

An overview of reward prediction error and its links with dopamine

Tianxiang Yao

College of Life Sciences, Sichuan University, Chengdu, 610000, China

tianxiang_yao@outlook.com

Abstract. Reward prediction error (RPE) refers to the discrepancy between expected and actual rewards received during an event, signaling the difference between what was predicted and what actually happened. Dopaminergic neuron encodes RPE signals in the brain, and is responsible for updating reward expectations and influencing decision-making processes. The relationship between RPE and dopamine has led to research in understanding reward-driven learning and its implications on cognition and behavior. In this review, I will provide an overview of the principles, and task models used to quantify RPE. I will also discuss about the neural mechanisms underlying RPE generation, with a particular focus on the role of dopamine in encoding and transmitting RPE signals and enhance learning. Overall, this review highlights the importance of RPE in reward processing, decision-making, memory and so on, emphasizing its connection to dopamine.

Keywords: reward prediction, dopamine, circuit, learning

1. Introduction

Reward processing and decision-making are essential aspects of human behavior, playing a vital role in our daily lives. Our brain makes predictions of rewards based on experience, and constantly evaluates the outcomes of our actions and adjusts our behavior based on the actual rewards we receive. Understanding how our brains process and respond to rewards is crucial for understanding how we learn and make decisions.

One concept that researchers have focused on in the study of reward processing is the reward prediction error (RPE), referring to the difference between the expected reward and the actual reward received [1,2]. RPE represents a signal in the brain that tells us when something unexpected happens, either better or worse than what we predicted [3].

Understanding RPE is important because it helps us understand how our brains learn from rewards and make decisions based on experiences. By studying RPE, researchers can gain insights into the computational and neural processes that underlie our ability to learn and adapt in response to rewards [4,5].

Studying RPE also has implications for understanding cognitive processes and neurological diseases. It can help us understand why psychiatric disorders (such as addiction, depression, and etc.) affect our ability to process rewards [6]. By studying RPE, researchers may be able to develop new treatments for these disorders.

Dopamine, a neurotransmitter concentrated in the midbrain, is closely linked to the brain's reward system. It plays a vital role in signaling rewards, reinforcing behaviors, and influencing decision-making [7]. Understanding the connection between dopamine and the brain's reward system is crucial for comprehending reward prediction error and its impact on learning and decision-making processes. This review aims to provide an overview of RPE and its relationship with dopamine, including the principles of RPE, task models used to measure RPE, the neural mechanisms generating RPE and how it affects decision-making behavior, and with future explorations and potential applications of RPE research. By studying RPE and its connection to reward processing, this review draws a picture of how the brain learns and makes decisions.

2. Principles of RPE

RPE derives from the comparison between the predicted rewards based on expectations and the actual rewards observed. It serves as a key mechanism for learning from rewards and adjusting behavior accordingly.

2.1. Basic principles in the RPE reaction process

- a.* Expectation formation: The brain constantly makes predictions about the rewards it expects based on prior knowledge, learned associations, and contextual cues.
- b.* Reward outcome evaluation: When the actual reward outcome is obtained, the brain compares it with the expected reward, and the difference between is computed as RPE.
- c.* Prediction error signal: RPE signals indicate the extent to which the actual reward differentiates from the expected reward. Positive RPEs occur when the reward is better than expected, while negative RPEs occur when the reward is worse than expected.
- d.* Neural representation: RPE signals are encoded by dopaminergic neurons in the brain. These neurons detect and transmit RPE information by producing dopamine, and update reward expectations.
- e.* Learning and adaptation: RPE signals affect learning and adaptation processes by influencing the updating of reward expectations and influence future decision-making.

2.2. Evidence from neuroscience

- a.* Dopaminergic neurons: Studies in animals and humans have shown that dopamine neurons play a crucial role in encoding and transmitting RPE signals [8]. These neurons show increased activity in response to unexpected rewards, encoding positive RPEs, and decreased activity when expected rewards are omitted, encoding negative RPEs.
- b.* Brain regions: RPE-related activity has been observed in various brain regions, including the striatum, prefrontal cortex, and anterior cingulate cortex. The striatum evaluates rewards and initiates behavior. The prefrontal cortex integrates reward information and guides decision-making. The anterior cingulate cortex monitors conflicts and adjusts behavior. Together, they are involved in reward processing, reinforcement learning, and decision-making [6].
- c.* Neuroimaging Studies: Functional magnetic resonance imaging (fMRI) studies, a non-invasive brain imaging technique that measures changes in blood flow to detect and map brain activity, have demonstrated a correlation between RPE signals, dopamine release, and brain activity in regions associated with reward processing [9]. These findings provide further support for the role of dopamine and RPE in shaping neural responses to rewards.
- d.* Clinical Implications: Dysfunctions in RPE processing and dopamine signaling have been implicated in several psychiatric disorders, such as addiction and depression [10]. Understanding RPE-related mechanisms may contribute to the development of therapeutic interventions targeting reward processing deficits.

3. Task models and methods used to measure RPE

3.1. Reinforcement learning paradigm

The reinforcement learning paradigm is widely used to study RPE. For example, a study by Liebenow et al. in 2022 employed this paradigm to investigate the role of RPE in the encoding of value-based decisions in the prefrontal cortex. In this task, participants or animals are presented with a series of choices or actions and receive feedback in the form of rewards or punishments. By comparing the expected reward based on previous experience with the actual outcome, researchers can quantify the RPE associated with each choice [11]. Tasks like the multi-armed bandit task is an example of reinforcement learning paradigms used to study RPE. The choice of the reinforcement learning paradigm is justified due to its ability to capture the dynamic nature of reward processing and decision-making.

3.2. Prediction error manipulation

In this task, researchers manipulate the expected outcome of an action to create prediction errors. For example, they may unexpectedly increase or decrease the probability of receiving a reward for a particular action. By comparing participants' expectations with the actual outcomes, researchers can measure the RPE associated with the manipulated prediction errors [12].

3.3. Pavlovian conditioning

Pavlovian conditioning tasks involve pairing neutral stimuli (such as sounds or images) with rewarding or aversive outcomes. For example, participants could be presented with a series of visual stimuli, each followed by a rewarding outcome (e.g., money) or no outcome. Over time, participants learn to associate certain visual stimuli with rewards and develop expectations about the forthcoming rewards. To examine RPE, the researchers will manipulate the association between stimuli and outcomes by occasionally presenting a visual stimulus that was previously associated with a reward but withholding the reward on those trials. This violation of the expected reward could lead to a negative prediction error, indicating a deviation from participants' expectations. And at the same time the participants' neural responses were recorded [13].

3.4. Brain imaging studies

Functional magnetic resonance imaging (fMRI) and electrophysiological techniques can be employed to investigate RPE-related brain activity. During these studies, when participants perform tasks that elicit RPE (such as decision-making tasks), brain regions, such as the prefrontal cortex, are examined for fMRI signals. These signals have been found to correlate with RPE in brain regions involved in reward processing. The blood oxygen level-dependent (BOLD) response, reflecting changes in blood flow and oxygenation, can vary based on RPE. Increased activity could be observed when outcomes go away from expectations (positive RPE), while decreased activity refers to outcomes better than expected (negative RPE). Model-based fMRI analyses further highlight the relationship between RPE signals and neural responses [14]. However, fMRI is an indirect measure of neuronal activity and requires integration with other techniques for a comprehensive understanding of RPE's neural mechanisms.

These tasks and methods provide researchers with various means to test RPE in both human and animal subjects. By employing them, scientists can gain insights into the neural mechanisms underlying RPE and its role in learning, decision-making, and reward-related behaviors.

4. Linkage between dopamine and RPE and its influence on learning

RPE, or reward prediction error, can be classified into three types: positive error, zero error, and negative error [15]. Positive error occurs when an unexpected reward is received without any prediction. Dopaminergic neurons exhibit a positive signal by increasing their firing rate, indicating the presence of an unexpected reward. Negative error happens when a predicted reward is not received. Dopaminergic neurons generate a negative signal by decreasing their firing rate. In such cases, the activity of dopaminergic neurons is suppressed. When both the expectation and reward match the prediction, there

is no prediction error. Dopaminergic neurons do not exhibit any specific changes in firing rate. The reward is anticipated and matches the prediction, resulting in no prediction error.

Dopamine neurons release dopamine to influence the RPE circuit [16,17,18]. However, the exact role of dopamine neurons in this circuit remains unclear. Previous research has uncovered several fragments of information regarding this role.

Schultz et al. proposed the “reward prediction error hypothesis” to explain the function of dopamine neurons in the reward circuit [19]. This theory was based on extracellular recordings of single dopamine neurons in monkeys. Initially, the dopamine neurons showed no response to the conditioned stimuli (CS) but exhibited a rapid discharge to the unconditioned stimuli (US), such as juice. However, in the later stages of training, the dopamine neurons responded only to the CS and not the US. Moreover, if the expected reward (US) was omitted after the training, the dopamine neurons exhibited inhibitory responses. By studying the dopamine neuron response to the US, it became clear that dopamine discharge was determined by the “actual reward received minus the expected reward.” In detail, during the initial stage of training, when the expected reward was 0 and the actual reward was 1 (resulting in a prediction error of 1), the dopamine neurons were excited. During the later stage of training, when both the expected and actual rewards equaled 1 (resulting in a prediction error of 0), the dopamine neurons did not respond. If the reward was unexpectedly omitted (expected reward = 1, actual reward = 0) (resulting in a prediction error of -1), the dopamine neurons showed inhibition. This simple description explains how dopamine neurons encode reward prediction error [20].

In the reward system, the neurons that release dopamine are located in the ventral tegmental area (VTA) and substantia nigra (SN) of the midbrain [21]. Early researchers conducted experiments known as intracranial self-stimulation by implanting stimulating electrodes in the VTA. They observed that whenever a subject pressed a lever, electrical stimulation of this area led to excessive lever pressing. It was also discovered that dopamine release increased when animals received rewarding stimuli such as food, water, or sexual stimulation. After the administration of addictive drugs, dopamine release was further heightened. Consequently, the early theory suggested that dopamine release was responsible for generating pleasure in animals [22].

There is substantial supporting evidence for the relationship between dopamine and learning. Recently, two laboratories used optogenetics to demonstrate the causal relationship between dopamine neurons and associative learning [23,24]. The basic principle involves optogenetic stimulation or inhibition to simulate the reward prediction error signal [10]. But above all, the underlying mechanisms are still full of complexity in the real-world incentives [17].

5. Further explorations on RPE and dopamine

5.1. Neural mechanisms of RPE

While significant progress has been made in understanding the role of dopamine and RPE in reward processing, there is still much to learn about the neural mechanisms underlying them, such as the feedback loops and potential modulatory factors in the RPE circuit [25]. Future research could investigate the specific neural circuits involved in RPE generation, including interactions between the prefrontal cortex, nucleus accumbens, striatum, and midbrain dopamine neurons [26]. By getting to know more pieces of these mechanisms, we can gain a more comprehensive understanding of how RPE is encoded and processed in the brain.

5.2. Individual differences in RPE

Exploring individual differences in RPE signals and their relationship to personality traits, genetic variations, and environmental factors could provide valuable insights. Future research could be investigating why some individuals show higher or lower RPE responses compared to others, and this may help explain differences in reward sensitivity, neural plasticity, decision-making strategies, and so on.

5.3. Developmental perspectives

Studying RPE and dopamine-related processes across different stages of development is also an important area. Research could focus on how RPE mechanisms grow and change during childhood, adolescence, adulthood or senescence. Apart from this, we could also investigate how early-life experiences impact RPE functioning in the long-term consequences. Understanding RPE during different developmental phases may help with our understanding of neurological disorders with an onset during specific life stages.

5.4. Neuroimaging techniques

The advancement of neuroimaging techniques, such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET), provides opportunities to investigate RPE-related brain activity in humans with higher resolution. Yet, indirect measures, limited spatial resolution are still limitations of these techniques. Combining these techniques with computational models and behavioral measures can help deepen research of the neural mechanisms of RPE. Additionally, exploring other neurotransmitter systems different from dopamine, may provide a more comprehensive picture of reward processing. Optogenetics can also offer powerful tools for exploring neural circuits and manipulating specific amounts of neurons in vivo. Applying these techniques to the investigation can inspire more ideas and methods, thereby enhancing our chance of understanding the relationship between RPE signal and dopamine.

In conclusion, further explorations of reward RPE and its association with dopamine are crucial for advancing our understanding of reward processing and decision-making.

6. Potential directions

6.1. Addiction treatment

The understanding of RPE and its connection to dopamine has implications for addiction treatment. By uncovering the neural mechanisms underlying RPE, we may develop interventions to modulate the reward system and reduce addictive behaviors. This may involve higher-class techniques or pharmacological interventions that specifically target dopamine-related processes [27].

6.2. Depression and anxiety interventions

RPE and dopamine dysregulation have been implicated in various mental health conditions such as depression, anxiety disorders, and schizophrenia [28]. Potential applications include developing interventions that directly address aberrant RPE signals and dopamine functioning to alleviate symptoms and improve overall mental well-being. Also, RPE-based computational models may assist in predicting treatment response and optimizing therapeutic strategies for mental health disorders.

6.3. Cognitive Enhancement

Since RPE plays a role in learning and decision-making processes, it is possible to manipulate RPE signals, to enhance motivation, concentration, and learning outcomes [29]. This could have implications for educational purposes, rehabilitation, and etc.

6.4. Entertainment and athletic sports:

RPE-based algorithms can be applied in the gaming and entertainment industry to improve user experiences. By dynamically adjusting game mechanics and rewards based on individual RPE responses, games can be designed to provide optimal satisfaction for users, increasing engagement and enjoyment.

As for sports performance, by integrating RPE measures with training protocols, coaches and athletes can manipulate interventions to improve their performance [30].

6.5. Artificial Intelligence and Robotics

RPE-based computational models can improve reinforcement learning algorithms in artificial intelligence (AI) systems and robotics. By incorporating RPE signals, AI agents can make more accurate predictions and optimize their behavior in dynamic environments, enabling applications in auto-vehicles, robot assistants, and smart home systems [31].

The understanding of reward prediction error (RPE) and its association with dopamine opens up numerous potential applications across various fields. From addiction treatment and mental health interventions to cognitive enhancement and gaming, RPE-related principles can be utilized to optimize interventions, enhance performance, and improve overall well-being. As these applications develop, it is essential to value the ethical considerations and ensure responsible use for the benefit of individuals and society.

7. Conclusion

In conclusion, this review article provides a comprehensive analysis of current understanding of RPE and its association with dopamine. The article highlights the principles of RPE, discusses the neural mechanisms generating RPE and decision-making behavior, and highlights future explorations and potential applications of RPE research. However, this review predominantly focuses on the role of dopamine, leaving limited space to for other neurotransmitters that may also influence RPE processing. A more comprehensive analysis of the entire neural network involved in RPE may be needed to provide a broader perspective. But overall, this review article provides an overview of RPE and its links with dopamine. By studying RPE and its connection to reward processing, we can gain a better understanding of how our brains learn and make decisions.

References

- [1] Glimcher PW. 2011. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* 108 Suppl 3: 15647-54
- [2] Lerner TN, Holloway AL, Seiler JL. 2021. Dopamine, Updated: Reward Prediction Error and Beyond. *Current opinion in neurobiology* 67: 123-30
- [3] Montague PR, Dayan P, Sejnowski TJ. 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 16: 1936-47
- [4] Schultz W. 2016a. Dopamine reward prediction-error signalling: a two-component response. *Nature reviews. Neuroscience* 17: 183-95
- [5] Schultz W. 2016b. Dopamine reward prediction error coding. *Dialogues in clinical neuroscience* 18: 23-32
- [6] Geugies H, Groenewold NA, Meurs M, Doornbos B, de Klerk-Sluis JM, et al. 2022. Decreased reward circuit connectivity during reward anticipation in major depression. *NeuroImage. Clinical* 36: 103226
- [7] Lammel S, Lim BK, Malenka RC. 2014. Reward and aversion in a heterogeneous midbrain dopamine system. *Neuropharmacology* 76 Pt B: 351-9
- [8] Berridge KC. 2007. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology* 191: 391-431
- [9] Fouragnan E, Retzler C, Philiastides MG. 2018. Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis. *Human brain mapping* 39: 2887-906
- [10] Keiflin R, Janak PH. 2015. Dopamine Prediction Errors in Reward Learning and Addiction: From Theory to Neural Circuitry. *Neuron* 88: 247-63
- [11] Liebenow B, Jones R, DiMarco E, Trattner JD, Humphries J, et al. 2022. Computational reinforcement learning, reward (and punishment), and dopamine in psychiatric disorders. *Frontiers in psychiatry* 13: 886297

- [12] Pfabigan DM, Alexopoulos J, Bauer H, Sailer U. 2011. Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology* 48: 656-64
- [13] Noritake A, Nakamura K. 2019. Encoding prediction signals during appetitive and aversive Pavlovian conditioning in the primate lateral hypothalamus. *Journal of neurophysiology* 121: 396-417
- [14] O'Callaghan G, Stringaris A. 2019. Reward Processing in Adolescent Depression Across Neuroimaging Modalities. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie* 47: 535-41
- [15] Matsumoto M, Hikosaka O. 2009. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459: 837-41
- [16] Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N. 2015. Arithmetic and local circuitry underlying dopamine prediction errors. *Nature* 525: 243-6
- [17] Rothenhoefer KM, Hong T, Alikaya A, Stauffer WR. 2021. Rare rewards amplify dopamine responses. *Nature neuroscience* 24: 465-69
- [18] Watabe-Uchida M, Eshel N, Uchida N. 2017. Neural Circuitry of Reward Prediction Error. *Annual review of neuroscience* 40: 373-94
- [19] Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. *Science (New York, N.Y.)* 275: 1593-9
- [20] Schultz W. 1998. Predictive reward signal of dopamine neurons. *Journal of neurophysiology* 80: 1-27
- [21] Solié C, Girard B, Righetti B, Tapparel M, Bellone C. 2022. VTA dopamine neuron activity encodes social interaction and promotes reinforcement learning through social prediction error. *Nature neuroscience* 25: 86-97
- [22] Zaghoul KA, Blanco JA, Weidemann CT, McGill K, Jaggi JL, et al. 2009. Human substantia nigra neurons encode unexpected financial rewards. *Science (New York, N.Y.)* 323: 1496-9
- [23] Chang CY, Esber GR, Marrero-Garcia Y, Yau HJ, Bonci A, Schoenbaum G. 2016. Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. *Nature neuroscience* 19: 111-6
- [24] Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH. 2013. A causal link between prediction errors, dopamine neurons and learning. *Nature neuroscience* 16: 966-73
- [25] Gerfen CR, Surmeier DJ. 2011. Modulation of striatal projection systems by dopamine. *Annual review of neuroscience* 34: 441-66
- [26] Pan WX, Coddington LT, Dudman JT. 2021. Dissociable contributions of phasic dopamine activity to reward and prediction. *Cell reports* 36: 109684
- [27] García-García I, Zeighami Y, Dagher A. 2017. Reward Prediction Errors in Drug Addiction and Parkinson's Disease: from Neurophysiology to Neuroimaging. *Current neurology and neuroscience reports* 17: 46
- [28] Papalini S, Beckers T, Vervliet B. 2020. Dopamine: from prediction error to psychotherapy. *Translational Psychiatry* 10: 164
- [29] Jang AI, Nassar MR, Dillon DG, Frank MJ. 2019. Positive reward prediction errors during decision-making strengthen memory encoding. *Nature Human Behaviour* 3: 719-32
- [30] Mohebi A, Pettibone JR, Hamid AA, Wong JT, Vinson LT, et al. 2019. Dissociable dopamine dynamics for learning and motivation. *Nature* 570: 65-70
- [31] Kawato M, Cortese A. 2021. From internal models toward metacognitive AI. *Biological cybernetics* 115: 415-30