PROCEEDINGS OF THE LATVIAN ACADEMY OF SCIENCES. Section B, Vol. 71 (2017), No. 5 (710), pp. 20–30.
DOI: 10.1515/prolas-2017-00XX

OPTIMAL OPTOTYPE STRUCTURE FOR MONITORING VISUAL ACUITY

Galina Rozhkova#, Dmitry Lebedev, Maria Gracheva, and Svetlana Rychkova

Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Sciences,
Bolshoy Karetny per. 19, build.1, Moscow 127051, RUSSIA

Corresponding author, gir@iitp.ru

Communicated by Ivars Lācis

To date, there are no generally accepted optotypes for monitoring visual acuity. All common optotypes are not completely suitable for some reasons. The tasks requiring visual monitoring — investigation of visual development, early diagnostics, assessment of treatment — impose heavy demands on the test stimuli. They must be: (1) suitable for patients of any age; (2) convenient for repeatable examinations; and (3) accurate enough for revealing the smallest physiologically significant changes of visual acuity. From theoretical consideration, one could conclude that the optotypes for monitoring visual acuity should be designed for measuring visual resolution but not recognition, unlike most popular optotypes. The best optotypes for visual resolution are grating-like stimuli whose recognition could only be based on the high frequency part of the Fourier spectrum around the characteristic frequency (not on the low-frequency components). On the basis of theoretical analysis we elaborated modified 3-bar optotypes, which minimise the possibility of using low-frequency cues for stimulus recognition. In this paper we present the results of theoretical and experimental comparison of these optotypes with the two widely used ones: tumbling-E and standard 3-bar targets. According to the data obtained, our modified optotypes seem to be better than other investigated ones for monitoring visual acuity.

Key words: *visual acuity, resolution, recognition, tumbling-E, 3-bar target, low-frequency cues.*

INTRODUCTION

Assessment of visual acuity is an obligatory procedure in any examination of the subject's visual capabilities. However, up to date, there is no consensus among optometrists and visual scientists regarding the best test stimuli, measuring procedures and conditions of measurements. In one and the same individual, the result of visual acuity assessment depends on many physical, physiological and cognitive factors. Due to the absence of a standard, universally adopted, convenient and not time-consuming technique, visual acuity is rarely measured accurately if it is not crucial for the study. Regular optometric screening studies are mainly aimed at revealing negative deviations from conventional norms. Visual scientists carrying out various experiments not related to investigation of the mechanisms of visual acuity, usually perform only rough examination to become convinced that their subjects have *conventionally normal vision*. The opticians often obviate the need for measuring visual acuity at all, since the main objective they pursue is to prescribe the best optical correction for the patient and it is evident that this objective can be achieved without accurate visual acuity assessment. One can simply ask the patient to try various eye-glasses while looking at a proper set of images differing in size. In particular, if a patient needs eyeglasses mainly for reading, the optimal test stimuli would evidently be some fragments of the printed text with letters of various sizes. For preschool children, the optimal stimuli can be simplified images of familiar objects or animals.

Herman Snellen was the first to design special images for the purpose of visual acuity measurements and named them "optotypes" (Snellen, 1862). Intuitively, a proper choice of the optotypes seemed to be very important for the accuracy and reliability of measurements, but many optotypes used in the past and nowadays were created without ą serious analysis of their properties and the information they can provide about the subject's visual system.

The majority of popular complex test stimuli — letters, pictures, Lea symbols, etc. require many parameters for description. Moreover, researchers are often unaware of which visual task is solved by the subjects in their investigation: detection, resolution, discrimination, or recognition? This presents serious difficulties in correlating the results of measurements with the stimulus parameters and in choosing an adequate quantitative measure of visual acuity. In this paper we consider the problem of choosing the most practicable optotypes for monitoring visual acuity, taking into account both theoretical demands and the constraints determined by physiological, psychological and technical factors.

To characterise differences between different optotypes, the following questions need to be addressed:

- Which parameters of the test stimuli determine visual detection, resolution, discrimination, recognition, etc.?
- How does the final decision emerge from the results of analysing various stimulus parameters in visual pathways?
- Is the decision rule the same or different in different subjects?

It is rather difficult to answer these queries due to very complex organisation of image processing in the human visual system. In the first approximation, one can distinguish the following principal steps in any procedure of measuring visual acuity.

Before the measurements, a proper instruction is given to the subject: all the test stimuli selected for examination are presented and an examiner makes the subject aware about the required responses. The subject's visual system uses the information gained to find or create the conforming neural images (templates) in their memory.

At the next step, one of the test stimuli (chosen at random from the set used) is presented to the subject and the projected retinal image undergoes transformations at the first level of information processing. The outputs of this level can be considered as the primary signals that undergo processing at the elementary first level modules and transmitting to the second level modules selective to certain combinations of the primary signals. In its turn, at the second level, the outputs of the first level undergo processing and transmitting to the third level modules, and so on. In addition, there is need to consider feedbacks between different levels and between modules within the levels. As a result, a given stimulus can be represented in the visual system as a dynamic pattern of selectively activated pathways in the heterarchically organised neural net with feedforward and feedback connections.

It is evident that, in the case of complex test images, a complete analysis of image processing requires consideration of many neural pathways and voluminous calculations. Taking into account specific features of various optotypes, it is clear that each of them can excite its own specific combination of pathways, and, therefore, in a general case, it is not legitimate to consider the results obtained with different optotypes as representing one and the same visual capability. This point was realised by the thoughtful researchers long ago, and it was Pirenne who offered what might be the best formulation of this idea: "There are in fact as many 'visual acuities' as there are types of test objects" (Pirenne, 1962). Thus, the analysis of information processing during visual acuity measurement should include consideration of many functional modules at several levels of the human visual system. For this reason, different specific terms have come into use to distinguish different kinds of visual acuities measured by means of different optotypes and procedures. Popular classification implies the following four categories: detection, resolution, discrimination, and recognition. However, this classification cannot be considered as satisfactory. In view of a potentially infinite number of various complex test images, the possible combinations of the functioning modules can hardly be classified into four definite classes.

A lot of neurophysiologic and psychophysical data, as well as clinical examination of patients with visual brain injuries, provide evidence that local impairment can have very selective effect on stimulus recognition. For example, some visual deficits make it impossible to recognise letters (*alexia literalis*) or faces (*prosopagnosia*). This means that it is not always sufficient to say what kind of visual acuity the authors were trying to measure: it is important to add more information about the stimulus parameters, its semantics, and familiarity to the subject.

However, if visual acuity has to be measured for purposes like monitoring visual system development, early diagnostics, or controlling visual rehabilitation, it is reasonable to use certain universal test stimuli that are suitable for repeatable accurate examinations over a long period of time and must provide a possibility to notice small changes in visual acuity (either negative due to progression of a pathology or positive due to maturation or rehabilitation). It would be ideal if identical test stimuli could be used over the whole span of life including infancy and early childhood. Taking into account that the higher visual mechanisms gradually develop during many years and require learning, it is evident that the universal stimuli for monitoring visual acuity should be simple enough to be processed by the low-level visual modules functioning from infancy.

It is obvious that, at threshold, recognition of the test figures should be determined by their fine structure — the size of the smallest details (W) — but should not be based on indirect cues like asymmetry, difference in total area, horizontal/vertical proportions and other more rough features. The inadequacy of some test images in this respect are easily seen in the charts and slides for children and illiterate subjects (Fig. 1).

In terms of Fourier spectral analysis, the fine structure of the image is reflected mostly in the parameters of the high frequency part of its spectrum around the characteristic spatial frequency of the image, $F_c = 1/2W$, while more rough features reveal themselves in the low frequency parts of the spectrum. Ideally, recognition of the test stimuli should be based on one critical parameter — F_c. Among the test stimuli used for visual acuity assessment, there is only one type of stimuli that can be characterised by one critical parameter: the extended sine wave grating. Its spatial frequency can be used as a direct measure of visual resolution.

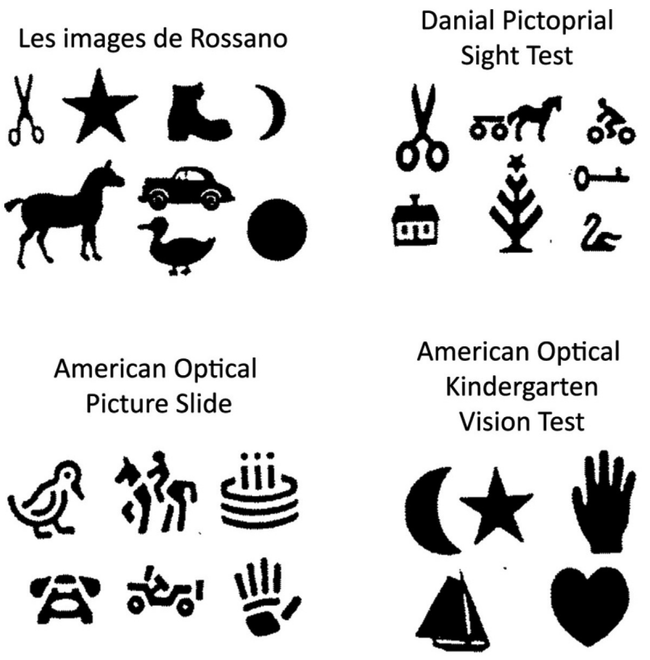

Fig. 1. Fragments of pictorial test sets for measuring visual acuity in children and illiterate subjects. Many optotypes can be distinguished using some indirect cues for recognition: asymmetry, difference in total area, horizontal/vertical proportions and other rougher features.

Since the extended sine wave gratings have serious disadvantages in practical usage, we decided to search for the more appropriate grating-style test stimuli among the commonly used and practicable ones and dwelled on the well-known 3-bar resolution target (USAF-1951; Anonymous, 1951). “The 3-bar target is a transition stimulus which blends the features of spatially extended sinusoidal gratings (with harmonically pure spectra) and spatially compact letters (with rich Fourier spectra)” (Anderson and Thibos, 1999a).

We performed several series of experiments to investigate visual processing of the standard 3-bar target and the two other types of simple stimuli similar to it in structure: slightly modified 3-bar target and the widely used tumbling-E optotype. Some results of these experimental investigations as well as the results of their theoretical analyses and modelling have already been published in Russian (Lebedev *et al*., 2010; Rozhkova *et al*., 2012; Lebedev, 2015). Here we describe here unpublished or very briefly published data (Rozhkova and Lebedev, 2010; Rozhkova, 2013; Rozhkova *et al*., 2014), and also include some previous results for completeness and logic of presentation. Some of these results appeared to be unexpected and crucial for understanding the essence of visual acuity measurements.

DETAILED CHARACTERISTICS OF THE OPTOTYPES INVESTIGATED

The important characteristics of the grating-style optotypes used in our comparative study are presented in Figures 2–5. The test stimuli were: standard (SV, SH) and modified (MV, MH) 3-bar resolution targets with vertical and horizontal bar orientation (Fig. 2, A and B) and tumbing-E images (Fig. 2, C). In MV and MH the bars were somewhat longer than in SV and SH. The reason for such modification of the standard 3-bar optotype is explained below. For all these optotypes, the characteristic frequencies F_c are determined by the width of the lines W, i.e. $F_c = 1/(2W)$. At first

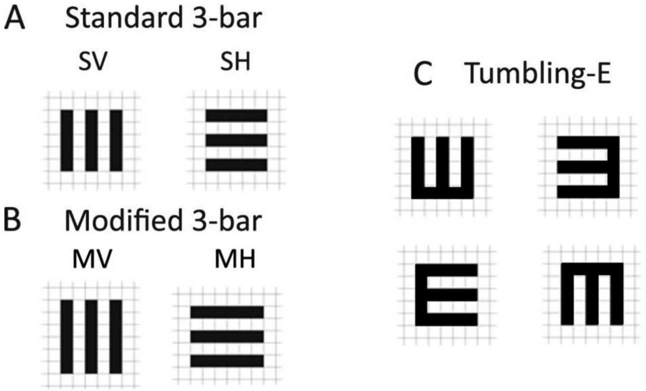

Fig. 2. The optotypes investigated in our work: standard (A) and modified (B) 3-bar targets and tumbling-E (C).

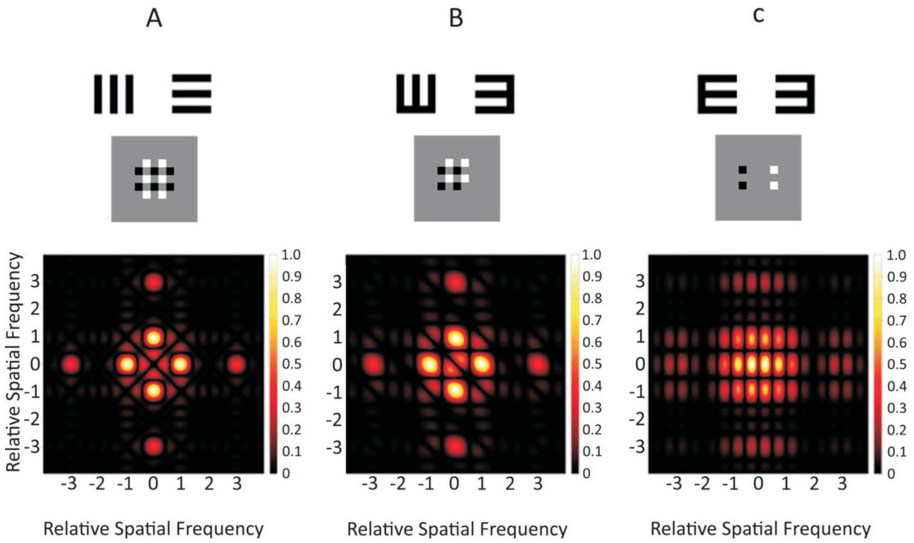

Fig. 3. Difference Fourier power spectra calculated for widely used optotypes: standard 3-bar targets (A) and tumbling-E (B, C). The frequency scales are normalised in relation to the characteristic spatial frequency F_c. The upper rows show the pairs taken for calculation; the intermediate rows show their luminance differences (grey background corresponds to zero value).

approximation, it seems natural to suppose that different simple optotypes having the same width of the lines W should be recognised equally easy (or hardly). This means that, supposedly, examination of one and the same subject using these three types of the test stimuli should give one and the same threshold size value, i.e. the same value of visual acuity. Experimental investigations revealed that this was not the case: the results obtained with different optotypes appeared to be different. Moreover, inter-individual differences indicated principally different ways of information processing in different subjects.

The following analysis based on the difference Fourier spectra might explain the cause of the discrepancy between anticipated and really obtained results.

The subject's task was to determine the orientation of the stimulus presented. To see which of the spectral components can contribute to the stimulus recognition, the difference spectra for each pair of the stimuli in the set can be calculated. Taking into account the specific features and symmetry of our simple optotypes, for such an analysis it was sufficient to calculate one difference spectrum for each pair of the 3-bar optotypes (standard and modified) and two difference spectra for the tumbling-E.

Figure 3, A–C shows two-dimensional difference power Fourier spectra for the standard 3-bar and tumbling-E pairs of images. Similar patterns are found in earlier papers of authors dealing with the same stimuli (Anderson and Thibos, 1999a; 1999b). These patterns contain many clusters of relatively powerful frequency components. The four brightest blobs in the difference spectrum calculated for the vertical and horizontal standard 3-bar stimuli (Fig. 3, A), as well as for the vertically and horizontally oriented tumbling-E stimuli (Fig. 3, B), are located on the horizontal and vertical axes at the positions (±1) corresponding to the characteristic frequency Fc. In the case of the left-right tumbling-E pair (Fig. 3, C), such composition of the four bright blobs is not expressed: in the central area within ±1 corresponding to F_c, the spectral power is distributed more evenly and the largest values are concentrated around the centre. These difference spectra indicate that, in each pair of the stimuli, recognition can be based both on the high-frequency and low-frequency (HF and LF) information and that, in the case of the left-right (and, by analogy, up-down) tumbling-E pair, the LF components can provide an essential contribution to the process of recognition. Regarding HF spectral components with frequencies significantly higher than F_c, they can be ignored at threshold: their peaks are situated at 3 F_c and farther, and therefore, they cannot be perceived because of the optical limitations of the human visual system since, near threshold, F_c reaches the upper limit of the optical transfer function (Campbell and Green, 1965).

In order to eliminate or reduce possibility of using LF information for the stimulus recognition, we modified the standard 3-bar stimuli, slightly elongating the bars (by 15–20%) in accordance with the results of our preliminary theoretical analysis and experiments. It was found that such a modification can provide elimination of LF components in the difference Fourier spectrum (Rozhkova *et al.*, 2012). For better illustration, in Figure 4, we show the amplitude but not the power difference Fourier spectra of the standard and modified 3-bar stimuli and have clipped the pictures excluding ineffective higher components in order to use a larger scale. In this figure, a neutral green colour means zero amplitude of the spectral components, while yellow-orange and blue blobs correspond to the well expressed sine wave components with the opposite phases.

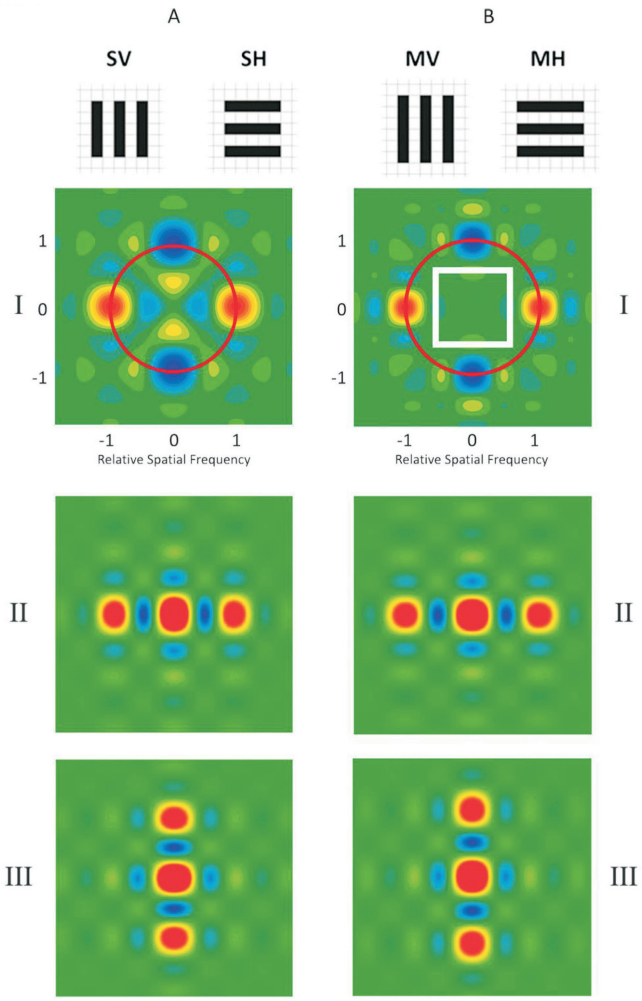

Fig. 4.

(I) Difference amplitude Fourier spectra calculated for the standard 3-bar stimuli (column A) and for the modified 3-bar (with the bars elongated by 16%) stimuli (column B). Red circular lines encompass the peaks of positive and negative values of the difference around the characteristic spatial frequency F_c. The white square frame indicates the area where LF components are almost absent.

(II) The initial amplitude Fourier spectra of SH (column A) and MH (column B).

(III) The initial amplitude Fourier spectra of SV (column A) and MV (column B).

Figure 4 contains not only the difference Fourier spectra for SH-SV and MH-MV but also the initial Fourier spectra of SH, SV, MH, MV. The difference Fourier spectra for SH-SV (Fig. 4, A, I) and MH-MV (Fig. 4, B, I) can be dis-

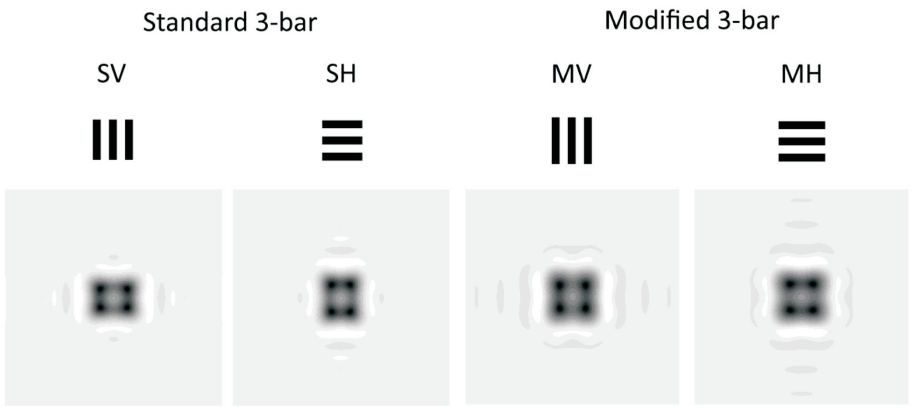

Fig. 5. Schematic illustration of contribution of LF-components to the perceived blurred images at resolution threshold. The images were calculated by means of inverse Fourier transform. Note the difference in the image orientations in the case of SV and SH (elongation across bars) and the absence of a similar difference in the case of MV and MH. The results of modelling (Lebedev, 2015).

tinguished rather easily by comparing their central areas: in the case of the modified stimuli, LF components are almost absent whereas, in the case of the standard stimuli, the composition of the four blobs is similar to the main composition at F_c, but having the opposite colours in the blobs of corresponding locations.

The effect of the described modification on the stimulus visibility at threshold is illustrated in Figure 5, which presents the results of modelling (Lebedev, 2015). This figure shows the blurred images calculated by means of inverse Fourier transform for the LF components of the standard and modified 3-bar stimuli. It is clearly seen that SV and SH produce images elongated across the bars and, therefore, they can be easily distinguished whereas MV and MH produce square images that are indistinguishable.

It should be particularly noted that the whole LF-images of SH and SV have orientations opposite to the orientations of the bars, i.e. blurred image elongation is orthogonal to the stimulus bars. Because of this feature, LF-information can be interpreted erroneously as indicating the opposite stimulus orientation: the response to SV would be SH and vice versa. However, there is also a possibility to use this LF-information properly taking into account the indicated orthogonality and thereby increasing probability of right responses to SV and SH. In the case of MV and MH, LF-information does not contain such cues of the stimulus orientation and, therefore, does not contribute to the stimulus recognition.

EXPERIMENTAL SERIES I: PSYCHOMETRIC FUNCTIONS FOR THE OPTOTYPES INVESTIGATED

From the above consideration, it follows that one can anticipate significant inter-individual variability of psychometric functions in the case of the standard 3-bar stimuli (S3B), since the LF-components of their Fourier spectra can be used for stimulus recognition and, probably, used differently by different subjects. In the case of the modified 3-bar stimuli (M3B), one might anticipate lesser variability of the psychometric functions, since the LF-components of their Fourier spectra cannot be used for stimulus recognition. Regarding the tumbling-E, their might be an intermediate degree of variability, taking into account rich patterns of LF-components in their difference Fourier spectra and, at the same time, the absence of clear evidence of their possible misuse. It seems likely, that, in most cases, the LF-information contained in tumbling-E can be used properly for stimulus recognition, thereby increasing probability of right responses and leading to some overestimation of the visual acuity in comparison with the M3B, which provides more correct values. Unfortunately, the direct comparison of the psychometric functions for the 3-bar and tumbling-E stimuli is problematic because of the difference in chance levels of the stimulus recognition: 50% for the 3-bar and 25% for the tumbling-E stimuli. However, this circumstance is not essential for the analysis of inter-individual differences.

METHODS

The psychometric functions for the S3B and M3B stimuli and tumbling-E were obtained by means of conventional procedures using orientation discrimination task. The aims and objectives of the tasks to be performed were fully explained to participants and consent was obtained. All procedures conformed to the tenets of the Declaration of Helsinki.

Subjects. The subjects were 15 adult volunteers aged 18–61 years; six of them were students (18–25 years). The subjects with refractive anomalies used proper glasses.

Stimuli. Original test charts containing S3B, M3B, and tumbling-E were created. Black images were printed on white paper using a professional printer HP LaserJet 1200 with 1200 dpi resolution. To avoid crowding effects, the distances between any adjacent optotypes were made not less than twice the size of the optotype. The level of the chart illumination was 1000 lx; the level of the ambient light in the room corresponded to 60-70 lx.

Procedure. The test charts with tumbling-E, S3B and M3B stimuli were presented in a quasi-random sequence at distance of 4 m. The subject viewed the chart binocularly. In each subject, 40–100 responses were collected for each point of the psychometric functions in the task of discrimination of the stimulus orientation. Several experimental sessions (2–4) were performed with each subject.

RESULTS

Paired comparisons of the psychometric functions for the optotypes investigated are presented in Figures 6 and 7. These data illustrate principal inter-subject variability: the psychometric curves for the S3B and M3B and the tumbling-E had different interposition in different subjects. Figure 6 shows the psychometric functions obtained in nine subjects for the S3B and tumbling-E. Each graph includes a pair of the psychometric functions belonging to one subject; the three columns present triplets of graphs belonging to the subjects with principally different interposition of the curves for the two optotypes. In the left column, the curves obtained for tumbling-E run lower than the curves for S3B type; in the central column, the interposition of the curves is inverted; in the right column, the pairs of the curves run together.

In the case of tumbling E, after reaching the chance level of 25%, the subjects usually refused to continue examination with the smaller stimuli saying that they could not see any difference between the stimuli presented. However, in the case of S3B stimuli, some subjects continued to respond readily after crossing the theoretical chance level (50%) and showed 70–90% of wrong responses. Such different behaviouvr of different subjects can be interpreted as follows: the subjects of the right column only relied on HF-information contained in both optotypes and did not use LF-information, while the subjects of the left and central columns certainly used it and differently. Proper usage of the LF-information contained in the S3B stimuli increased the probability of right responses (the left column), whereas misuse of this information increased the probability of wrong responses and sometimes, led to clear prevalence of the "opposite" responses resulting in paradoxical shape of the psychometric curves (central column). The most interesting cases of paradoxical psychometric curves obtained for the S3B optotype demonstrated that, at threshold, certain subjects can systematically give responses opposite to the right ones but not random.

Similar data were obtained when the S3B and M3B optotypes were compared. Figure 7 shows the individual data of three subjects with typical interpositions of the psychometric curves analogous to Figure 6. Here, one can also notice manifestations of proper use of LF information (I), misuse of this information (II), and ignoring it (III) in the case of S3B stimuli.

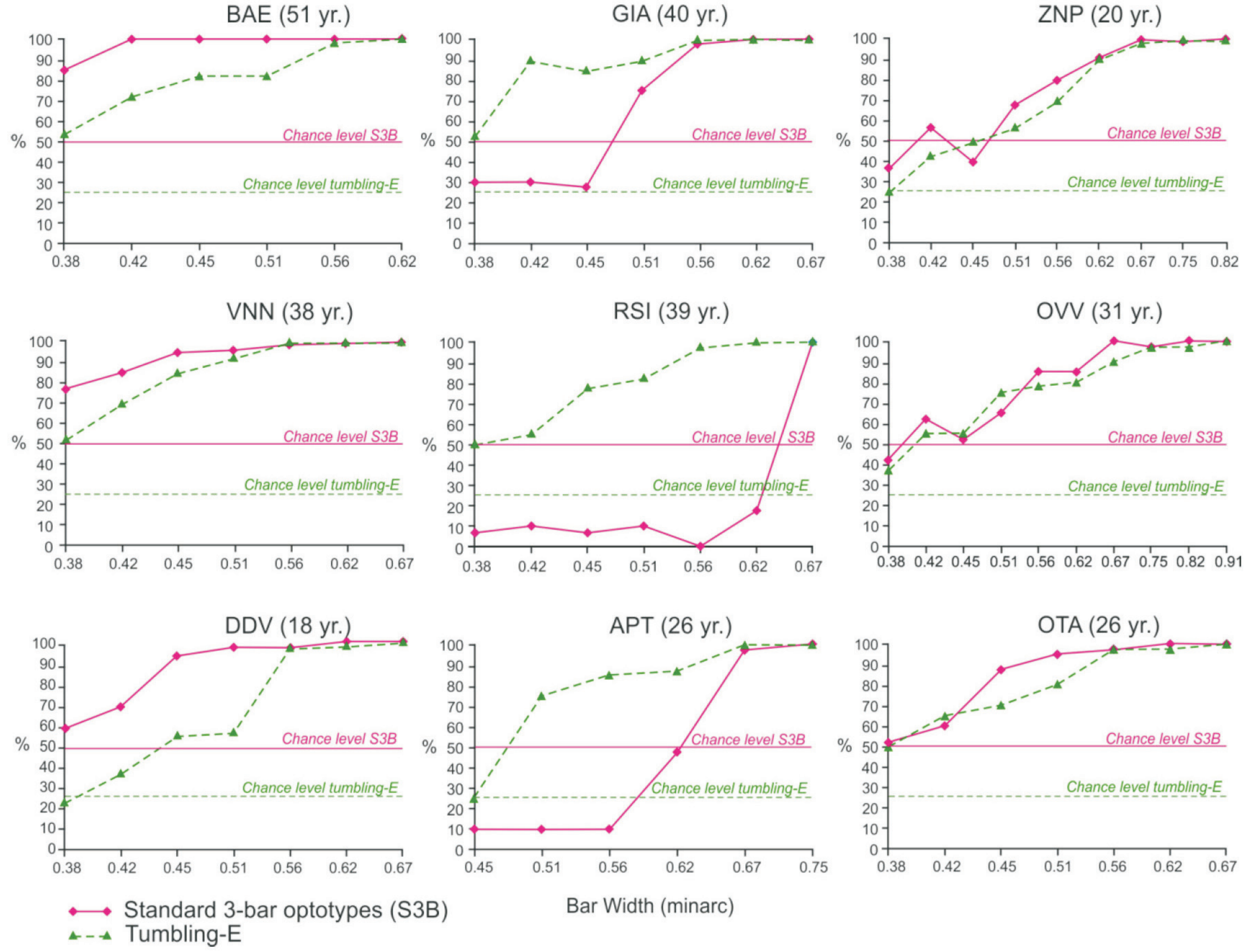

Fig. 6. Psychometric functions obtained in nine subjects for tumbling-E and S3B stimuli. The three vertical columns correspond to the three types of subjects with different interpositions of the curves for the two optotypes.

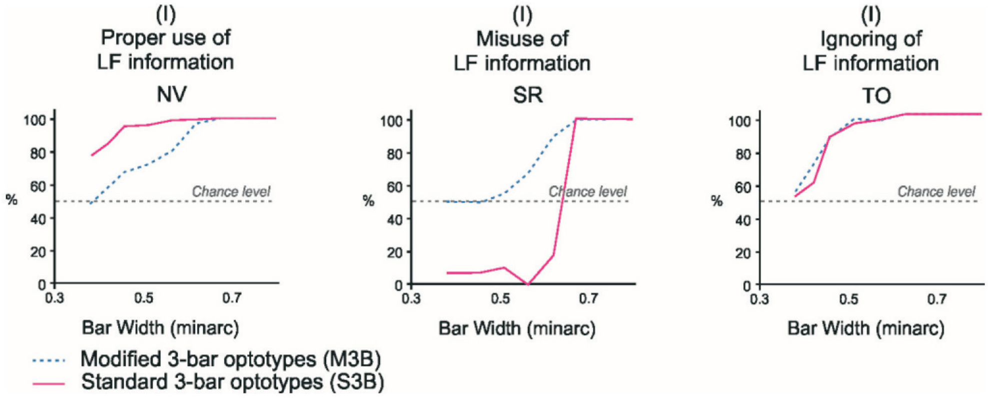

Fig. 7. Typical interpositions of psychometric functions for the S3B and M3B stimuli in subjects demonstrating proper use of low frequency information (I), misuse if this information (II), and ignoring it (III).

Thus, our experimental data on the psychometric curves for the three investigated optotypes indicated a possibility of an essential and individually varying influence of the LF-components on the psychometric curves and, therefore, on visual acuity measurements. This possibility was further investigated in the following two series of experiments.

EXPERIMENTAL SERIES II: THE DATA ON THE EFFECT OF LEARNING

A special series of experiments was devoted to study the influence of LF components of the optotypes on changes of visual acuity scores with repetition due to learning effects. It was supposed that such effects could manifest themselves as specific changes in the shapes of psychometric functions in repeated sessions.

METHODS

Subjects. The subjects were five naive young adults (21–23 yrs.) having no experience in psychometric experiments and one experienced subject (51 yrs.).

Stimuli. The charts used for obtaining psychometric function were the same as in the experimental series I.

Procedure. The psychometric functions for the S3B and M3B stimuli were obtained repeatedly with short time intervals (1–3 days). In this series of experiments, the data for psychometric functions were collected in conditions of presenting stimuli of the test chart through a rectangular window, line by line, starting from the smallest size.

The purpose of such an order was to eliminate a possibility to compare the orientation of the bars and the whole image shapes in well seen large optotypes on the same chart.

All procedures conformed to the tenets of the Declaration of Helsinki.

RESULTS

In Figure 8, the most demonstrative graphs are presented showing the effect of learning on recognition of the S3B and M3B stimuli in two naļve subjects.

The upper row contains three pairs of psychometric functions obtained in one and the same subject (MG) at the first, second and third successive experimental sessions. Taking into account the analysis presented above, it can be concluded that, initially, during the first session, this subject misused LF-information contained in S3B and, during the second session, the situation did not change: in both cases, the psychometric curves for S3B ran below the curves for M3B and crossed the chance level. There was even an increase in the efficiency of misusage: in the second session, the S3B curve was steeper than in the first one. However, at the third session, the subject suddenly understood that his interpretation of the threshold blurred image orientation was wrong and began to interpret it properly. Naturally, this led to an abrupt change of responses and inversion of the curve interposition indicating transition to proper usage of the LF-information due to some “insight” as a result of learning. Similar data (not presented in Figure 8) were also obtained in another naļve subject EF.

The lower row shows another typical case: from the very beginning, the subject AB was capable of using properly the LF-information contained in the S3B stimuli and this capability significantly improved with learning.

As for M3B, in the results of these three subjects, there was no evidence of dramatic changes and the data obtained did not allow to judge if there were any learning effects.

In the data of the nest two naļve subjects and in the data of the experienced subject there was no manifestation of learning, but the reasons appeared to be different. The experienced subject BA belongs to the type I of Figure 7 and demonstrated proper usage of the LF information, equally effective from the beginning to the end of the experimental series, while the two naļve subjects EK and NZ belong to the type III of Figure 7 and had not used LF information.

It is obvious that subjects are not always aware of the intimate details of visual processing in the course of image recognition. Changes in the way and the extent of using LF-information can occur unconsciously. Further, it is reasonable to think that, in a certain range of near-threshold stimulus sizes, visual image processing is inevitably unconscious and uncontrollable. Thus, the structural features of S3B and a possibility to compare the data obtained for the

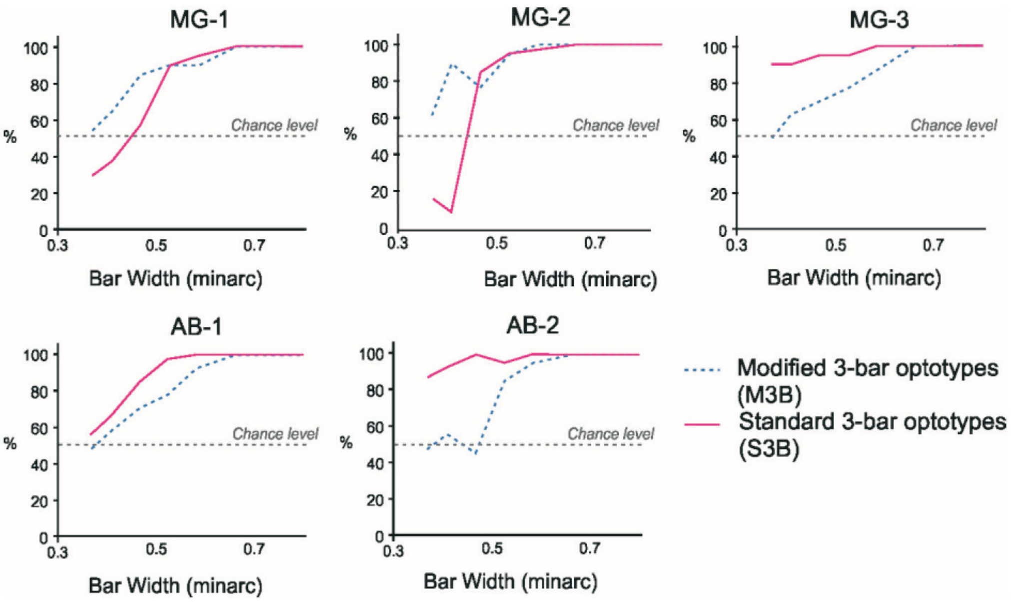

Fig. 8. Graphs showing the effect of learning on recognition of the S3B and M3B stimuli. The upper row: three pairs of psychometric functions obtained in one and the same naïve subject MG at the first, second and third successive experimental sessions. One can notice abrupt inversion of the curve interposition at the third session, indicating misuse of the LF information at the beginning and proper use of it later due to some "insight" as a result of learning. The lower row: the data of another subject obtained at two successive sessions; subject AB was capable to use the LF information properly from the very beginning but this capability significantly improved with learning.

S3B and M3B optotypes provided a chance to clarify the influence of the LF-information on the results of threshold measurements.

EXPERIMENTAL SERIES III: TEST-RETEST DATA FOR THE THREE OPTOTYPES INVESTIGATED

One of the best ways to compare the qualities of the optotypes is to assess test-retest reliability of the results provided by each optotype. Here, we present test-retest data obtained in healthy adults and in children with ophthalmopathology to demonstrate concordance of general conclusions regarding the optotypes investigated. In brief, our test-retest experiments showed significantly better reproducibility of measurements with M3B targets in comparison to two other optotypes: S3B and tumbling-E.

METHODS

Subjects. The testing was performed differently in two groups of subjects. The first group consisted of 17 adult volunteers (18–75 yrs., median — 38), and the second group consisted of 65 children (7–18 yrs., median — 13) with ophthalmopathology. The subjects used proper glasses if needed.

The purpose and the tasks to be performed were fully explained to participants and consent was obtained. All procedures conformed to the tenets of the Declaration of Helsinki.

Procedures. (1) In the group of adults, visual acuity was assessed in monocular viewing conditions by means of interactive software (elaborated by A. Terekhin) providing presentation of various optotypes (S3B, M3B and tumbling-E) and recording subject responses. The stimuli were generated on the display with pixel size 0.18 mm. The display luminance was 100 cd/m^2. The viewing distance was 6 m. Twelve sessions were performed with each adult subject: twice (test and retest) for each eye and each stimulus type (tumbling-E, S3B and M3B). The interval between test and retest sessions was variable (depending on subject fatigue), but enough for recovery.

(2) In the group of children with ophthalmopathology, visual acuity was assessed by means of our original printed test-charts containing the lines either with M3B or with tumbling-E stimuli. The test charts and conditions of measurements were the same as in the experimental series I. We excluded S3B from this clinical work to avoid subject fatigue. The steps in size were about 10% (approximately 0.5 logMAR). The viewing distance was 4 m. Eight sessions were performed with each subject: twice (test and retest) for each eye and for each stimulus type (tumbling-E and M3B). The interval between test and retest sessions in children group also was variable (depending on subject fatigue), but enough for recovery.

RESULTS

The average values of visual acuity obtained in the group of adults are presented in Table 1. Analysing this table and paying special attention to the right column (SD for difference retest-test), it can be concluded that M3B provided the best test-retest reproducibility of the visual acuity measurements. The tumbling-E optotype provided somewhat larger values of visual acuity than M3B in both sessions — this can be treated as overestimation of resolution capabilities in

Table 1

TEST-RETEST RESULTS OF MONOCULAR VISUAL ACUITY MEASUREMENTS (DEC. UNITS) IN ADULTS (17 subjects, 34 eyes)

Optotype	Mean ±SD, TEST	Mean ±SD, RETEST	Mean difference	SD for difference
Modified 3-bar	1.11 ± 0.23	1.12 ± 0.25	0.017	0.14
Standard 3-bar	1.08 ± 0.32	1.10 ± 0.27	0.024	0.24
Tumbling-E	1.21 ± 0.39	1.19 ± 0.32	–0.021	0.22

Table 2

TEST-RETEST RESULTS OF MONOCULAR VISUAL ACUITY MEASUREMENTS (DEC. UNITS) IN CHILDREN WITH OPHTHALMOPATHOLOGY (65 subjects, 130 eyes)

Optotype	Mean ±SD, TEST	Mean ±SD, RETEST	Mean difference	SD for difference
Modified 3-bar	0.66 ± 0.27	0.66 ± 0.27	0.00	0.017
Tumbling-E	0.61 ± 0.26	0.67 ± 0.27	0.06	0.074

the case of tumbling-E due to usage of LF-information. At the same time, despite a possibility to use LF-information, the average values for S3B appeared to be almost equal to those for M3B, supposedly, because the LF-information contained in S3B was used properly by some subjects and misused by the others, leading to neutralisation of the positive and negative deviations in the course of averaging. It's interesting that, in S3B and tumbling-E optotypes, the total negative effects of using LF-information on the test-retest reliability were quantitatively similar (see the right column).

The results of visual acuity measurements in the group of children with ophthalmopathology presented in Table 2 demonstrate an evident advantage of M3B concerning test-retest reliability. It is noteworthy that, in contrast to adults, lower values of visual acuity were obtained with tumbling-E than with M3B at the first session. This difference might be due to the difficulties for some children to name correctly the left and right configurations of tumbling-E. At the second session, the visual acuity scores (values) obtained with tumbling-E appeared to be higher than at the first one, evidently indicating the effect of learning. This tendency is also visible on the histograms of retest-test differences shown in Figure 9: note that, in the case of tumbling-E, the majority of the difference values are positive.

The analysis of the Tables 1, 2 and the histograms leads to the conclusion that M3B can provide better test-retest reproducibility than tumbling-E and S3B both in adults and children.

DISCUSSION

Visual acuity measurements with various optotypes. Our investigation presents direct evidence of principal inter-subject differences concerning the way of using low-frequency information contained in the widely used optotypes in addition to the main high-frequency information that, ideally, should determine the true thresholds measured in the course of visual acuity assessment. The low-frequency components of the stimulus Fourier spectra manifest themselves in a general shape of the whole image, its symmetry and proportions, providing some indirect cues for stimulus recognition, although it is supposed that recognition should be based on successful perception of fine stimulus details corresponding to the high-frequency components. Intuitively, researchers have understood the importance of eliminating indirect cues for test stimulus recognition and, in fact, the history of search for optimal optotypes reflects the progress in excluding such indirect cues. Gradually, it was also realised that different optotypes imply solving different visual tasks and activation of different nervous circuits in the visual brain whose structure appeared to be much more complicated than one could imagine at the onset of visual science.

Review of the literature available indicates that, on one hand, many various optotypes were proposed and used by ophthalmologists and visual scientists over time (Snellen, 1862; Green, 1868; Landolt, 1889; Sloan, 1959; Bennet, 1965; Baily and Lovie, 1976; Hyvarinen, 1980; Shelepin *et al*., 1985; 1987; 1992; Polat and Sagi, 1993; Plainis *et al*., 2007; Colenbrander, 2008; Koskin, 2009; etc.) but, on the other hand, the principal achievements were limited in number. The overall variety of the optotypes includes letters, numbers, geometrical figures, pictorial images in silhouette and contour realizations, periodic black-white and greyscale patterns (Fig. 10).

Starting from the Snellen optotype, one can create a not so long sequence of principally different optotypes trying to range their quality for measuring visual resolution (Fig. 11): standardised letters with serifs – standardised non-serif letters – Landolt rings – tumbling-E – 3-bar resolution targets – sine wave gratings – Gabor patches.

The main tendencies of modifications were: (1) simplifying the optotype shape, (2) increasing similarity of the optotype elements in the set, and (3) reducing the number of parameters that distinguish each optotype in the set from the others.

Despite clear evidence of a great diversity of the pathways that can be activated in the course of processing different optotypes, there still remains a belief that the terms "resolution acuity" ("grating acuity") and "recognition acuity"

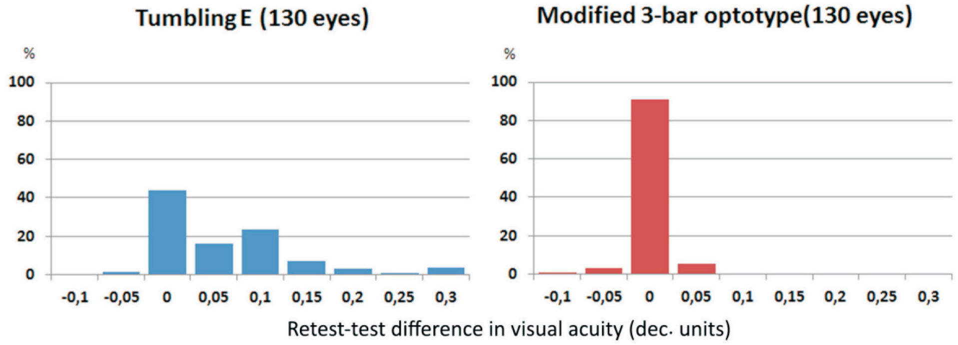

Fig. 9. Histograms showing differences in "retest-test" for tumbling E and M3B stimuli obtained for the group of 65 children with ophthalmopatology (130 eyes). Note that, in the case of tumbling-E, most differences are positive, i.e. visual acuity values in retest were larger than in the test, indicating the effect of learning.

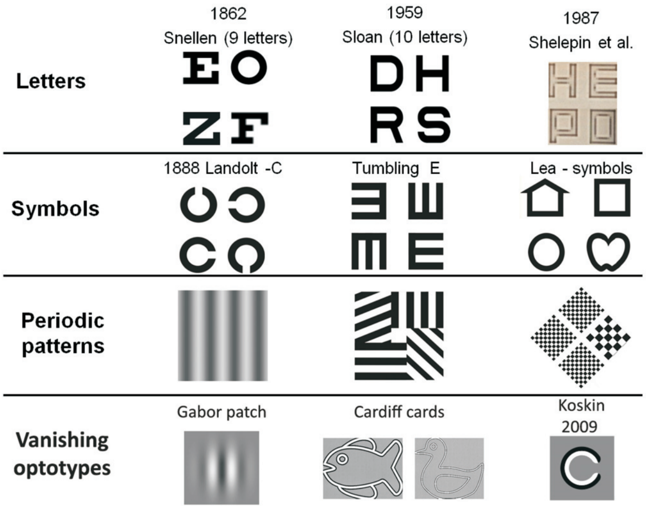

Fig. 10. Examples of various optotypes used in practice.

Fig. 11. Progressively improving optotypes regarding the assessment of visual resolution.

(“optotype acuity”) can be treated as certain unambiguous concepts (e.g., Wittich *et al.*, 2006; Stiers *et al.*, 2003; 2004). In a recent study (Heinrich and Bach, 2013) the authors analysed their data on measuring visual acuity with a set of Landolt-C-style optotypes including test stimuli with imbalanced and balanced luminance across potential gap locations. They concluded that “it is recognition acuity, rather than resolution acuity, which is measured with standard Landolt-style optoptypes, with the imbalanced luminance serving as a cue”, and that “luminance-balanced optotypes may help to obtain a more veridical estimate of resolution acuity”. That is true but it is not the whole truth. Firstly, it does not mean that the results obtained with other “recognition stimuli” (e.g., pictorial images, letters, Lea symbols, tumbling-E) will be the same as with standard Landolt-C. Secondly, any Landolt-style optotypes, including the luminance-balanced ones, are far from being optimal for measuring resolution acuity because of high proportion of low-frequency components in their Fourier spectra.

In fact, it is even not clear which of the stimulus parameters should be used as a critical one for quantitative characterisation and comparison of the Landolt-C stimuli (Bondarko and Danilova, 1997). There are two natural approaches to answer this question: the theoretical one, based on calculations, and the empirical one, based on clinical experiments.

The first approach suggests establishing the equivalence of the test stimuli (in view of their recognition difficulty) on the basis of certain characteristic spatial frequency F_c. The convention among the followers of this approach is to equate the size of the smallest detail W (e.g. the stroke width in the letter or the gap width in the Landolt-C) to half a period of a sine wave function with “characteristic frequency” $F_{c,}$ thus defining F_c as 1/(2W). This convention seems rather reasonable in many cases since the Fourier spectra of most optotypes reveal a prominent peak of power at F_c. However, this approach ignores the fact that, as a rule, the optotypes contain a broad spectrum of spatial frequencies above and below F_c, which may be important for recognition. This shortage was justly considered by Anderson and Thibos in investigation of relationship between acuity for gratings and tumbling-E letters (Anderson and Thibos, 1999a; 1999b).

The second approach (recommended by ISO; Anonymous, 1994, 2009) implies the employment of a specific calibration (correlation) procedure aimed at matching the investigated set of optotypes to the approved standard optotypes – Landolt Cs – concerning the probabilities of right responses.

It should be noted that Landolt C is far from appropriate for both approaches. The main two objections against Landolt C can be formulated as follows:

- The structure of the Landolt C, per se, is not easy to correlate theoretically with visual acuity: it is not clear what critical frequency determines optotype recognition in the case of such stimuli (Bondarko and Danilova, 1997).

- For many optotypes, the results of their matching with Landolt C will be essentially dependent on the subject’s age and experience.

Thus, it is apparent that in order to analyse the problem of visual acuity measurement one should start from less complex optotypes than Landolt C. That is why we did not use Landolt C in our experiments.

In our previous study (Rozhkova *et al.*, 2012), we found that even much more simple optotypes — the well-known standard 3-bar resolution targets — cannot be considered as the proper stimuli for an accurate visual acuity measurement. Taking into account the data obtained, we proposed to employ a slightly modified version of the standard 3-bar stimuli, which seemed to be more suitable for visual acuity monitoring (Lebedev *et al.*, 2012). In this paper, we briefly summarised the advantages of modified 3-bar optotypes revealed in our previous and new experiments on comparative investigation of modified 3-bar optotypes and other optotypes: (1) the uniformity of individual psychometric functions; (2) better test-retest reliability — reproducibility of visual acuity measurements; and (3) minimal effect of learning. The analysis of our own data and the relevant data available in literature allow us to formulate some general conclusions regarding the structure of any practicable optotypes designed for measuring visual acuity of any kind (detection, resolution, discrimination, recognition, etc.) and for any purpose (screening, monitoring, expertise).

The preliminary list of essential features required of good optotypes for assessment of visual acuity

1. Easiness of luminance matching for all optotypes in the set.

2. Structural uniformity (similarity of components, elementary parts): all the stimuli in the set should excite similar functional modules (detectors) at the periphery of the human visual system.

3. Gnosiological equivalence (involvement of the same higher brain centres in discrimination of all symbols in the set).

4. Equality of information content (to exclude *a priori* inequality in symbol significance and corresponding subject motivation).

5. High power of the Fourier spectral components around F_c — characteristic frequency reflecting the fine structure of the optotype.

6. Identical values of the characteristic frequencies F_c for all symbols in the set.

7. Minimal inter-stimulus differences in the low frequency domains of their Fourier spectra.

8. Equal probabilities of right responses to the test stimuli of equal sizes and identical arrays of error probabilities (probabilities of confusing each given stimulus with each another one in the set).

9. Steep psychometric function: for an accurate assessment of the threshold size, the probability of right responses should rise rapidly with increasing optotype size.

10. Easiness of computer generation, presentation, and printing (if necessary).

In addition to these general requirements, the tasks that are based on visual monitoring — following visual development, early diagnostics, assessment of treatment efficiency — impose rather heavy demands on the test stimuli. They must be: (1) suitable for patients of any age; (2) convenient for repeatable examinations; and (3) accurate enough for revealing the smallest physiologically significant changes of visual acuity.

From our investigation it follows that the prospective optotypes for visual acuity monitoring could be simple grating-like stimuli specially modified to exclude contribution of low-frequency Fourier spectrum components to the stimulus recognition.

ACKNOWLEDGEMENTS

The authors are grateful to A. P. Terekhin for creating the test software, to A. E. Belozerov for taking part in some earlier investigations, and to E. N. Kruttsova for technical assistance.
Supported by the Program III.3 of DNIT Russ Acad Sci.

REFERENCES

Anderson, R. S., Thibos, L. N. (1999a). The relationship between acuity for gratings and for tumbling-E letters in peripheral vision. *J. Opt. Soc. Amer. A,* **16**, 2321–2333.

Anderson, R. S., Thibos, L. N. (1999b). Sampling limits and critical bandwidth for letter discrimination in peripheral vision. *J. Opt. Soc. Amer. A,* **16**, 2334–2342.

Anonymous (1994). *ISO 8596. International Standard.* Ophthalmic optics. Visual acuity testing. Standard optotype and its presentation. Geneve. (2nd edition: Geneve, 2009).

Anonymous (1994). *ISO 8597. International Standard.* Optics and optical instruments. Visual acuity testing. Method of correlating optotypes. Geneve.

Anonymous (1951). USAF-1951. United States Air Force 3-bar resolution test chart.

Bailey, I. L., Lovie, J. E. (1976). New design principles for visual acuity letter charts. *Amer. J. Optom. Physiol. Opt.*, **53**, 740–745.

Bennett, A. G. (1964). Ophthalmic test types. A review of previous work and discussions on some controversial questions.*Brit. J. Physiol. Opt.*,**22** (4), 238–271.

Bondarko, V. M., Danilova, M. V. (1997). What spatial frequency do we use to detect the orientation of a Landolt C? *Vis. Res.*, **37**, 2153–2156.

Campbell, F. W., Green, D. G. (1965). Optical and retinal factors affecting visual resolution. *J. Physiol*., **181** (3), 576.

Colenbrander, A. (2008) The historical evolution of visual acuity measurement. *Vis. Impair. Res*., **10** (2–3), 57–66.

Green, J. (1868). On a new series of test-letters for determining the acuteness of vision. *Trans. Amer. Ophthalmol. Soc.*, **1** (4–5), 68–71.

Heinrich, S. P., Bach, M. (2013). Resolution acuity versus recognition acuity with Landolt-style optotypes. Graefes Arch. Clin. Exp. Ophthalmol., 251 (9), 2235–2241.

Hyvarinen, L., Nasanen, R., Laurinen, P. (1980). New visual acuity tests for pre-school children. *Acta Ophthalmol*., **58** (4), 507–511.

Koskin, S. A. (2009). The system of visual acuity measurements for medical expertise [Коскин, С. А. Система определения остроты зрения в целях врачебной экспертизы]. MD Thesis. St Petersburg. 48 pp. (in Russian).

Landolt, E. (1889). *Tableau d'optotypes pour la determination de l'acuite visuelle*. Societe Francais.

Lebedev, D. S. (2015). A model of orientation recognition mechanisms for the 3-bar two-grade optotypes [Модель механизма распознавания ориентации 3-полосных двухградационных оптотипов]. *Sensory Systems* [Сенсорные системы], **29** (4), 309–320 (in Russian).

Lebedev, D. S., Belozerov, A. E., Rozhkova, G. I. (2010). The optotypes for an accurate assessment of visual acuity [Оптотипы для точной оценки остроты зрения]. Patent № 2447826; 07.12.10 (in Russian).

Pirenne, M. H. (1962). Visual acuity. *The Eye*, **2**, 175–195.

Plainis, S., Tzatzala, P., Orphanos, Y., Tsilimbaris, M. K. (2007) A modified ETDRS visual acuity chart for European-wide use. *Optom. Vis. Sci*., **84** (7), 647–653.

Polat, U., Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vis. Res.*, **33** (7), 993–999.

Rozhkova, G. I. (2013). Visual acuity measurement: Account of the optotype structure. 36th European Conference on Visual Perception 2013. Bremen, Germany. 25–29 August. *Perception*, **42**, Suppl., 69.

Rozhkova, G. I., Lebedev, D. S. (2010). Is it rational to use Landolt C and Snellen E as the optotypes in the modern-day visual acuity measurements for early diagnostics? *Proceedings of the 1st World Congress on Controversies in Ophthalmology, Prague, Czech Republic, March 4–7*, p. 16.

Rozhkova, G. I., Belozerov, A. E., Lebedev, D. S. (2012).Visual acuity measurement: uncertain effect of the low-frequency components of the optotype Fourier-spectra [Рожкова, Г. И., Белозеров, А. Е., Лебедев, Д. С. Измерение остроты зрения: неоднозначность влияния низкочастотных составляющих спектра Фурье оптотипов]. *Sensory Systems* [Сенсорные Системы], **26** (2), 160–171 (in Russian).

Rozhkova, G., Lebedev, D, Gracheva, M., Rychkova, S. (2014). Advantages of employing specially modified 3-bar stimuli for visual acuity monitoring in adults and children: Test-retest reliability. 37th European Conference on Visual Perception, Belgrad, Serbia, 24–28 August. *Perception*, **43**, Suppl., 34.

Shelepin, Ju. E., Kolesnikova, L. N., Levkovich, Ju. I. (1985). *Visiocontrastometry* [Шелепин, Ю. Е., Колесникова, Л. Н., Левкович, Ю. И. *Визоконтрастометрия.*] Nauka, Leningrad. 103 pp. (in Russian).

Shelepin, Ju. E., Glezer, V. D., Bondarko, V. M., Pavlovskaja, M. B., Vol, I. A., Danolov, Ju. P. (1992) Spatial Vision [Шелепин, Ю. Е., Глезер, В. Д., Бондарко, В. М., Павловская, М. Б., Вол, И. А., Данилов, Ю. П. Пространственное зрение] In: Byzov, A. L. (Ed.). *Vision Physiology* [*Физиология зрения.* Под ред. А. Л. Бызова]. Nauka, Moscow, pp. 528–585. (in Russian).

Shelepin, Ju. E., Volkov, V. V., Makulov, V. B., Kolesnikova, L. N., Pauk, V. N. (1987). Measuring of functional possibilities of human visual system [Шелепин, Ю. Е., Волков, В. В., Макулов, В. Б., Колесникова, Л. Н., Паук, В. Н. Измерение функциональных возможностей зрительной системы человека]. *Academy of Sciences Messenger, USSR* [Вестник АН, СССР], **9**, 63–72 (in Russian).

Sloan, L. L. (1959). New test charts for the measurement of visual acuity at far and near distances. *Amer. J. Ophthalmol*., **48** (6), 807–813.

Snellen, H. (1862). *Test-types for the Determination of the Acuteness of Vision*. P. W. van de Weijer, Utrecht. 44 pp.

Stiers, P., Vanderkelen, R., Vandenbussche, E. (2003). Optotype and grating visual acuity in preschool children. *Invest. Ophthalmol. Vis. Sci.*, **44**, 4123–4130.

Stiers, P., Vanderkelen, R., Vandenbussche, E. (2004). Optotype and grating visual acuity in patients with ocular and cerebral visual impairment. *Invest. Ophthalmol. Vis. Sci.*, **45**, 4333–4339.

Wittich, W., Overbury, O., Kapusta, M. A., Watanabe, D. H. (2006). Differences between recognition and resolution acuity in patients undergoing macular hole surgery. *Invest. Ophthalmol. Vis. Sci*., **47**, 3690–3694.

Received 8 November 2016
Accepted in the final form 17 July 2017