evoked potentials to mirror

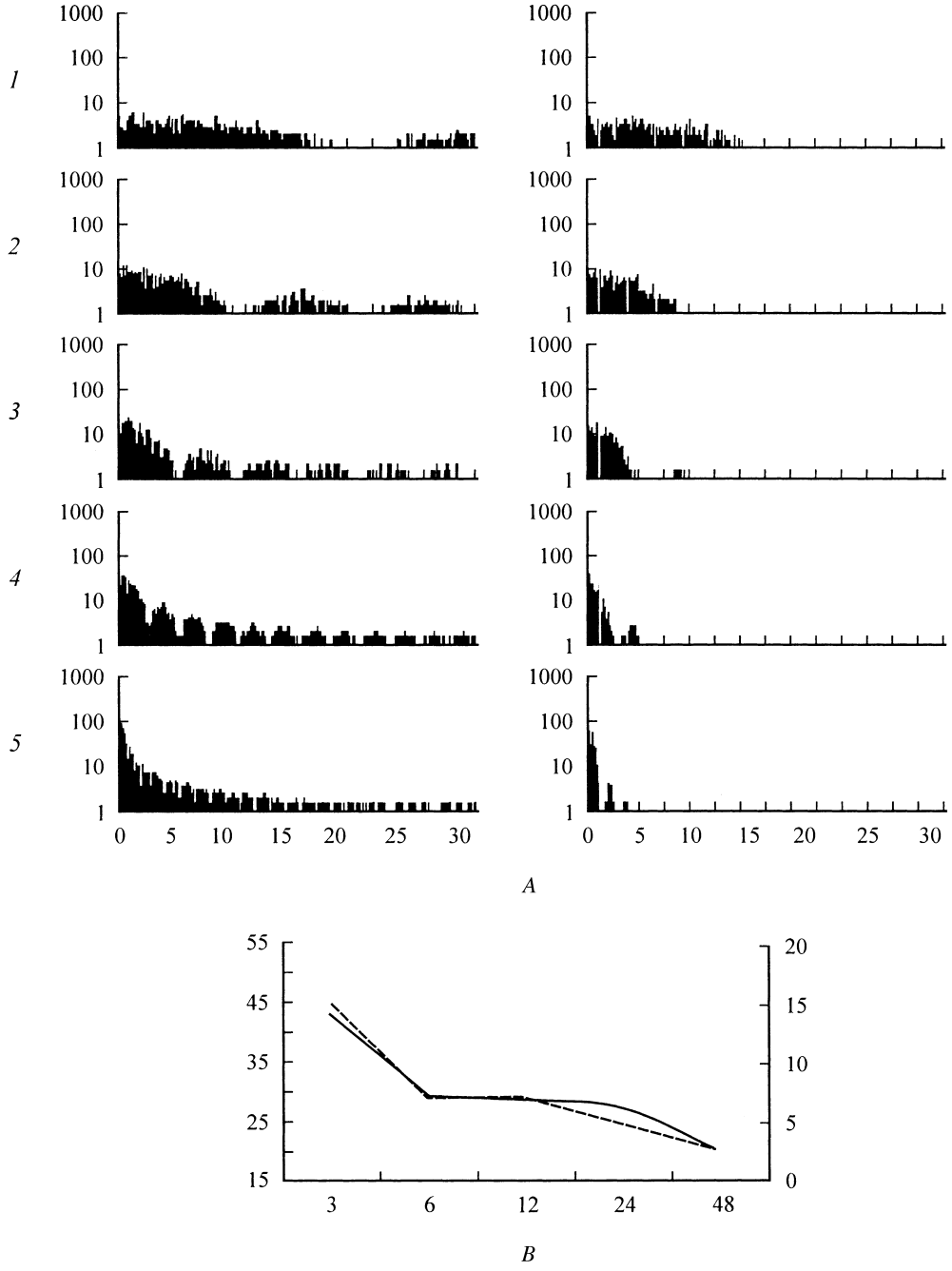

Fig. 4. Comparison of changes introduced by noises to the shapes of evoked potentials and stimulus spectra. *A*) Amplitude-frequency spectra of noise-masked checkerboard patterns. Noise element sizes: *1*) 3 × 3; *2*) 6 × 6; *3*) 12 × 12; *4*) 24 × 24; *5*) 48 × 48 pixels. The abscissas show spatial frequency, cycles/°; the ordinates show amplitude, arbitrary units. Left: averaged verticals and horizontals; right: averaged diagonals of spectra. Checkerboard pattern components are shown in gray and noise components in black. *B*) Dynamics of changes in the shapes of evoked potentials and the diagonal components of the spectrum of the pattern with increases in noise element size. The abscissa shows noise element size, pixels; the right ordinate shows the width of the range of the diagonals of the spectrum occupied by noise, cycles/° (dotted plot); the left ordinate shows squares of the differences between the evoked potentials to the noise-masked and non-noise-masked patterns (continuous plot).

reflections of images [33]. Electrophysiological studies in which stimuli are presented in a reversion regime therefore generally involve averaging of the left and right parts of the evoked potentials. It should be noted that amplitude-frequency spectrum is invariant to image reversion, which is not the case for arithmetic descriptions.

TABLE 3. Mean Squares of Differences between Spectra of Noise-Masked and Non-Noise-Masked Checkerboard Patterns for Different Noise Element Sizes

Noise element size, pixels		3 × 3	6 × 6	12 × 12	24 × 24	48 × 48
Orientation	0 and 90°	2.82	7.66	11.55	31.04	59.39
	45 and 135°	1.83	4.33	7.26	16.44	30.54

TABLE 4. Distribution of Noise Energy in the Diagonal Orientations of the Spectrum of the Noise-Masked Checkerboard Pattern

Noise element size, pixels		3 × 3	6 × 6	12 × 12	24 × 24	48 × 48
Centers of diagonals, cycles/°	0.61	2.28	5.50	9.94	17.61	22.89
	2.44	2.16	4.11	5.13	1.61	0.53
	4.88	2.26	2.58	0.05	0.21	–
	7.32	1.42	0.53	0.13	–	–
	9.76	1.32	–	–	–	–
	12.20	0.47	–	–	–	–
	14.64	0.03	–	–	–	–
	17.08	–	–	–	–	–

For "sample comparison" models, inversion of the stimulus makes it "unrecognizable." This is illustrated in Fig. 5, *A*. The image of the white square on a black background (the image) is compared point by point with three other patterns: a point-by-point copy of the image, the image of a white circle on a black background, and an inverted copy of the image. Comparisons were performed using the "sample comparison" model included in Adobe Photoshop (the "difference" procedure). Subtraction of the arithmetic matrix of the first pattern from the matrix of the image gives only null values, as the two images are identical. The resulting matrix on the monitor is apparent as a completely black image, as the brightness value 0 corresponds to the black color. A different result is obtained by comparison with the image of the second pattern. Some of the corresponding pixels in this image have brightness 0 (0 = black, 255 = white). Thus, after subtraction of the matrix of the pattern from the matrix of the image, some of the resulting arithmetic values will be 255 (the sign of the difference is disregarded). Correspondingly, the total brightness of the resulting image differs from the null. However, the greatest difference from the image is obtained with the third pattern. That is, a figure of the same shape and size as the image differs from it more than the figure of a different shape because of the differences in the brightnesses of the corresponding points. Models operating with arithmetic descriptions to not "see" shape.

Thus, if the visual system operates with arithmetic descriptions of objects, then the left and right parts of the evoked potentials to the reversive stimulus must differ, which is not seen in practice (see above). The gestalt percept is invariant to absolute brightness values. For example, Lashley's studies [31] involved training of rats to select one of two figures presented and to avoid the other (Fig. 5, *B*, above). After the skill was stably fixed, the rats were presented with another pair of figures with different brightnesses compared with the figures of the first pair (Fig. 5, *B*, below). The previously acquired reaction was found to be transferred quickly to the new figures.

Gestalt psychology, which before the appearance of engineering psychology attempted to seek mathematical means of describing the phenomena of perception [1], initially came out against arithmetical descriptions. A human perceives an object rather than a set of points with different brightnesses. Point-by-point descriptions are no more than a mosaic of unconnected elements (inert atoms) [29, 34].

Engineering psychology, closely associated with cognitive psychology, was also against point-by-point descriptions, as they lack any recognition of a "pathway" to being combined. For example, the junction points in the image of a cat's head shown in Fig. 5, *C* can be combined such that the resulting shape is completely unrecognizable [2].

Associative psychology, the underlying concept of which is the "sample comparison" model, did not agree with this position. It took the view that the elements of a mosaic are combined into a gestalt by associations [3], i.e., by temporal links. Gestalt psychology in turn denied the role of associations in the gestalt percept [29]. The image even of a not particularly complex visual scene can contain several objects. If the visual system is based on point-by-point

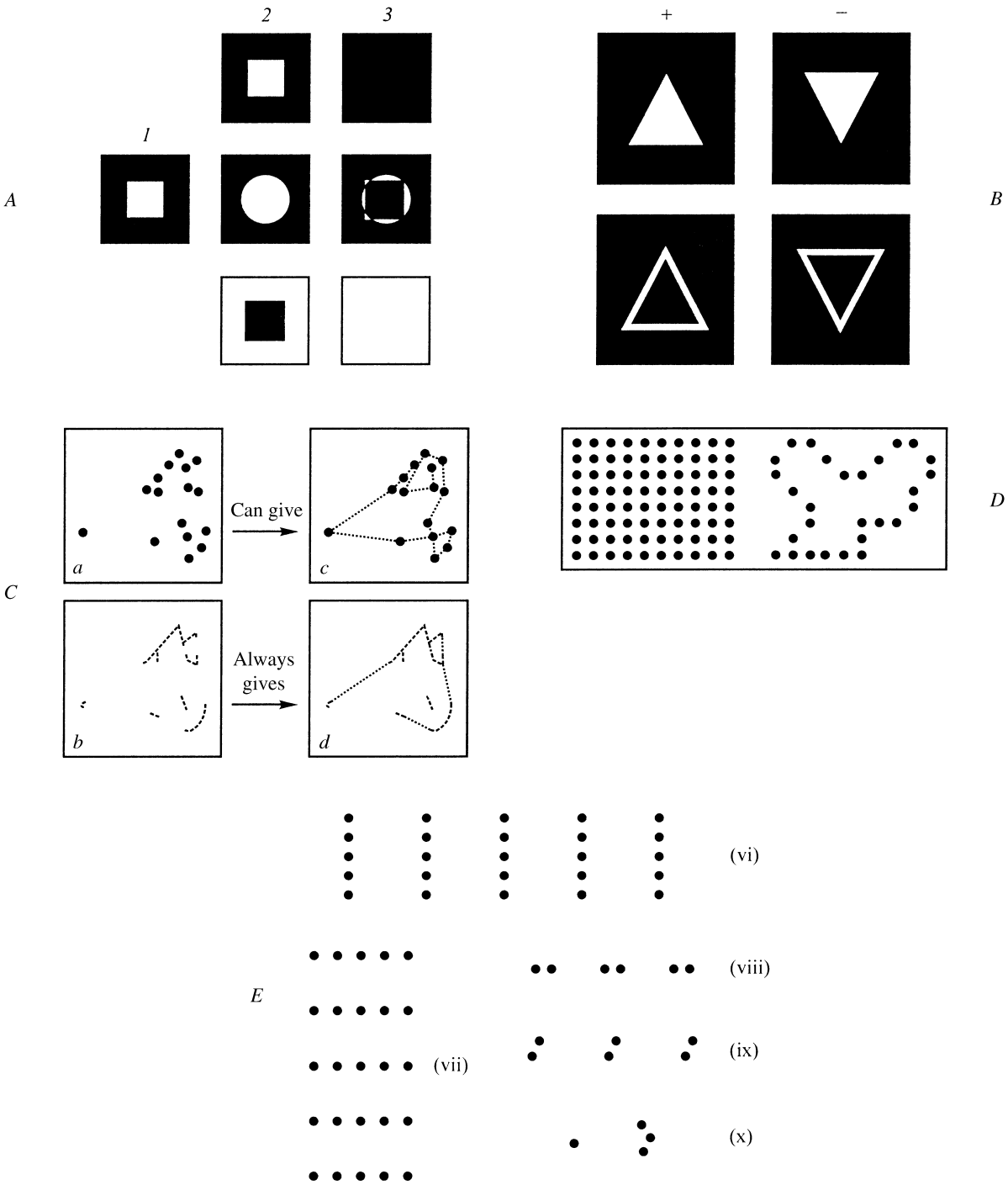

Fig. 5. Examples countering the view that the visual system uses arithmetical descriptions. *A*) Illustration showing the information on object shape ignored by "sample comparison" models. *1*) Sample; *2*) images to be compared with sample; *3*) results of comparisons. For explanation see text. *B*) Illustration of the invariance of the gestalt percept to stimulus brightness. Above: stimuli for which positive ("+") and negative ("–") reactions generated reinforcements. Below: stimuli to which rats were able to "transfer" the corresponding reactions without reinforcement (from [31]). *C*, *D*) Illustrations of the ambiguity of point-by-point descriptions; *C*) images of "Attneave's cat," consisting of corner points (above) or angles (below) and possible ways of combining them (from [2]); *D*) illustration of the Kanysza paradox (from [24]). For explanation see text. *E*) Examples of punctate images used by gestalt psychologists to illustrate one of the rules of gestalt perception, the "nearness rule." There is a link between close-lying elements which combines them into a single gestalt. For example, the result of this linkage is that figure vi is always perceived as a set of columns and figure vii as a set of rows (from [34]).

descriptions, then the simultaneous perception of such a scene will involve the linkage of all its elements by associations. This is illustrated by the Kanysza paradox (Fig. 5, *D*): the matrix shown at left potentially contains an enormous number of different shapes, for example the figure shown in Fig. 5, *D* at right [24]. If the visual gestalt percept depended on associations, a human would see anything he wanted to see in this matrix.

However, there are images which are initially presented in point-by-point discrete ways. This does not apply only to computer and television images and typographic reproductions, but even to photographs, whose discrete nature arises from the graininess of the light-sensitive material. This led to the criticism of gestalt psychology that in criticizing arithmetical descriptions, it did not take account of the discrete (pixel) nature of many images perceived by humans without any difficulty [24].

Kliks made a more general comment in addressing gestalt psychology. He pointed out the fact that the retina is a mosaic of receptors. Therefore, all stimuli falling on the retina and having physiological actions are accumulations of punctate and identical particles distributed as mosaics and unconnected to each other [12]. That is, the image of the object at the input of the visual system is a point-by-point description. And this, in the framework of gestalt psychology, is not suitable for the perception of shapes.

Both comments could also be directed against engineering psychology, which is dominated by the concept that perception uses local geometrical descriptions. However, engineering psychology avoided this criticism, by taking the so-called computational approach to visual perception as its main approach, as developed by Marr. This approach is based on Gibson's ideas regarding the restoration of the "pure" properties of three-dimensional objects by the visual system [19]. Gibson's view was that the perception of any two-dimensional image is not unique. Even an image formed by the optics of the eye on the retina has no relationship to perception, as the visual system gathers information directly from the light stream [5] (however, these ideas of Gibson were not discussed in [17]). Thus, the computational approach derives discrete images from its studies.

At first sight, gestalt psychology holds contradictory positions. On the one hand, in gestalt theory, the shape (gestalt) is regarded as an objective reality; it does not impose reasoning on sensory elements, but extracts the percept from the background [16]. However, filtered by the mosaic "sieve" of the retina, it should disappear in accordance with the views of gestalt psychology of arithmetical descriptions. Gestalts should also be absent from punctate images. However, the gestalt is somehow present in the percept although, from the gestalt psychology point of view, it cannot be synthesized de novo. The gestalt theory denies the existence of any kind of synthetic (or analytical) process in perception [3]. Everything we perceive is already present in our sensations, so there is no need to introduce additional acts of ordering a sensory structure [23].

However, gestalt psychology made no discrimination between discrete and analog images. As a result, Wertheimer [34] illustrated gestalt phenomena of perception using punctate, discrete textures (Fig. 5, *E*). From the point of view of gestalt psychology, gestalts are also present in punctate images, though they are obscured by the background. Thus, the problem of the perception of gestalts is not one of the discrete nature of images, but one of the form of their description. The mosaic nature disappears at the very beginning of sensory reception [28], while the representation of information in the brain is distributed (antilocation or equipotential) in nature [30]. That is, gestalt perception is based on some kind of global geometrical description which can be obtained from an arithmetical description.

This idea can be understood using a simple analogy. We assume that some algorithm exists to allow unambiguous translation of Russian texts into English and, conversely, from English into Russian. If this hypothetical algorithm is used to translate a Russian poem, then the rhyme present in the Russian version will be absent from the English. However, if the resulting English text is translated back into Russian, the rhyme will reappear. The rhyme never disappeared – it was merely "latent" in the English version of the text, in the same way that a gestalt is latent in the arithmetical description.

The mathematical tool proposed by the French physicist Fourier, i.e., Fourier transformation, provides for unambiguous translation of any function consisting of discrete components into a Fourier spectrum, along with the reverse translation [20]. That is, it is a method for the direct progression from arithmetical to global geometrical descriptions. Unlike arithmetical matrixes, the Fourier spectrum is a distribution (holographic) description of the image. Every grating contains some information on every point in the image. Every point in the image is distributed over the whole spectrum.

This distribution of information across the spectrum is not uniform. For example, the low-frequency components located in the center of the Fourier spectrum carry information relating predominantly to the larger details of the image, while the smaller details are described mainly by the high-frequency components located at the periphery of the spectrum. This feature provided an explanation of several visual gestalt phenomena in terms of spatial frequency filtration [25]. The question of extracting gestalts will be addressed in a separate report. Here we note only that the extraction of gestalts can occur only by using spatial frequency descriptions. We do not believe that these points provide adequate grounds for proposing that the visual system uses Fourier descriptions of images. Different types of a single class of descriptions should have common properties, which must lead to essentially identical results when

these are used. Thus, the question of the actual form of a global geometrical description used by the visual system remains unclear [9, 32].

According to one of the propositions of Aristotelian logic, arithmetic should not be used to solve geometrical problems [3]. In arithmetic, unlike geometry, the concepts of "shape" and "figure" do not exist, so studies and simulations of visual gestalt perception should not use arithmetical, but rather geometrical descriptions of stimuli. The results obtained in the present study provide further support for this view.

REFERENCES

1. B. G. Anan'ev, M. D. Dvoryashina, and N. A. Kudryavtseva, *Individual Human Development and the Constancy of Perception* [in Russian], Prosveshchenie, Moscow (1968).
2. M. Arbib, *The Metaphorical Brain* [Russian translation], Mir, Moscow (1976).
3. Aristotle, *Posterior Analytics. Collected Writings* [Russian translation], Mysl', Moscow, Vol. 2, pp. 255–346.
4. É. G. Vatsuro, *I. P. Pavlov's Studies of Higher Nervous Activity* [in Russian], Ministry of Education of the RSFSR Press, Moscow (1955).
5. D. Gibson, *An Ecological Approach to Visual Perception* [Russian translation], Progress, Moscow (1988).
6. V. D. Glezer, K. N. Dudkin, A. M. Kuperman, L. I. Leushina, A. A. Nevskaya, N. F. Podvigin, and N. V. Prazdnikova, *Visual Recognition and Its Neurophysiological Mechanisms* [in Russian], Nauka, Leningrad (1975).
7. D. V. Glezer, N. V. Prazdnikova, L. I. Leushina, A. A. Nevskaya, and M. B. Pavlovskaya, "The description of visual images," in: *Visual Physiology (Handbook of Physiology)*, A. L. Byzov (ed.), Nauka, Moscow (1992), pp. 466–527.
8. R. M. Granovskaya, *Perception and Models of Memory* [in Russian], Nauka, Leningrad (1974).
9. K. N. Dudkin, *Visual Perception and Memory* [in Russian], Nauka, Leningrad (1985).
10. D. Keisasent, "Recognition of objects and features," in: *Optical Holography* [Russian translation] G. Colfield (ed.), Mir, Moscow (1982), Vol. 2, pp. 550–594.
11. D. I. Kirvyalis, "Theoretical basis of visual analyzers," in: *Visual Systems. Proceedings of the All-Union Symposium "Vision in Organisms and Robots"* [in Russian], Vilnius (1987), pp. 6–46.
12. F. Kliks, *Problems in the Psychophysiology of the Perception of Space* [in Russian], Progress, Moscow (1965).
13. N. N. Krasil'nikov, *Theory of Image Transmission and Perception* [in Russian], Radio i Svyaz, Moscow (1986).
14. G. F. Lakin, *Textbook for Biological Specialists in Colleges* [in Russian], Vysshaya Shkola (1990).
15. L. I. Leushina, A. A. Nevskaya, and M. B. Pavlovskaya, "Asymmetry in the hemispheres from the point of view of the recognition of visual images," *Sensor. Sistemy*, **3**, No. 4, 76–92 (1982).
16. T. Likhi, *The History of Contemporary Psychology* [in Russian], Piter, St. Petersburg (2003).
17. D. Marr, *An Informational Approach to Studies of the Representation and Processing of Visual Images* [Russian translation], Radio i Svyaz, Moscow (1987).
18. A. V. Merkul'ev, S. V. Pronin, L. A. Semenov, N. Foreman, V. N. Chikhman, and Yu. E. Shelepin, "Threshold signal:noise ratios in the perception of fragmented figures," *Ros. Fiziol. Zh. im. I. M. Sechenova*, **90**, No. 11, 1348–1355 (2004).
19. L. F. Petrova, V. P. Smirnov, O. V. Alekseev, Yu. K. Allik, and A. V. Pul'ver, "Visual perception of images," in: *Processing of Optical Images* [in Russian], M. M. Miroshnikov (ed.), S. I. Vavilov GOI, St. Petersburg (1990), No. 10.
20. A. Poincare, *Science* [Russian translation], Nauka, Moscow (1990).
21. A. For, *Image Perception and Recognition* [in Russian], Mashinostroenie, Moscow (1989).
22. A. K. Kharauzov, Yu. E. Shelepin, S. V. Pronin, N. N. Krasil'nikov, and S. V. Murav'eva, "Visual evoked potentials on dichotic presentation of test sinusoidal gratings and interference," *Ros. Fiziol. Zh. im. I. M. Sechenova*, **87**, No. 2, 261–270 (2001).
23. M. G. Yaroshevskii, *The History of Psychology* [in Russian], Mysl, Moscow (1976).
24. A. Desolneux, L. Moisan, and J.-M. Morel, *From Gestalt Theory to Image Analysis. A Probabilistic Approach*, New York (2006).
25. A. Ginsburg, *The Perception of Visual Form: a Two-Dimensional Filter Analysis. Information Processing in the Visual System*, Leningrad (1976), pp. 46–51.
26. L. D. Harmon and B. Julesz, "Masking in visual recognition. Effects of two-dimensional filtered noise," *Science*, **180**, 1194–1197 (1973).
27. B. Julesz, E. N. Gilbert, and J. D. Victor, "Visual discrimination of textures with identical third-order statistics," *Biol. Cybernet.*, **31**, 137–140 (1978).
28. K. Koffka, *Principles of Gestalt Psychology*, New York (1978).
29. W. Kohler, "Gestalt psychology today," *Amer. Psychol.*, **14**, 727–734 (1959).
30. K. S. Lashley, *Brain Mechanisms and Intelligence*, Chicago University Press (1929).
31. K. S. Lashley, "The mechanism of vision. Preliminary studies of the rat's capacity for detail vision," *J. Gen. Psychol.*, **18**, 123–193 (1938).
32. A. L. Ochs, "Is Fourier analysis performed by the visual system or by the visual investigator?" *J. Opt. Soc. Amer.*, **69**, No. 1, 95–98 (1979).
33. L. H. Van der Tweed, "Pattern evoked potentials: facts and considerations," in: *Proc. of the 16th ISCEV Symposium*, Morioka (1979), pp. 27–46.
34. M. Wertheimer, "Laws of organization in perceptual forms," first published as "Untersuchungen zur Lehre von der Gestalt II," in: *Psycologische Forschung*, **4**, 301–350 (1923). Translation published in *A Source Book of Gestalt Psychology*, Ellis, W. (ed.), Routledge & Kegan Paul, London (1938), pp. 71–88.