Neuroplasticity in economic decision making under active choice

A.N. Shestakova
Institute for Cognitive Neuroscience
HSE University
Moscow, Russia
a.shestakova@hse.ru

Grigory Kopytin
Institute for Cognitive Neuroscience
HSE University
Moscow, Russia
kopytin.kg@yandex.ru

Alla Simova
Institute for Cognitive Neuroscience
HSE University
Moscow, Russia
akondratenko@hse.ru

***Abstract*—Neuroeconomics, an emerging field blending neurobiology and economic theory, investigates decision-making and reinforcement learning. Grounded in early Russian and Soviet research, this field examines how dopaminergic circuits assess and choose favorable options based on sensory, motivational, and cognitive inputs. Traditional decision-making models often overlook perceptual changes linked to reinforcement learning. Modern theories suggest that sensory cortex adaptability, influenced by reinforcement, plays a critical role in decision-making.**

Our study aims to integrate sensory plasticity into neuroeconomic models, providing a comprehensive understanding of how experiences and repeated behaviors affect sensory cortex reorganization and decision-making. We developed an audio version of a lottery-like one-armed bandit task to explore the dynamics of reinforcement learning and its relationship with neuroplasticity markers (MMN, P3a, and brain oscillatory activity).

The study involved 29 participants who chose between two audio-encoded options with different reward probabilities. EEG data were recorded and analyzed using custom Python scripts and the MNE library. Artifact removal and segmentation were performed, focusing on the Cz electrode due to its prominent amplitude expression.

A repeated measures ANOVA revealed significant differences in responses to standard and deviant sounds before and after the game, and between sounds encoding large and small losses. These findings highlight the role of sensory neuroplasticity in decision-making and underscore the importance of incorporating sensory changes in reinforcement learning models. This study enhances our understanding of the interaction between sensory input and reinforcement learning in shaping behavior.

***Index Terms*—one-arm bandit task, P2, decision-making, neuroplasticity**

I. Introduction

Neuroeconomics, a relatively new interdisciplinary field at the crossroads of neurobiology and economic theory [1], emerged from foundational research in Russian and Soviet psychophysiology. Pavlov's theories on conditioned reflex activity and Anokhin's functional systems theory have laid the groundwork for contemporary studies into decision-making and the neurobiology of reinforcement learning [2]. Modern neuroeconomics emphasizes a mechanistic model where specialized dopaminergic neural circuits evaluate options and select the most favorable one based on various sensory, motivational, and cognitive inputs [3].

Key areas of current research in neuroeconomics include understanding how stimulus-reinforcement associations form, analyzing the dynamics of such learning, and assessing its impact on behavior and physiological responses. [4, 5] Studies have shown the critical role of dopaminergic neurons in decision-making and behavioral adjustments. Additionally, cognitive research has highlighted the adaptability of the sensory cortex due to exposure and repeated actions. [6-10] However, classical decision-making models do not fully account for perceptual changes at the sensory input level associated with reinforcement learning [2].

Modern neuroeconomic theory identifies several stages in decision-making [2]: forming a representation of the task and context, evaluating the expected value of behavioral options, comparing options, executing the chosen action, and learning from the outcomes to improve future decisions. Most studies focus on the evaluation phase, aiming to understand how organisms from microorganisms to humans optimize their strategies for survival by forecasting the potential benefits and probabilities of different options.

Previous research on neuroplasticity related to decision-making, especially concerning monetary incentives, often employs the anticipated utility/value doctrine, which integrates reward size and probability [11-13]. However, the practical selection process between alternatives is more complex, as shown by Kahneman and Tversky. Despite this, the size of the reward and the probability of obtaining it are crucial in basic experimental scenarios [14].

While biologists and economists have extensively studied the neurobiology of expected value [2, 18], the specific mechanisms of the brain's computation of these parameters remain unclear.

Reinforcement learning theory formalizes flexible decision-making by focusing on prediction error—the difference between expected and actual outcomes. Feedback allows individuals to adjust their predictions and choose options based on

Russian Science Foundation, Grant No. 22-18-00660

past experiences and future expectations. Recent studies also highlight the significance of sensory cortex plasticity, influenced by the reinforcement of stimuli, in affecting subsequent behavior.

Despite numerous studies on sensory system plasticity, particularly the auditory system, many current reinforcement learning models do not consider changes in sensory input. However, evidence shows that sensory neuroplasticity plays a role in decision-making, with sensory cortex adaptations occurring even with unconscious stimulus recognition [8-10, 16-17]. This calls for a decision-making model that incorporates sensory information reorganization due to learning outcomes.

So, integrating sensory plasticity into neuroeconomic models of decision-making can enhance our understanding of how experience and repeated behavior influence sensory cortex reorganization, ultimately affecting decision-making processes. This approach will provide a more comprehensive understanding of the dynamic interplay between sensory input and reinforcement learning in shaping behavior.

Our previous research [18-20] has examined neuroplasticity related to auditory reception in the human brain. Thas study utilized an adapted version of the monetary incentive delay (MID) task [21], where auditory cues replaced visual ones to explore stimulus-reinforcement relationships. Using electroencephalography (EEG), it was observed that reinforcement learning correlates with adaptive changes in the auditory cortex. An auditory adaptation of the MID task revealed increased neural activity and sensory plasticity linked to reinforcement signals [22].

Regression analyses indicated a positive correlation between sensory plasticity and control processes, suggesting that association learning dynamics may predict sensory cortex plasticity. Notably, when the MID task involved monetary loss, an increase in mismatch negativity (MMN) for auditory cues predicting significant loss was observed [20].

Despite evidence of sensory cortical plasticity induced by learning, current models of reinforcement learning and decision-making often assume constant sensory input. Research suggests sensory plasticity can be induced by repeated decision-making, with early neural encoding of expected rewards affecting initial sensory processing.

In the auditory domain, research on reinforcement learning-induced neuroplasticity is limited, with notable exceptions from the HSE University. The MID task, primarily an instrumental learning task, lacks elements of choice and accurate task performance assessment crucial for reward-based learning models.

We developed a new audio version of the lottery, a modified one-armed bandit task. This task required participants to choose between two audio-encoded options with different reward probabilities, allowing for the study of reinforcement learning dynamics.

The relationship between operant learning and neuroplasticity markers (MMN, P3a, and brain oscillatory activity) was investigated. Two sessions of the oddball task, conducted before and after the modified one-armed bandit task, aimed to record electrophysiological correlates of neuroplastic changes due to repeated decision-making.

We created a novel audio version of a lottery-like one-armed bandit task. Participants had to choose between two options, each with a distinct expected outcome, allowing the examination of behavioral patterns in reinforcement learning (RL).

The new task overcomes the limitations of the MID task because it involves active choice on the part of the participant.

Additionally, the task allows for the examination of evoked potential dynamics in mixed feedback scenarios, contributing to understanding cortical feedback processing mechanisms. This can enhance the assessment of learning efficacy and the development of novel methodologies in neuroeconomic studies.

We created a novel audio version of a lottery-like one-armed bandit task. Participants had to choose between two options, each with a distinct expected outcome, allowing the examination of behavioral patterns in reinforcement learning (RL). This adaptation enabled to link learning and neuroplasticity in active choice.

II. Methodology

A. Participants

The study involved a total of 29 participants, comprising 13 males and 14 females. The average age of the participants was 24 years. All participants were right-handed, had normal or corrected vision, and did not have any neuropsychiatric disorders.

B. Equipment

The electroencephalogram (EEG) was recorded using a setup of 32 active electrodes placed according to the 10-20 system. The average value from electrodes 10 and 21, positioned on the ears, served as the referent. The ground electrode (ground) was positioned in place of electrode Fpz. Additionally, an oculogram was recorded using electrode 27 under the right eye and electrode 5 on the outer corner of the left eye. The electrodes were maintained at a resistance level of no more than 10 kOhm.

The recording was conducted using a BrainVision actiCHamp amplifier (Brain Products GmbH) at a sampling frequency of 500 Hz. A 50 Hz notch filter was applied to eliminate electrical device frequency, and the data were further filtered within the 1-40 Hz band. This was followed by a block of the game itself, after which the subjects were once again presented with the oddball task.

C. Stimulus

The stimulus material consisted of a set of ten sounds, which were divided into four blocks. These sounds were created using PRAAT software exclusively for this study. Each block had specific frequencies assigned to the sounds.

The frequencies employed in Block 1 were 272 Hz and 502 Hz, in Block 2 – 381 Hz and 637 Hz, in Block 3 – 325 Hz and 568 Hz, and in Block 4 – 440 Hz and 711 Hz. Additionally, a specific oddball-task paradigm utilised frequencies of 208 Hz and 772 Hz.

Each sound was presented for a duration of 200 ms, with a volume of 70 dB, as measured using a sound intensity meter.

D. Procedure

The game comprised four blocks. The first block comprised an explanation of the task, followed by two blocks of the game, a break, two more blocks of the game, and finally the results were announced. Each block of the game comprised 25 trials. At the commencement of the experiment, each participant was provided with an instruction sheet outlining the procedure. The trials involved the subject pressing either the "left" or "right" buttons on the keyboard to select one of the sound options. Initially, both sounds were played, after which the participant was required to select the desired sound.

It was explained to the participants that their choices would result in a deduction of a specific amount of money from the initial sum.

The remaining amount after each trial was recorded, and at the conclusion of the experiment, 10 randomly selected amounts were presented to the participants as rewards. If the participants did not make a choice within five seconds, they lost all the money they had started the trial with. Additionally, the instructions indicated that the order of the sounds in each trial could change.

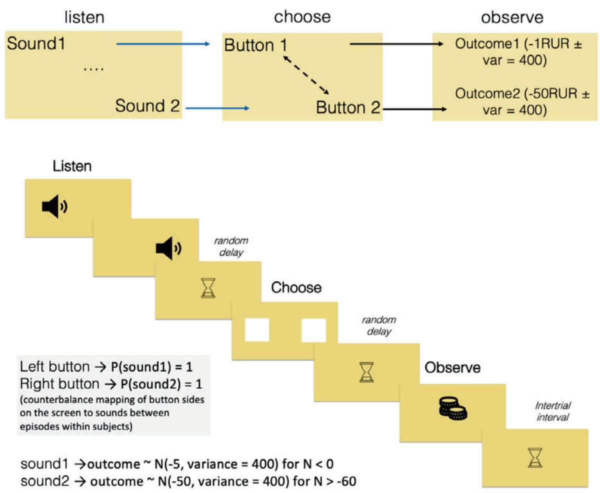

Fig. 1. The layout of the monetary lottery involves the participant hearing two consecutive sounds. Subsequently, they choose between the two options displayed on the screen using the "right" or "left" keys, representing the first or second sound. The screen then displays the amount of money lost associated with the selected sound.

The participant's financial loss was determined by the button selected. If the button indicated a larger amount, the loss was equal to a randomly selected amount from a distribution with a mean value of 5 rubles. Conversely, if the button indicated a smaller amount, the loss was equal to a randomly selected amount from a distribution with a mean value of 50 rubles.

The loss distributions were generated using a Python script, which created arrays of monetary losses within specified ranges for each sound.

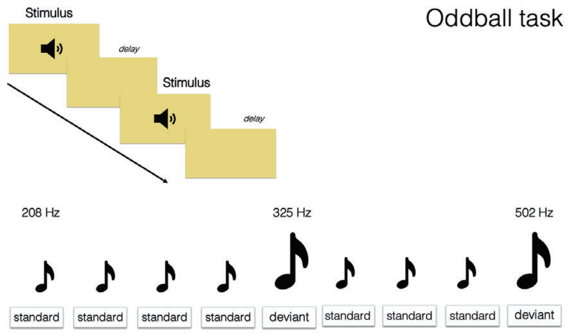

Fig. 2. The scheme of oddball-task.

III. Results

A custom Python script using the MNE library was employed to analyze the acquired EEG data. The data were cleaned of artifacts both manually and algorithmically using the ICA algorithm. The data were then segmented based on the stimuli.

For a detailed analysis, the average magnitude of the components on the Cz electrode was used, as this channel showed the most prominent amplitude expression. The mean amplitude values within the relevant evoked potential window (-200-800 ms) were analyzed.

A repeated measures analysis of variance (ANOVA) was conducted to assess differences in component amplitude. The factors included in the analysis were Stimulus/Control, Large Loss/Small Loss, and Before Game/After Game.

The analyses revealed significant differences in responses to standard and deviant sounds before and after the game. Additionally, significant differences were observed between the responses to sounds encoding large losses and small losses.

TABLE I
RESULTS OF THE ANALYSIS OF VARIANCE. DIFFERENCES IN COMPONENTS OF EVOKED POTENTIALS IN RESPONSE TO STANDARD SOUND AND DEVIANT SOUND BEFORE PERFORMING THE "TWO-HANDED BANDIT" TASK.

	Df	*Sum Sq*	*Mean Sq*	*F-value*	*Pr(>F)*
loss	2	1.78e-10	8.919e-11	31.76	3.1e-14
Residuals	1485	4.17e-09	2.810e-12		

TABLE II
RESULTS OF THE ANALYSIS OF VARIANCE. DIFFERENCES IN COMPONENTS OF EVOKED POTENTIALS IN RESPONSE TO STANDARD SOUND AND DEVIANT SOUND AFTER PERFORMING THE "TWO-HANDED BANDIT" TASK.

	Df	*Sum Sq*	*Mean Sq*	*F-value*	*Pr(>F)*
loss	2	1.050e-10	8.919e-11	5.272e-11	13.13
Residuals	1485	5.962e-09	4.010e-12		

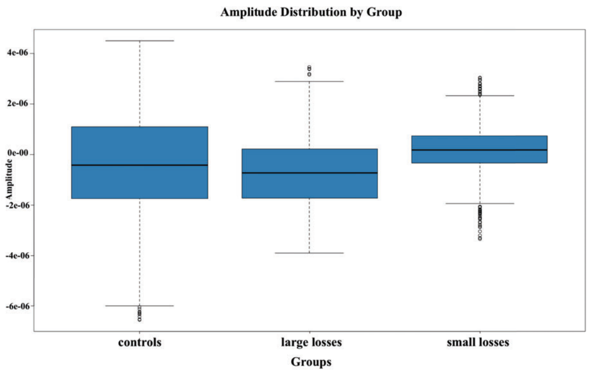

Fig. 3. Differences in the components of evoked potentials in response to standard sound and deviant sound prior to performing the two-armed bandit task in the Cz lead.

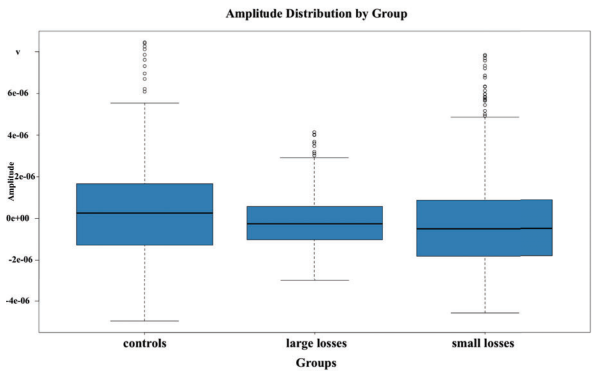

Fig. 4. Differences in components of evoked potentials in response to standard sound and deviant sound after performing the two-armed bandit task in the Cz lead.

TABLE III
RESULTS OF THE ANALYSIS OF VARIANCE. DIFFERENCES IN COMPONENTS OF EVOKED POTENTIALS IN RESPONSE TO STANDARD SOUND AND DEVIANT SOUND FOR CONTROL AUDITORY STIMULI.

	Df	*Sum Sq*	*Mean Sq*	*F-value*	*Pr(>F)*
loss	1	1.890e-10	1.889e-10	38.39	8.47e-10
Residuals	990	4.871e-09	4.920e-12		

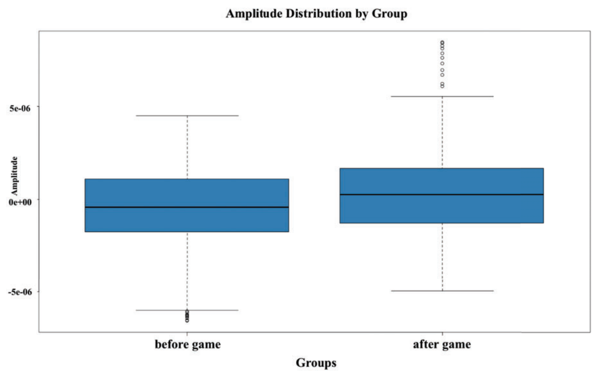

Fig. 5. Differences in components of evoked potentials in response to standard sound and deviant sound for control auditory stimuli in the Cz lead.

TABLE IV
RESULTS OF THE ANALYSIS OF VARIANCE. DIFFERENCES IN COMPONENTS OF EVOKED POTENTIALS IN RESPONSE TO STANDARD SOUND AND DEVIANT SOUND FOR AUDITORY STIMULI ENCODING LARGE LOSSES.

	Df	*Sum Sq*	*Mean Sq*	*F-value*	*Pr(>F)*
loss	1	4.530e-11	4.531e-11	22.27	2.71e-06
Residuals	990	2.015e-09	2.040e-12		

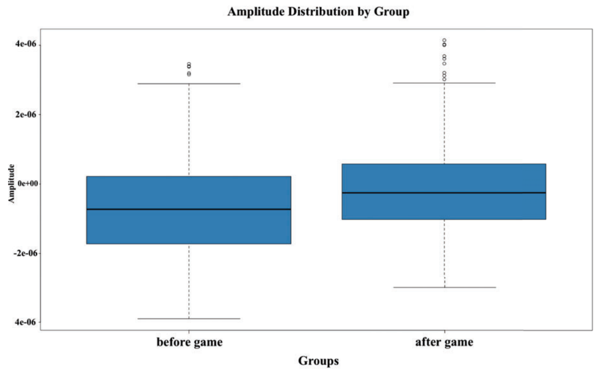

Fig. 6. Differences in components of evoked potentials in response to standard sound and deviant sound for auditory stimuli encoding high loss, in the Cz lead.

TABLE V
RESULTS OF THE ANALYSIS OF VARIANCE. DIFFERENCES IN THE COMPONENTS OF EVOKED POTENTIALS IN RESPONSE TO STANDARD SOUND AND DEVIANT SOUND FOR AUDITORY STIMULI ENCODING LESS LOSS.

	Df	*Sum Sq*	*Mean Sq*	*F-value*	*Pr(>F)*
loss	1	4.700e-11	4.693e-11	14.32	0.000164
Residuals	990	3.246e-09	3.280e-12		

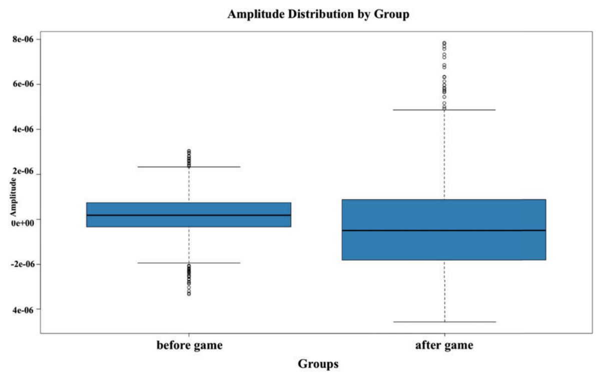

Fig. 7. Differences in the components of evoked potentials in response to standard sound and deviant sound for auditory stimuli encoding less loss, in the Cz lead.

IV. Discussion

The results demonstrated that the changes associated with recognizing sounds encoding larger money losses were more pronounced than those associated with recognizing sounds encoding smaller money losses. In other words, during an economic game with active choice, respondents were trained to better recognise the sounds they were trying to avoid in order to avoid financial losses.

The results of the study show that in situations of active choice and repeated decision-making in the one-handed bandit game, changes in the auditory cortex occur. Furthermore, the magnitude of these changes correlates with the magnitude of the monetary rewards encoded by the sounds.

The results of the study indicated that participants demonstrated an ability to distinguish between sounds encoding monetary losses during an economic game with active choice. The evoked potentials in response to the sound encoding smaller losses were more pronounced than to the sound encoding larger losses. This finding is consistent with the results of previous studies, which demonstrated that participants were better at learning to discriminate between widely varying losses [20]. This is partly explained by Kahneman's theory of reduced loss sensitivity [23].

In order to gain further insight, we intend to analyse the data in greater detail in order to compare the change in evoked potentials with behavioural manifestations of learning, namely reaction speed and correctness of choice.

From a practical standpoint, the dynamic characteristics of learning and their influence on plastic changes can significantly assist in comprehending the learning process of the auditory system. For instance, a certain degree of advancement is essential for the design of tailored learning programmes that facilitate high levels of learner engagement. Research has demonstrated that the rapid acquisition of a new language is accompanied by plastic changes in auditory cortical activity. However, without an understanding of the individual differences in the dynamics of these changes, it is impossible to determine the exact time to begin the process and the duration of exercise required to effectively teach new words or sounds. If the dynamics of reinforcement learning can be used as a predictor of plastic changes, then the proposed model can be used to predict the necessary exercise duration for each individual. Furthermore, the successful outcomes of such a study may encourage the development of neurotechnological approaches to education, which becomes especially crucial in the context of accelerating robotisation and the necessity to change professions, requiring continuous learning for an increasing number of people worldwide.

Acknowledgment

This work was supported by the Russian Science Foundation, Grant No. 22-18-00660 (Neurocomputational mechanism of sensory plasticity).

References

[1] P. W. Glimcher and A. Rustichini, 'Neuroeconomics: The Consilience of Brain and Decision', Science, vol. 306, no. 5695, pp. 447–452, Oct. 2004, doi: 10.1126/science.1102566.

[2] A. Rangel, C. Camerer, and P. R. Montague, 'A framework for studying the neurobiology of value-based decision making', Nat Rev Neurosci, vol. 9, no. 7, pp. 545–556, Jul. 2008, doi: 10.1038/nrn2357.

[3] J. K. Rilling, B. King-Casas, and A. G. Sanfey, 'The neurobiology of social decision-making', Current Opinion in Neurobiology, vol. 18, no. 2, pp. 159–165, Apr. 2008, doi: 10.1016/j.conb.2008.06.003.

[4] Rescorla, R. A., Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H., Black,W. F. Prokasy, (Eds.), Classical conditioning II. New York: Appleton-Century-Crofts, 1972.

[5] C. B. Holroyd and M. G. H. Coles, 'The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity.', Psychological Review, vol. 109, no. 4, pp. 679–709, Oct. 2002, doi: 10.1037/0033-295X.109.4.679.

[6] W. Schultz, 'Behavioral Theories and the Neurophysiology of Reward', Annu. Rev. Psychol., vol. 57, no. 1, pp. 87–115, Jan. 2006, doi: 10.1146/annurev.psych.56.091103.070229.

[7] E. S. Bromberg-Martin, M. Matsumoto, and O. Hikosaka, 'Dopamine in Motivational Control: Rewarding, Aversive, and Alerting', Neuron, vol. 68, no. 5, pp. 815–834, Dec. 2010, doi: 10.1016/j.neuron.2010.11.022.

[8] J. R. Whitlock, A. J. Heynen, M. G. Shuler, and M. F. Bear, 'Learning Induces Long-Term Potentiation in the Hippocampus', Science, vol. 313, no. 5790, pp. 1093–1097, Aug. 2006, doi: 10.1126/science.1128134.

[9] T. J. Vickery, M. M. Chun, and D. Lee, 'Ubiquity and Specificity of Reinforcement Signals throughout the Human Brain', Neuron, vol. 72, no. 1, pp. 166–177, Oct. 2011, doi: 10.1016/j.neuron.2011.08.011.

[10] Y. I. Alexandrov, V. Klucharev, and M. Sams, 'Effect of emotional context in auditory-cortex processing', International Journal of Psychophysiology, vol. 65, no. 3, pp. 261–271, Sep. 2007, doi: 10.1016/j.ijpsycho.2007.05.004.

[11] Von Neumann, J., Morgenstern, O. (2007). Theory of games and economic behavior (60th anniversary ed). Princeton University Press.

[12] J. B. Rotter, J. E. Chance, and E. J. Phares, Applications of a Social Learning Theory of Personality. Holt, Rinehart and Winston, 1972.

[13] A. Bandura, 'Self-efficacy: Toward a unifying theory of behavioral change.', Psychological Review, vol. 84, no. 2, pp. 191–215, 1977, doi: 10.1037/0033-295X.84.2.191.

[14] D. Kahneman and A. Tversky, 'Choices, values, and frames.', American Psychologist, vol. 39, no. 4, pp. 341–350, Apr. 1984, doi: 10.1037/0003-066X.39.4.341.

[15] T. A. Hare, C. F. Camerer, and A. Rangel, 'Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System', Science, vol. 324, no. 5927, pp. 646–648, May 2009, doi: 10.1126/science.1168450.

[16] O. E. Güler and K. M. Thomas, 'Developmental differences in the neural correlates of relational encoding and recall in children: An event-related fMRI study', Developmental Cognitive Neuroscience, vol. 3, pp. 106–116, Jan. 2013, doi: 10.1016/j.dcn.2012.07.001.

[17] A. S. Persichetti, G. K. Aguirre, and S. L. Thompson-Schill, 'Value Is in the Eye of the Beholder: Early Visual Cortex Codes Monetary Value of Objects during a Diverted Attention Task', Journal of Cognitive Neuroscience, vol. 27, no. 5, pp. 893–901, May 2015

[18] E. Krugliakova, V. Klucharev, T. Fedele, A. Gorin, A. Kuznetsova, and A. Shestakova, 'Correlation of cue-locked FRN and feedback-locked FRN in the auditory monetary incentive delay task', Exp Brain Res, vol. 236, no. 1, pp. 141–151, Jan. 2018, doi: 10.1007/s00221-017-5113-2.

[19] E. Krugliakova et al., 'The Monetary Incentive Delay (MID) Task Induces Changes in Sensory Processing: ERP Evidence', Front. Hum. Neurosci., vol. 13, p. 382, Nov. 2019, doi: 10.3389/fnhum.2019.00382.

[20] A. Gorin et al., 'Cortical plasticity elicited by acoustically cued monetary losses: an ERP study', Sci Rep, vol. 10, no. 1, p. 21161, Dec. 2020, doi: 10.1038/s41598-020-78211-7.

[21] B. Knutson, A. Westdorp, E. Kaiser, and D. Hommer, 'FMRI Visualization of Brain Activity during a Monetary Incentive Delay Task', NeuroImage, vol. 12, no. 1, pp. 20–27, Jul. 2000, doi: 10.1006/nimg.2000.0593.

[22] D. B. Polley, E. E. Steinberg, and M. M. Merzenich, 'Perceptual Learning Directs Auditory Cortical Map Reorganization through Top-Down Influences', J. Neurosci., vol. 26, no. 18, pp. 4970–4982, May 2006, doi: 10.1523/JNEUROSCI.3771-05.2006.
[23] D. Kahneman and A. Tversky, 'Prospect Theory: An Analysis of Decision under Risk', Econometrica, vol. 47, no. 2, p. 263, Mar. 1979, doi: 10.2307/1914185.