The impact of multimodal and polyvalent audiovisual stimuli on the emotional state of an individual

line 1: Alexander Y. Petukhov
line 2-3: *ANO «Neymark»*
line 4: Nizhni Novgorod, Russia
line 5: https://orcid.org/0000-0002-7412-5397

line 1: Nikita S. Morozov
line 2: *Laboratory of digital radio engineering*
line 3: *ANO «Neymark»*
line 4: Nizhni Novgorod, Russia
line 5: nsmorozov@rf.unn.ru

line 1: 3rd Olga Khaldina
line 2: *Department of Psychophysiology*
line 3: *Lobachevsky Nizhny Novgorod State University*
line 4: Nizhni Novgorod, Russia
line 5: hurrycane.1905@gmail.com

line 1: Nikolay V. Krasnitskiy
line 2: *Department of Psychophysiology*
line 3: *Lobachevsky Nizhny Novgorod State University*
line 4: Nizhni Novgorod, Russia
line 5: nvkrasnitskiy@gmail.com

line 1: Sofia A. Polevaya
line 2: *Department of Psychophysiology*
line 3: *Lobachevsky Nizhny Novgorod State University*
line 4: Nizhni Novgorod, Russia
line 5: s453383@mail.ru

line 1: Ivan V. Loskot
line 2: *Department of Psychophysiology*
line 3: *Lobachevsky Nizhny Novgorod State University*
line 4: Nizhni Novgorod, Russia
line 5: ivan@loskot.ru

***Abstract* — In this study, we study the features of the influence of various audiovisual stimuli on the emotional state of an individual. Special attention is paid to the peculiarities of the formation of subjective emotional images that arise as a result of interaction with multimodal (sounds, images) and polyvalent (positive/negative valence) affective audiovisual stimuli from the international regulatory framework IAPS and IADS.**

Keywords— emotional perception; information images; audiovisual stimuli; IAPS; IADS; SAM.

I. Introduction

The modern world with overflowing information causes year over year increase in stress and emotional negativity of the population, which significantly affects both psychological and physiological states of individuals. This is facilitated by external factors (pandemics, wars, migration crises, etc.) along with expanding social pressure due to complexity of informational networks and processes relevant for each human and a growing number of corresponding external incentives. That's why the challenges of personalized emotional state diagnostics, including phenomena such as emotional maladjustment, increased anxiety, etc., have become ever-so acute. Given the vast scale and prevalence of the aforementioned phenomenon across the population, this kind of diagnostics should be accessible, easy to implement and en-sure sufficient accuracy.

There is an immense number of available approaches to determining emotions such as the SAM (Self-Assessment Manikin) questionnaire [1-2]. Such approaches are linking physiological and self-reported indicators. Oftentimes these studies use images as stimulus material, and in some cases, they resort to sound stimuli (music, sounds of nature, etc.) [3-8].

One of the distinctive aspects of human cognitive activity is the fact that individuals do not think in codes (similar to a computer); our mind operates with multiple information images/representations (further - IR). These images have a very specific physical structure based on their functional core (namely, electrical and chemical activity in the human brain). However, their description from the point of view of conventional mathematical models is difficult for a number of reasons [9-12].

A separate issue for comparative research (experimental and model description) is the interaction of images that are in active state, i.e. how much the simultaneous activation of two or more IIs affects the perception of each individual one.

II. Experimental Methods of Emotional Diagnostics

An important parameter in psychological diagnostics is ambivalence, which allows researchers to keep the subject focused on the experiment without any distractions and simultaneously collect reliable data. One of these methods is the analysis of heart rate variability. A heart rate variability study includes tracking the heart rate, the average du-ration of the RR interval, the number of RR-intervalpairs, and the power of the high-frequency band [2, 13-14].

Currently most researchers use the IAPS (developed by P. Lang, consists of 1182 im-ages) and IADS (developed by M. Bradley, the second edition consists of 168 sound com-positions) databases to study the effects of auditory and visual analyzers [15-16].

Besides, the validity of subjective reports is strictly limited in time: the longer the time between the assessment of the state and the emotional event, the less its reliability. Therefore, the study of the emotional sphere of a person should rely on the objective methods for assessing emotions obtained both from the central (EEG) and from the autonomic nervous system (ECG, GSR, respiration). With an integrated approach to research of emotions, special attention should be paid to the assessment of the behavioral reactions of the subject.

In that case the cognitive component of the assessment of emotions is based on as-sessing the stimulus and individual's state both before and after exposure to the stimulus. These databases have proven to be effective for studying HRV and emotional manifestations triggered by stimuli of different valences. Therefore, we can assume that HRV can represent the occurrence of emotions or even their valence. It can be quite challenging for subjects to describe their emotions due to semantics (incomplete understanding of meaning) or lexis (not knowing how to describe the mani-fested emotion), etc. That's why it is better to determine the actual genesis of

emotions of individuals using standardized express methods based on pictographic design, which simplifies the correlation of the actual state of the subject and the parameters of the meth-od (scale).

III. Experimental Study

The research hypothesis was as follows: depending on the order of presentation of stimuli, as well as in case of adding new modalities, significant differences in the assessment of the emotional perception of the original stimulus may be observed.

227 people were selected as subjects: men and women from 17 to 58 years old. All subjects were examined to ensure they had no problems with cardiovascular, psychiatric, or respiratory diseases, and that they did not take any medications. All subjects were informed and signed an informed consent to participate in the study. As stimulus material for the experiment, 24 images (12 extremely positive and 12 extremely negative according to the SAM valence scale) were selected from the international IAPS database and 24 sound stimuli (12 extremely positive and 12 extremely negative according to the SAM valence scale) were chosen from the international IADS database.

These visual and audio effective stimuli were randomly formed into specific blocks:

- Negative sound (S-) (12 negative sounds);
- Positive sound (S+) (12 positive sounds);
- Negative Image (P-) (12 negative images);
- Positive Image (P+) (12 positive images).

The subjects were asked to listen/watch the block and then evaluate it using the SAM technique. In addition, 4 audiovisual stimulus sets were compiled from the selected stimuli:

- Negative video (image + sound) (P-S-) (12 negative sounds and 12 negative images);
- Positive video (image + sound) (P+S+) (12 positive sounds and 12 positive images);
- Dissonant video 1 (image + sound) (P-S+) (12 negative images and 12 positive sounds);
- Dissonance video 2 (image + sound) (P+S-) (12 positive images and 12 negative sounds).

These stimuli were grouped into 8 pairs of combinations / sessions:

№	Stimulus 1	Stimulus 2
1	P+	P+S-
2	P-	P-S+
3	S+	P-S+
4	S-	P+S-
5	P+S-	P+
6	P+S-	S-
7	P-S+	P-
8	P-S+	S+

The exposure time was 6 seconds (in accordance with the playing time of one sound stimulus from the IADS database). After demonstrating each stimulus in combination, the subjects were assessed using the SAM method on three scales: valence, arousal, dominance.

There was a pause of 20 minutes between the demonstrations of the first and second stimuli within the combination. The sessions were measured on different days to exclude the factor of interference from other stimuli on the individual's emotional perception.

The study was conducted in a room isolated from external noise, with a comfortable workplace, in accordance with the basic requirements for conducting psychological testing.

IV. Results

Based on the ratings on three SAM scales (valence, arousal, dominance) for each pair of stimuli within a combination for all 8 sessions, the total absolute values of the difference in ratings were calculated. Within each session, basic statistics (mean, standard deviation, median, and mode) were calculated for all assessments.

To assess significant differences in comparison of study blocks, the paired Wilcoxon W test was used, $p<0.05$ (Table 1). As the next step, the average values of ratings on the valence scale for multimodal blocks (P-S- and P+S+) were compared. The combination of a negative image and sound led to an effect of enhancing the evaluation of negative stimuli. However, the negative sound had a greater effect than the negative image. No significant differences were found between positive stimuli. However, when comparing variations of multimodal stimuli from the basic and dissonant sets, significant differences were obtained (Table 2).

Based on assessments of the valence of the dissonant set, a tendency of subjects to negatively evaluate stimulus blocks was noted. Regardless of the type of presentation of bimodal stimuli in a given set, equal negative values were obtained in the sample means. Obviously, the addition of any negative information image (visual or auditory) inclines the subject in a subjective assessment to a negative attitude towards the environment.

Next, we compared the ratings on the valence scale between the first stimuli in combinations of unimodal and monovalent stimuli with their combinations into single multimodal videos. The subjects who gave at least one stimulus a rating that did not correspond to its valence obtained from the average values during stimulus validation (e.g. rating a P+ stimulus (positive image) as negative or neutral) were eliminated. The share of such estimates in a variety of combination options is significant – 47%.

It was further found that unimodal sounds and images were rated at the same level according to valence. The combination of these components in a single video led to a negative assessment of all stimuli within the integrative image.

As a result, it was possible to identify the effect of the dominance of a negative emotional image over a positive one. When integrating at least one negative unimodal component into the structure of a positive image, a reduction in assessment occurs.

TABLE I.

Pairs for comparison	Valid	T	Z	p-value
Sbad+Pbad & Sbad+Pgood	11	3.00	2.67	0.01
Sbad+Pbad & Sgood+Pbad	21	9.00	3.70	0.00
Sbad+Pbad & Sgood+Pgood	13	9.00	2.55	0.01
Sbad+Pbad & PBad	11	0.00	2.93	0.00
Sbad+Pbad & SBad	12	10.50	2.24	0.03
Sbad+Pbad & Pgood	15	0.00	3.41	0.00
Sbad+Pbad & Sgood	15	0.00	3.41	0.00
Sbad+Pgood & Sgood+Pbad	15	4.00	3.18	0.00
Sbad+Pgood & PGood	13	2.00	3.04	0.00
Sbad+Pgood & SGood	14	1.50	3.20	0.00
Sgood+Pbad & Sgood+Pgood	14	4.50	3.01	0.00
Sgood+Pbad & PBad	15	2.00	3.29	0.00
Sgood+Pbad & SBad	15	1.00	3.35	0.00
Sgood+Pgood & PGood	13	7.50	2.66	0.01
Sgood+Pgood & SGood	13	0.00	3.18	0.00
PBad & PGood	14	2.00	3.17	0.00
PBad & SGood	15	3.00	3.24	0.00
SBad & PGood	15	0.00	3.41	0.00
SBad & SGood	14	0.00	3.30	0.00

TABLE II.

stimulus	Image valency		
Sbad+Pbad	-2.27	±	0.38
Sbad+Pgood	-1.20	±	0.49
Sgood+Pgood	1.45	±	0.36
Sgood+Pbad	-1.20	±	0.49

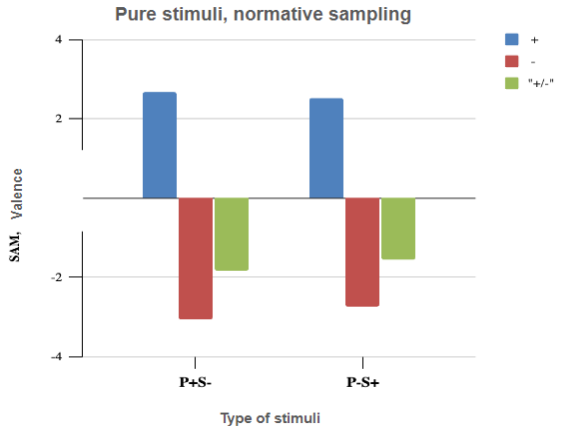

Fig. 1. Average values on the valence scale for the sample with standard answers. Multimodal combinations with color display of the valence of each component and combination are noted.

Next, an analysis was carried out of the second stimuli in connection with the prior ones displayed before them. In combinations where the demonstration of a negative image came first and the positive image came second, its rating decreased and approached zero. The described effect was associated with the order of content presentation: the leading negative element reduces the positive element of the image.

When compared with other variants of combinations, where the first stimulus was positive, and then came a negative one, it was revealed that there was no influence of order on the assessment of negative content. The negative image is resistant to the positive one that outstrips it and does not lead to a significant change in the assessment).

As we compared ratings by magnitude by modality, it was revealed that negative unimodal stimuli received the highest ratings (Figure 2). It is not possible to accurately assess the reason for this effect at this stage of research. There is always a chance of dependence on exposure time.

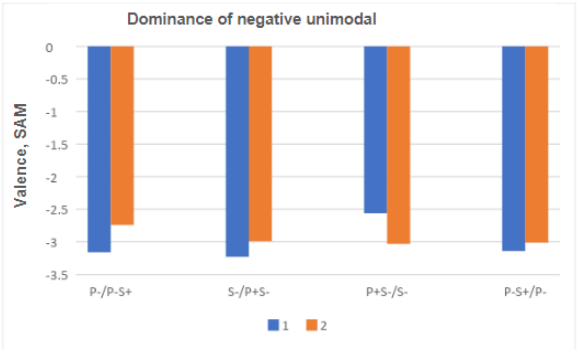

Fig. 2. Comparison of data means with standard scores among combinations with a unimodal negative stimulus.

To test the effect of the predominance of a negative emotional image over a positive one, the experimental model of the main study was tested with the demonstration of 3-minute videos. In this case, each subject was asked to preliminary evaluate, using the SAM method, each stimulus that was part of the video (12 negative and 12 positive sounds and the same number of valence images). The exposure time was 6 seconds, consistent with the duration of the IADS stimuli. The stimuli were presented in a random order.

As the result, the effect of "accumulation of emotions" was revealed: with the sequential presentation of a set of univalent affective stimuli, an integrative emotional image is formed, which corresponds to the average sample valence in sign and is greater in modulus for assessment on the valence scale. This effect occurs for negative and positive affective stimuli.

A linear relationship ($y = 1.7027x + 0.1267$) was revealed between the valence of the integrative emotional image and the average sample valence of affective stimuli ($r2 = 0.88$) (Fig. 3). Black indicates the expected trend line if the average score corresponded to the same integrative one.

Trend line in linear form according to the function $y = 1.7027x + 0.1267$ with $R^2 = 0.8841$.

When a set of differently valent stimuli is presented, there is no linear relationship and the effect of "accumulation of emotions" does not appear. There is a dominance of the negative image (Fig. 4).

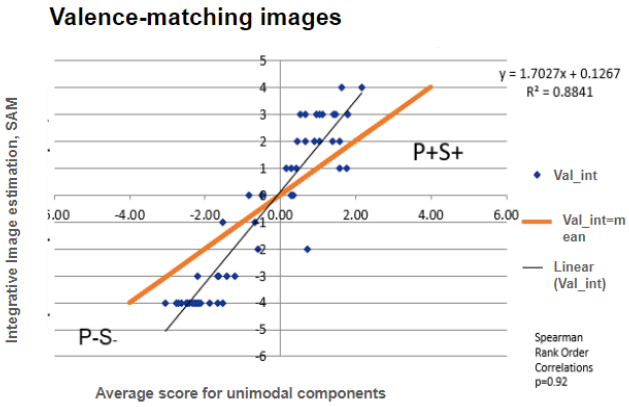

Fig. 3. Phase space according to integrative image SAM valence estimates with the subject's average rating for unimodal components that were included in the integral negative and positive videos.

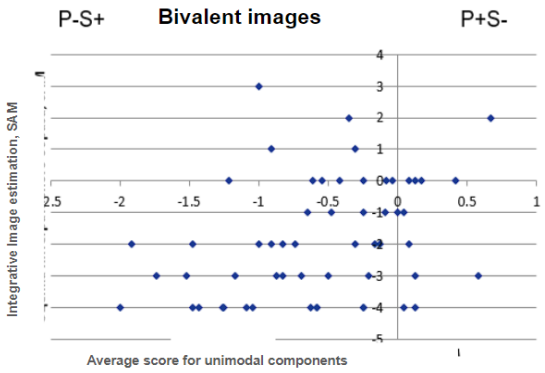

Fig. 4. Phase space according to the integrative images SAM valence estimates with the subject's average rating for unimodal components that were included in the integral bivalent videos.

Here it is also important to note the specificity and nonlinearity of the effect of the influence of the bimodal component on the state of the individual, taking into account the effect of "accumulation of emotions." This is due to the fact that activated information images influence the individual's perception of each other, distorting both quantitative and qualitative values of the impact. This leads to interesting effects, which are well used by masters of the visual industry in cinema, computer games, etc. [17].

V. Conclusions

The research hypothesis that the assessment depends on the emotional perception of the original stimulus and the order of presentation of stimuli, as well as on the addition of new modalities, was confirmed. The authors identified the following features:

- The predominance of a negative information image associated with emotions over a positive one. If the structure of a positive image includes at least one negative unimodal component, then a decrease in assessment occurs.

- The evaluation of a positive stimulus depends on the preceding stimuli. A negative stimulus preceding a positive one reduces the evaluation of the positive element of the image, while a negative image is resistant to the anticipating positive one.

- There is an increase in the evaluation of a negative univalent unimodal stimulus, regardless of the preceding stimulus. The exact reason for this effect cannot be determined by comparison with the bivalent stimuli from previous phases of the study, since the mean ratings in this phase differ from those obtained in combinations of bivalent stimuli. This may be due to the dependence on exposure time.

- The effect of "accumulation of emotions" was revealed: with the sequential presentation of a set of univalent affective stimuli, an integrative emotional image is formed, which corresponds to the average sample valence in sign and more in modulus for assessment on the "Pleasure" scale. This effect occurs for negative and positive affective stimuli. A linear relationship ($y = 1.7027x + 0.1267$) was revealed between the valence of the integrative emotional image and the average sample valence of affective stimuli ($r2 = 0.88$). When a set of differently valent stimuli is presented, there is no linear relationship and the effect of "accumulation of emotions" does not appear. There is a dominance of the negative image.

The research was supported by a grant from the Russian Science Foundation (project No. 22-18-20075)

References

[1] Pavlov, S.V.; Reva, N.V.; Loktev, K.V.; Korenek, V.V.; Aftanas, L. Effect of meditation practice on cardiovascular response in the perception and cognitive reevaluation of emotiogenic stimuli of emotion. Russ. J. Physiol. 2015, 101, 360–373.

[2] Vasanov, A.Y.; Marchenko, O.P.; Sevostyanova, M.S. Selection of culturally specific emotionally colored photographs for experimental research. Exp. Psychol. 2013, 6, 105–114.

[3] Kozachuk I.V., Kirillova I.A., Vedernikova T.N., Simonov S.N. Effects of audio stimulation in the form of nature sounds on the regulation of heart rate. Bulletin of TSU. 2012; 17 (4); 1270-1272

[4] Matokhina, A.A. Study of the influence of classical music on the functional state of people of various professions. Faces Knowl. 2013, 2, 69–72.

[5] Efremova, G.I.; Kovaleva, M.A.; Bochkovskaya, I.A.; Novikova, G.V. Possibility of using the method of analysis of thermographic images in the diagnosis of emotional and stressful states. Personal. Educ. 2016, 4, 181–187.

[6] Podladchikova, L.N.; Koltunova, T.I.; Shaposhnikov, D.G.; Lomakina, O.V. Individual characteristics of the examination of emotionally significant images. Russ. J. Physiol. 2016, 102, 618–627.

[7] Kwang-Ho Choi, Junbeom Kim, O. Sang Kwon, Min Ji Kim, Yeon Hee Ryu, Ji-Eun Park Is heart rate variability (HRV) an adequate tool for evaluating human emotions? - A focus on the use of the International Affective Picture System (IAPS). Psychiatry Research. 2017;. 251; 192-196; https://doi.org/10.1016/j.psychres.2017.02.025

[8] Smolyaninova, S.V. Investigation of the relationship between the degree of extraversion and the effectiveness of the use of various musical programs in the correction of the mental state of operators. Leningr. State Univ. Bull. 2017, 2, 126–134.

[9] Anokhin, K.V. Gene Probes for Mapping Neural Networks during Training. Principles and Mechanisms of Human Brain Activity; Nauka: Leningrad, USSR, 1989.

[10] Petukhov A.Y. and Polevaya S.A. Modeling of cognitive brain activity through the Information Images Theory in terms of the bilingual Stroop test. International Journal of Biomathematics. 10(6) (2017) 1750092 DOI: 10.1142/S1793524517500929

[11] Petukhov A.Y., Petukhov Yu.V. Model of cognitive activity of the human brain based on the mathematical apparatus of self-oscillating quantum mechanics. Mathematics 2022, 10 (22), 4215; https://doi.org/10.3390/math10224215

[12] Petukhov, A.Y.; Polevaya,S.A.; Polevaya, A.V. Experimental Diagnostics of the Emotional State of Individuals Using External Stimuli and a Model of Neurocognitive Brain Activity. Diagnostics 2022, 12, 125. https://doi.org/10.3390/diagnostics12010125

[13] Kosonogov V.V., Martinez-Selva J.M., Sanchez-Navarro J.P. A review of modern methods for measuring physiological signs of the sign and strength of emotional states. Theoretical and experimental psychology., 2017; 10 (3); 63-78

[14] Edward F., Pace-Schott, E. F., Amole, M. C., Aue, T., Balconi, M., Bylsma, L. M., Critchley, H.,Van Elzakker, M. B Neuroscience and Biobehavioral Reviews. ELSEVIER. 2019; 103; 267-304; https://doi.org/10.1016Zj.neubiorev.2019.05.002

[15] Bradley M. M., Lang P. J. The International Affective Digitized Sounds (2-nd Edition; IADS-2): Affective Ratings of Sounds and Instruction Manual Gainesville, FL: The Center for Research in Psychophysiology. University of Florida. 2007; P. 29-46

[16] Lang, P. J., Bradley, M. M., & Cuthbert, B. N. International affective picture system (IAPS): Instruction manual and affective ratings. Technical Report A-4. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida 1999.

[17] Bogachev, A. M. Horror films and nightmares: some aspects of psychological analysis of phenomena. Psychology and psychotechnics. 2019. 4, 39-49.DOI 10.7256/2454-0722.2019.4.31301