OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

Dopamine and Reward Prediction Error Generalizability

Aurea Maria Marchini Charloux
auramarchini@gmail.com

ABSTRACT

It has been established that learning relies primarily on the brain's ability to compare expectations with outcomes. Depending on the results, the brain adjusts future behavior based on what it considers worth repeating. Dopamine has previously been associated with reward and pleasure. Newer research established, however, that its primary function was to encode reward prediction error (RPE), the difference between expected and actual outcomes. This paper presents a literature review addressing the question: How do dopamine neurons represent learning signals for diverse behaviors, and to what extent is reward prediction error maintained and adapted across different contexts? Firstly, the paper analyzes foundational experimental papers that have enabled the scientific understanding of RPE as a computational principle and of dopamine signaling's role in reward-based learning. Secondly, the paper presents evidence of RPE's generalizability beyond external reward by reporting an experiment on songbirds. Thirdly, the paper reviews studies on courtship and maternal behavior, thereby showing that the dopaminergic system is adaptable, prioritizing social outcomes while maintaining the prediction error framework. Fourthly, an analysis of synaptic plasticity - the brain's ability to strengthen or weaken connections between neurons - is conducted. This explains the drive for learning and behavioral change: dopamine-dependent prediction-error evaluation. Finally, an analysis of the pathological consequences of deficiency in Parkinson's and dopamine excess in addiction is conducted, respectively. As a whole, the analysis and synthesis of the experiment's findings strongly suggest a universal mechanism: reward prediction error. Furthermore, the experiments demonstrate RPE's flexibility in response to the agent's current goals. Additionally, RPE extends beyond reward acquisition to encompass performance evaluation, social behavior, parenting, and neurological conditions across species.

INTRODUCTION

Dopamine is often described as a simple "pleasure neurotransmitter"; however, modern research suggests it is instead the foundation of incentive salience, learning, and behavioral adaptation [1]. Dopamine enables our motivation, pleasure, memory, focus, and movement to be actively regulated by encoding discrepancies between expected and real outcomes. [2] Dopamine neurons allow our brains to learn from the outcomes of actions and decide whether to repeat them [3]. They are also involved in human

conditions that affect many people around us [4]. On the other hand, dopamine plays a crucial role in both addiction and Parkinson's. Addiction led to an estimated 92,000 overdose deaths in America (aged 15-64) between 2020 and 2021, an increase of 20% over the past decade.[5, 6]. Meanwhile, Parkinson's is expected to increase from affecting 6.9 million individuals in 2015 to 14.2 million by 2040 [7]. Over the past decade, research has advanced significantly, enabling scientists to discover a mechanism known as "reward prediction error" [8]. This mechanism relates to dopamine neurons in the context of reinforcement learning, where agents learn to make decisions through trial and error. Together, these ideas show that dopamine provides signals that enable the brain to update future behavior. While newer papers suggest that it may be associated with social interactions and parenting behavior [9, 10, 11], the generalizability and flexibility of reward Prediction Error (RPE) across different contexts have yet to be observed. The purpose of my paper is to address the fundamental question "How do dopamine neurons represent learning signals for diverse behaviors, and to what extent is reward prediction error maintained and adapted to different contexts?"

METHODS

I. Search strategy

This paper employs a structured search strategy alongside a narrative literature review approach. This paper aims to synthesize the fundamental experimental evidence relevant to the research question, rather than to list all published studies on dopamine. The most prominent data: Olds and Milner (1954), Shultz et al. (1997), Gadgakar et al. (2016), Shea et al. (2023), and Gadgakar et al. (2023) were collected from the National Center of Biotechnology, PubMed, Google Scholar, and ScienceDirect. The terms used to search for relevant experimental papers included "dopamine," "reward prediction error," "reinforcement learning," "dopaminergic signaling," "prediction error," "basal ganglia," "songbird learning," "social reward," "maternal behavior," "courtship behavior," "synaptic plasticity," "addiction," and "Parkinson's disease." In addition to the main papers used, a combination of additional database research was reviewed and continuously cited to support the development of the reward prediction error (RPE) theory.

II. Inclusion and Exclusion Criteria

The research process required a selective process to determine which studies to include or exclude. To be included in the research, specific criteria had to be met. These include: (1) The studies had to include a measurement of either dopamine activity, demonstrate learning progress, or represent reinforcement processes. (2) Original methodologies had to be used to collect data. Examples of acceptable methods for data collection included electrophysiology, fiber photometry, calcium imaging, and optogenetics. (3) Experts in the field scrutinized research, methods, and results before applying for publication. (4) The study brought significant input regarding RPE. (5) Studies were included regardless of publication date if they were considered foundational to the concepts of RPE.

If research met the following criteria, it would be excluded from the paper's sources. (1) A primary focus of dopamine as a pleasure neurotransmitter without encompassing its role as a learning mechanism. (2) lacked original experimental evidence. (3) did not include any new concepts or data that can be directly

associated with the research question. Based on these criteria, of the 40 initial sources, only 22 were used to develop the paper.

Among the selected papers, the studies that made the strongest contributions to the development of RPE theory (Olds & Milner, 1954; Schultz et al., 1997) were highlighted and analyzed in greater detail. Meanwhile, the other studies were used as support for the concepts observed.

The selection of more modern papers was determined by how they extended the concepts of RPE to new domains and contexts (song learning, socialization, parenting, and neurological pathologies)

III. Synthesize Method

Rather than analyzing the studies independently, a comparative analysis was used to synthesize the data and identify recurring principles. Studies were not organized randomly. Rather, they were thematically categorized into the main subsection of the paper: I. Foundation of reward prediction. II. Reward prediction error based on internal benchmarks. III. Dopamine in social and parenting. IV. Dopamine signaling connection to learning & plasticity. V. Addiction and its abnormal dopamine activity. VI. Parkinson's and its suppression of dopamine signaling.

The synthesis was centered around figure analysis. This brings originality to the paper, alongside a structured synthesis of complex neuroscience studies. The use of various figures invites a deeper understanding of the concepts discussed. A clearer comparison between the experimental figures and the author's interpretation emerges. It remains its focus on the computational mechanisms.

In terms of this paper's unfolding, it fits into a conceptual progression rather than a chronological one. Evidence first invited an understanding of the core RPE concepts, later delved into its generalizability across contexts, and finally discussed the neurological illnesses associated with RPE.

IV. Quality Assessment

A few factors determined the quality of the studies. (1) coherence between data evidence and the author's interpretation. (2) In addition to the content, attention was paid to the clarity and structure of the papers. (3) Preference was given to papers published in renowned journals, since this would ensure that extensive scrutiny was applied.

DISCUSSION

I. Foundation of Reward Prediction

Before the discovery of dopamine and its role as a key factor in the learning system, scientists wondered whether there might be a circuit in the animal brain whose ultimate role was to process reward. An initial attempt to reshape neuroscience was made in 1954 by James Olds and Peter Milner [12]. The experiment implanted electrodes in rats' brains and placed them in a Skinner box where pressing a lever would

electrically stimulate the implanted area. The experiment was repeated with various rats, using electrodes in distinct brain regions. The results eventually reshaped how scientists perceive motivation and reinforcement. Rats whose electrodes had been implanted in the septal area of the forebrain uninterruptedly pressed the lever. There was no additional reward other than the electrical stimulus; no water, nor food was given. A rat whose electrode had been placed in the septal area subsequently pressed the lever approximately 742 times per hour [12]. Table 1 of the Olds and Milner 1954 paper [12] compares the percentage of time spent responding during acquisition and extinction in rats with electrodes placed in either the septal or medial geniculate (m.g.) area.

As shown in the table, the response percentage decreased significantly from above 75% during acquisition to 6%-21% when the current was turned off during extinction periods [12]. As soon as the electrical stimulus was restarted, the response immediately resumed. The curve produced was strongly correlated with that of standard external rewarding.
Although dopamine was not directly measured, the experiment enabled scientists to identify a specific brain system that produces a rewarding effect during certain behaviors. The Olds and Milner [12] experiment did not explain how the brain actually learned from reward feedback. However, it laid the groundwork for many future studies aimed at understanding whether a single system could account for learning across all contexts. Wolfram Schultz, Peter Dayan, and P. Read Montague in 1997 [13], suggested that dopamine not only signals reward but also compares expected with actual reward. This is now known as ‘reward prediction error’. The experiment eventually provided biological evidence supporting reinforcement learning theory: behavior is refined by reward prediction error. If the prediction error is “better than expected”, the agent will reinforce the behavior and encourage it in the future. In contrast, if the prediction error is “worse than expected”, the brain will discourage the repetition of these actions.

The Schultz [13] experiment establishes how dopamine neurons encode a prediction-error mechanism for external rewards such as juice. The Shultz [13] experiment involved implanting electrodes for electrophysiological recording, with the core concept being to record the action potential produced whenever a neuron communicated. The implant was placed in two specific brain regions where dopamine neurons are most commonly found: the substantia nigra and the Ventral Tegmental Area (VTA). The scientist recorded dopaminergic neural activity as a primate received an unexpected positive reward (juice) versus a worse-than-expected surprise reward (air puffs or drops of saline). The results showed that between 55% and 80% of the recorded neurons exhibited the same behavior: a spike in electrical activity when the primate received a better-than-expected result, versus a dip in dopamine activity [13]. This demonstrates that dopamine neurons respond to positive rewards with phasic activation and to aversive ones with suppression.

The paper by Schultz, Dayan, and Montague [13] argued that dopamine neurons did not keep the prediction-error signals to themselves. Rather, they broadcast it throughout the brain to areas involved in motivation and decision-making. These areas include the striatum, nucleus accumbens, and the frontal cortex, which help the brain decide whether a behavior is worth repeating. This discovery led to the understanding that dopamine neurons are biological detectors of prediction error, constantly updating the

brain about whether an outcome is better than, worse than, or equivalent to the expectation.

Dopamine neurons, however, are not solely linked to the precise moment during which a reward is delivered. Based on electrophysiological recordings, the Schultz paper found that as an agent learns to associate a particular sensory cue with the arrival of the rewarding event, dopamine neurons activate when the sensory cue is presented. This argues that, contrary to previous thought, activity in dopamine neurons is not linked to pleasant stimuli.

Rather, dopamine is directly linked to prediction. This can be clearly demonstrated in Figure 1 of the Shultz 1997 paper [13], where the x-axis represents time (in seconds) relative to 0, the exact moment when the prediction cue occurs (labeled CS - conditioned stimulus-). Point "R" refers to the exact moment at which the reward takes place. The y-axis of the top graphs (histograms) demonstrates the firing rate of dopamine neurons. What should be noticed here is that in the first graph, where the animal does not expect to get any reward, there is a normal dopaminergic neural activity until after point "R" where the reward was given. At that point, we can see the bars get taller, indicating a significant increase in dopamine activity because the animal is receiving "better than expected" feedback.

In the second graph (during which the animal has already associated the cue with the arrival of the reward), we can see significant firing at point CS, followed by stable dopamine neural activity even when the actual reward is given. This confirms the suggestion that dopamine neurons do not encode the actual reward but rather predict the reward. Transitioning to the third graph, in which the cue is given but no reward follows, we can see once again a strong firing at the conditioned stimuli (cue), followed by a stable dopaminergic level and later a dip at point "R" (when no reward is given).

By examining the bottom section of the graphs (raster plots), where each row represents a different trial, and each dot represents an action potential ("spike"), we can see a similar pattern. Initially, the dots correlate with the reward distribution. As the monkey learns to associate the sensory cue with the arrival of a reward, a correlation between the dots and the conditioned stimuli develops. Followed by no particularly noticeable neural activity when the reward is delivered. In the last graph, the dots once again correlate with the conditioned stimuli and later completely disappear at point R when no reward is given. This data clearly supports the Shultz, Dayan, and Montague [13] statement that dopamine neurons do not fire as a result of a pleasant reward but rather encode the prediction of a "better than expected", "equivalent to expectation," or "worse than expected" outcome. [13] Thus, showing that the brain is actively updating an internal model of when rewards appear. The Shultz [13] paper demonstrates that Reward Prediction Error (RPE) serves as a learning signal for at least one category of reward: external reward. However, this statement alone is insufficient to determine whether RPE can be applied in all contexts. For instance, do contexts in the absence of external rewards apply RPE as a learning mechanism? This limitation is directly addressed in the upcoming section.

II. Reward Prediction Error Based on Internal Benchmarks

Various behaviors, such as learning to speak in humans or singing in songbirds, are not based on any

external reward. Instead, these complex processes involve the agent comparing its behavior to a set of internal models of how it should be. If reward prediction error is in fact generalizable - as the research question asks - then it must be applied to contexts in which there is no external reward but rather an internal benchmark. [9]

Songbirds such as zebra finches consistently adjust their pronunciation of syllables by comparing the sound they make with that of an adult bird. The Gadgakar (2016) [9] paper used zebra finches and implanted small electrodes in the ventral tegmental area (VTA) - a region where dopamine neurons are located [9]. The activity of individual neurons was recorded with these electrodes (single-unit recordings) while the bird was singing voluntarily and freely. The electrode placement is shown in Figure 1B of Gadgakar et Al, 2016 [9] (labeled 'Record VTA'). Area X is the nucleus of the basal ganglia where the process of learning songs takes place. To ensure that the recorded neurons projected in Area X, the scientists used antidromic stimulation. This process involves electrically stimulating neurons in Area X, forcing electrical signals to travel backward down the axon. If a neuron fired with a very fast latency, it meant that the neuron was indeed connected to Area X and could therefore be used in the experiment.

A total of 14 neurons were correctly tested, with 13 faulty. The experiment involved a computer program that recorded a specific syllable from the bird's song and projected a distorted version of it 50 milliseconds later. As a result, the bird heard the song's syllable at the wrong time, causing it to think it was singing wrong. 44% of the time, the computer would transmit the faulty sound. For the remaining 56% of the time, the animal heard its own undistorted song. The experiment's result showed that although the bird sings many syllables one after another with very little pausing, the brain is fast enough to interpret 1 syllable at a time. If the bird's brain could not detect precisely where the error was made when it sang a syllable wrong, the bird would have an overall idea that the sound was wrong, but would not be able to pinpoint it and correct it. However, as mentioned, the results show that the signal is precisely timed to each syllable and to the correctness of the feedback sound. This is shown in Figure 2I of the Gadgakar et Al. 2016 paper [9], which examines how the peak firing time of each VTA_error neuron when the sound was undistorted aligns precisely with the targeted moment in the song. The results, therefore, suggest that the error signal is time-locked to the targeted syllable. The production of specific syllable time-locked error calculating signals demonstrates that the RPE is applicable beyond external rewards. The algorithm remains the same regardless of whether the reward is internal or external.

By observing the particularly small latency between the sound feedback and dopaminergic response: activation for undistorted and suppression for distorted, the maintenance of RPE is reinforced. The dopamine neuron reaction occurs in 50 to 125 milliseconds following the sound feedback, with an average of 58 milliseconds for suppression and 51 milliseconds for activation. The response is roughly three to six times faster than the blink of an eye. The presence of RPE is further evidenced in the double-syllable-targeted section of the paper, where not one but two syllables are distorted: syllable B and D.

In Figure 3 D & E of the Gadgakar et Al. 2016 paper [9], two panels are shown: D represents the distorted

sound, and E represents the undistorted sound. Both panels have a top and a bottom section. The top section has the y-axis representing the normal firing rate in seconds, and the x-axis showing the target onset syllable. Two different lines are presented: the high distortion probability of approximately 49%: target -1, and the lower distortion probability of approximately 20%: target-2. The graph is relative to the normal neural baseline of 1.0. Values below it indicate suppression, and values above it indicate firing. The bottom graphs are scatter plots, where each triangle represents an individual neuron. The graph addresses the question: how long did it take for a neuron to activate after the target appeared? The experiment also compares reaction times across neurons, investigating whether some neurons respond more strongly to certain targets than others. Panel D shows, in the top graphs, suppression in both syllables, with similar minimum values. The scatter graph in panel D is congruent, showing that all neurons have relatively low values (under 1.0).

Therefore, regardless of whether the probability of it occurring and the prediction of its occurrence is higher or lower, a distorted sound will result in similar suppression. Looking at both the top and bottom graphs of panel E (undistorted), we can see that neurons in target-1 have a higher activation compared to target-2 neurons. Signaling that neurons are strongly correlated with the probability. Firing is much more significant when something bad was expected, and the feedback was better than expected. This probability-based scaling directly reflects the discovery by Schultz [13] in primates: when a reward is expected, the dopaminergic response approaches zero; when no reward is expected, but a reward is given, dopamine spikes. The similarity of results in the Shultz [13] and Gadgakar 2016 [9] papers not only shows that RPE is present in internal benchmark situations. Further strengthening the argument that RPE may be broadly conserved as a universal mechanism.

The Gadgakar (2016) [9] paper demonstrates that dopamine cannot be reduced to a simple reward chemical; rather, it should be considered in its entirety. Dopamine neurons seem to be permanently asking themselves: how does what I expected differ from what I got? What is important to conclude from this experiment, however, is that it does not merely mean that dopamine neurons are “good” or “bad” buttons. Instead, they have a complex algorithm that constantly calculates the degree of prediction error. The Gadgakar (2016) [9] paper does not merely suggest but directly demonstrates that this calculation can be applied to various contexts and species: from a primate getting juice to a bird singing well. This answers one aspect of the research question: RPE is maintained across distinct types of reward learning. However, whether RPE extends to social and parenting contexts in which the reward is neither an externally nor an internally set goal but is instead based purely on another individual remains unknown.

III . Dopamine In Social And Parenting Contexts

The previous section has determined that the brain uses reward prediction error in two contexts: external and internal rewards. However, that alone does not determine RPE’s universality. Naturally occurring contexts such as social and parenting should also be considered. The two previous contexts were rather unidimensional, involving only the agent itself. Parenting and social contexts are far more intricate; the reward depends on another individual’s behavior. Social and parenting behaviors involve distinct factors, including touch, sound, movement, and emotion. Feedback can be positive, negative, or even neutral. If the concept of RPE can, in fact, be applied to social contexts, then its universality is proven. To address

RPE's generalizability across social contexts, the paper will examine two experimental papers: Gadgakar et al. (2023) [11] on zebra finches and Shea et al. (2023) [10] on mouse and pup retrieval.

The Gadgakar et al. (2023) [11] paper builds on a previously published paper observed in section III (Gadgakar, 2016), which analyzed zebra finches while singing. In the 2016 paper, the zebra finches were primarily focused on the quality of their singing, whereas in the 2023 paper, the songbirds were also exposed to a female (courtship situation). The songbirds stopped responding to whether their singing was better or worse than expected; instead, they began to pay attention to the female they were courting. The brain area primarily focused on singing (Area X) shifted from evaluating its own singing to the female's call. Figure 3G of the Roeser, Gadgakar et Al. 2023 paper [11] compares the error signal in Area X when zebra finches sang alone, versus when they sang to a female. Each grey line in the graph represents a single recording site. The black lines represent the average response. The graph's y-axis shows the strength of dopamine neurons' response to the difference between expected and actual sound. Most of the lines in Figure 3G of Roeser, Gadgakar et Al. 2023 [11] slope downwards, indicating that dopamine neurons become less preoccupied and responsive to the bird's own song quality during courtship.

Conversely, Figure 4E of the Roeser, Gadgakar, et al, 2023 paper [11] illustrates the dopamine activity relative to a female's response during courtship. The y-axis represents the intensity of the neural reaction. An increase in dopamine activity can be seen when a female replies during courtship singing compared to when they occur during non-courtship singing periods. Figure 4E of the Roeser, Gadgakar, et al, 2023 paper [11] presents results for two brain regions that were tested: Area X and MST. The results show stronger neural response in Area X - basal ganglia nucleus involved in song learning, and where dopamine signals have been observed so far. Contrary to MST, the medial septum, another recorded brain region, does not exhibit the same courtship-dependent modulation. This difference in results suggests that the observed effects are specific to Area X, ruling out the possibility that it is simply a broader neural response to female vocalizations.

Together, these figures suggest that dopamine neurons do not merely respond to social rewards. Rather, the brain evaluates contexts to decide what matters most in a particular moment and responds accordingly. This concept is known as retuning. The concept of retuning directly addresses the "to what extent is RPE adapted" section of the research question. The RPE mechanism (expected vs. actual outcome) remains the same, but the algorithm's input changes. The relevant outcome shifts depending on the agent's priority.

These findings, however, did not address whether dopamine neurons encoded RPE during social interactions or simply fired in response to a social reward. The Shea paper [10] on dopamine activity during pup retrieval in mice uses a trial-by-trial design to address this. The paper may have used mouse pup retrieval because it is structurally analogous to previous experiments in which animals repeated an action to obtain a specific reward. The mice will become increasingly accustomed to retrieving their pups, thereby improving their performance. The setup enabled researchers to measure factors such as the mice's improvement in retrieval skills and whether the mice expect this. Consequently, proving the presence of RPE, rather than simple dopamine responsiveness to a social reward. The Shea et al. [10] paper uses fiber photometry with the calcium sensor jGCaMP7f, expressed exclusively in the Ventral Tegmental Area

(VTA), where dopamine neurons are infected, to record phasic dopamine activity during pup retrieval via fluorescence fluctuations.

Figure 1D of the Shea et Al. Dopamine and Parenting 2023 paper [10] demonstrates various dopamine activation (signaled by the four red arrows) each time the mice made contact with their pup. The x-axis is time (measured in seconds), and the y-axis is a change in fluorescence percentage (relative to baseline fluorescence).
Crucially, dopamine neuron activation does not necessarily indicate RPE learning. Dopamine neurons can be activated for diverse reasons, including motor vigor, arousal to an important event, or salience from seeing the pup. However, an evaluation of timing will confirm the cause of dopamine activation. The dopamine bursts lasted 2-3 seconds and were time-locked to the exact moment of contact with the pups. The timing ruled out motor performance and arousal, which would produce longer activation starting before contact. This suggests that dopamine activity is not related to motor behavior but rather to the behavioral outcome. Connecting back to the Roeser, Gadgkar et al. (2023) [11] paper: the bird's dopamine neurons would not fire when the bird sang (when the action was made), but rather when the female answered to his call. Here, the dopamine activity was tied to the social outcome rather than the motor behavior. These two experimental papers show how the brain can differentiate between the action and the outcome. In the context of social interaction, dopamine neurons respond to whether an important social event occurred rather than to the motor behavior to achieve it. Hinting that dopamine response is governed by the same mechanism: RPE.

Although establishing that dopamine neurons respond to outcomes rather than to preceding actions is important, this alone is not sufficient to prove the presence of RPE in social contexts. RPE involves an evaluation process in which neurons compute the difference between the expected and actual outcomes. If RPE does in fact apply to social contexts, it must decrease as the reward becomes predictable. Two key factors evidence this in the Shea paper [10]. The first evidence appears in a trial where two groups of mice were tested. The first group would only be exposed to approximately 5 pup retrieval exercises per day. The other group (the highly experienced group) was exposed to 24 pup-retrieval exercises per day. The results are shown in Figure 2C of the Shea et AL. Dopamine and Parenting 2023 [10].

The figure includes top heatmaps and bottom line graphs. The top graphs have the x-axis time, measured in seconds, from -4 to 4 seconds, relative to the white dashed bar (0s), which represents the exact moment of contact with the pup; the y-axis represents different trials (each row is a trial). The dark purple color represents low dopamine activity, whereas the bright yellow represents higher dopaminergic neural activity. What one should notice when observing the top trial of the non-co-house retrieval (the mice exposed to 5 exercises/day) is that the yellow color seems to cluster around the dashed line (representation of the exact moment of contact with the pup); this pattern is preserved throughout the days (P0, P1, P2, P3, P4, P5). When comparing it to the "high experience non-co-housed" mice, we can see that the pattern disappears. The bright yellow clusters at the retrieval point on the first day, but as the days go by, the dopamine neurons eventually stop responding to the pup retrieval. The same thing can be seen when observing the bottom (line graphs). The x-axis is still time in seconds relative to the dashed bar (pup retrieval), and the y-axis is the change in fluorescence relative to the baseline (showing the dopamine

activity level). In the less-experienced mice (top-line graphs), the fluorescence spikes at the point of contact with the pup on all days, whereas in the mice that are exposed to more trials per day (bottom-line graphs), the spike decreases over time. This claims that dopamine neurons respond not to the reward (which remains the same over days) but to whether the reward is expected. Further proving that dopamine neural activity is not a reward-based response but a prediction error response. The RPE framework is, in fact, applied to social contexts just as in primates getting juice and the singing quality of songbirds.

A second factor of Shea's experiment that supported the presence of RPE was the cued retrieval task [10]. In this experiment, the mice were cued to the pup's arrival. They learned that every time the door opened, the pup would appear. The door opening became the reward prediction, whereas the pup retrieval itself was merely the outcome. In Figure 4H of the Shea et Al. Dopamine and Parenting 2023 [10] , we can see a comparison of the neural behavior of the mice during the early and late learning stages.

The top graphs are a heatmap, where the y-axis (rows) represents each mouse (individual trials). The darker colors represent low dopamine activity, and the brighter ones demonstrate higher dopaminergic neural activity. The first dashed line represents the exact moment at which the door opens, whereas the second dashed line refers to the exact moment at which the mouse makes contact with its pup. When looking at the early learning mice, one can see a strong correlation with brighter color at the retrieve line, whereas there is little to no bright yellow at the door-opening dashed line. This is because the mice have not yet learned that opening the door signals the pup's arrival.

On the other hand, if we consider the late-learning-stage trials, we can see that the correlation shifts away from the "retrieve" line and towards the "door" line. This shows that the mice are learning to associate the door opening with the upcoming social reward. The same pattern is evident in the bottom-line graph, where the x-axis shows time on a relative scale and the y-axis shows the strength of dopamine-related neural activity. The black line represents the early stage, in which the mice have not yet learned and therefore exhibit a dopaminergic burst at retrieval. The red line represents the later stage, in which the mice have learned to associate the cue with the upcoming event and therefore show a stronger dopamine response after the door opens. This directly mirrors Shultz's discovery that dopamine activity migrates from reward to the earliest predictive cue [13]. The application of this temporal back-propagation occurs upon contact with the pup (social reward), in the same way as it does in primates with juice, once again confirming the universal learning algorithm. However, the most prominent evidence is found in the catch trials, which involve surprising the mice with the opposite of what they expected. Figure 5B of the Shea et Al. Dopamine and Parenting 2023 paper [10] shows the results.

At the top, we can see heatmaps representing the dopaminergic response to what happened. There were two scenarios: the first was represented in the left heatmaps, and the second in the right heatmaps. The left heatmap portrays the mouse finding its pup when entering the room. The right heatmaps represent the contrary. Two different groups of mice were exposed to the test. One group received negative conditioned

stimuli (CS-), in which the mice did not expect to receive a pup. The other group received a positive conditioned stimulus (CS+), in which the mice learned that they always got their pup and therefore expected to find it after the cue. The white dashed line in the heatmaps represents the moment the mouse entered the chamber. When observing the left heatmaps and comparing the results of CS- and CS+, we can see that the CS+ mice, who expected to find their pup, showed a mild dopamine response, whereas the CS- mice showed a dopamine burst. Yet the reward in both cases was the same. The dopamine neurons are therefore responding to the surprise of the reward. In the second scenario (right heatmaps), the mice did not find their pup. In the top CS- heatmap the mice were not expecting to find their pups and did not find them, resulting in no dopamine response. In the bottom heatmap, we can see significantly darker colors indicating dopamine suppression when the mice expected to find their pups and did not. The reward was the same across all cases, yet the dopamine response seemed to vary depending on whether the social reward exceeded, fell short of, or matched expectations. Dopamine neurons thus responded not to the presence of the pup but rather to the degree to which it was surprising. This correlates with the exact definition of the RPE system [14].

The catch trials are crucial in the Roeser, Gadagkar et al. paper because they illustrate the precise classical omission and surprise manipulations used to characterize dopamine neurons in previous primate studies. Instead, its principle extends to external rewards, internal performance evaluations, and social interactions across species such as primates, songbirds, and mice. The evidence that the same computational mechanism is applied across all scenarios strongly supports the view that RPE is generalizable rather than context-specific.

IV . Dopamine Signaling Connection To Learning & Plasticity

The previous sections have established that dopamine neurons compute reward prediction error across different contexts (external, internal, and social rewarding). However, for RPE to be fully considered a universal learning mechanism, as the research question asks, it must have a downstream system that translates the information and applies it. Encoding information is not the same as using it. How do dopamine neurons allow the brain to learn and therefore adjust future actions? The translation occurs through the basal ganglia –a group of subcortical nuclei with a crucial role in: motor learning, behavioral prioritizing, and selection of action– [15]. The signaling from dopamine neurons travels along their axons to the striatum, a crucial structure in the basal ganglia. This area is called Area X in songbirds. [16]. Mammals' striatum receives dopamine signaling from the Ventral Tegmental Area, and the substantia nigra (or from Area X in songbirds). Once the information is received, the brain not only reads it but also changes in response to the RPE. The dopamine signaling pathway involves three main factors: the cortex, the striatum, and behavior. When a behavior produces a better-than-expected outcome, dopamine strengthens the synaptic connections between the cortical inputs and striatal neurons that were activated during the behavior. Conversely, the synaptic connections are weakened when the outcome is negative (worse than expected). This process enables the brain to encourage rewarding behavior and discourage the unrewarding one. What is interesting here is that synaptic plasticity is context-independent. Regardless of the behavioral domain, the same dopamine signaling is produced. The basal ganglia do not differentiate between signaling the arrival of juice in primates, bird singing evaluation, or pup retrieval in mice. The reviewed evidence indicated that the basal ganglia may apply the same system regardless of context:

strengthened when a positive error occurs, weakened when a negative error occurs. This directly addresses the universality part of the research question. However, it does not enable any generalized conclusions to be made.

The Gadgakar et al. (2016) [9] study clearly illustrates the existence of dopamine-driven strengthening or weakening of corticostriatal signaling, allowing the songbird to connect each specific syllable to a moment in a song and, therefore, reproduce the vocal sound that felt rewarding [9]. This is the same flexible mechanism that appears in the monkey-and-juice experiment. The circuit's structure remains the same, and only the content being learned changes. Furthermore, not only does dopamine allow animals to learn, but it can retune itself depending on what the animal treats as success in the specific moment. This builds a direct bridge between plasticity and the retuning mechanism previously discussed. There is a specific role-devising mechanism that enables the dopaminergic system to guide learning across all contexts and species examined in this paper: dopamine changes what is being evaluated depending on the agent's context and current goals. At the same time, the basal ganglia (the downstream learning circuit) allocates the same plasticity across all contexts. Consequently, the dopamine system can be considered flexible in relation to the agent's context. In contrast, the universal point across all contexts examined is the plasticity mechanism - the final piece of the RPE framework.

With this established, the research question can take a step forward: what happens when this system, which has proven to be flexible and so robust until now, in different natural and healthy contexts, is disrupted by dopaminergic signaling malfunction? Is the RPE algorithm preserved?

V. Addiction And Its Abnormal Dopamine Activity

The previous sections have examined the RPE mechanism, its maintenance across different contexts, and how the basal ganglia translate dopamine signaling into learning through a universal rule. However, the paper has examined the system under healthy conditions, ruling out the possibility of dopamine signaling disruption. If RPE is truly a foundational method for learning - as argued so far - disrupting dopamine signals while leaving the computation intact must produce predictable and destructive outcomes. If RPE is the ultimate mechanism that enables our brains to digest the consequences of actions, adjust our future behavior accordingly, and predict the feedback, what would happen if it were disturbed? Addiction is a perfect example. Often misunderstood, addiction does not derive from a failure of willpower nor a simple pleasure-seeking behavior; instead, addiction is a reinforcement learning process that has mistakenly gone awry [17] [18].

As discussed earlier, the critical function of reinforcement learning is to actively calculate the difference between expectations and reality and adjust accordingly. The system can therefore be considered self-correcting. As the predictions approach the actual event, the error signal approaches zero and eventually plateaus. There is a particular elegance to the equilibrium at the heart of classic reinforcement learning. As established in section II, there is a convergence: as the brain learns to predict a reward, the dopamine activity approaches zero. All the scenarios explored in this paper so far have aligned with this convergence. However, the perfect equilibrium appears to be disrupted in substance addiction. The natural

convergence of dopamine signaling is disrupted, whilst the rest of the RPE machinery continues to function as before. Addiction varies from all the previous examples this paper has observed. The dopamine receptors never encode an equal prediction. Dopamine neurons consistently interpret substance intake as better than expected, regardless of how accurately the brain predicts its arrival. All substances produce a phasic increase in dopamine that bypasses prediction-error computations and disrupts the circuit. While cocaine blocks dopamine reuptake, nicotine stimulates dopamine release.[19] The disruption occurs because pharmacological effects are induced, creating an artificial, surprising signal. The brain is unable to distinguish between genuine prediction error and artificial signals produced by drugs. This is crucial to answer the research question. The RPE system is maintained, the learning rules operate in the same way, the only difference is its inability to adapt - converge to accurate prediction and redirect the signaling to the next priority, because the drug precludes the convergence required for adaptation to occur.

In Figure 2 of the Reddish Addiction 2004 paper [17], the x-axis represents the cost of obtaining a certain reward, and the y-axis represents how often the reward was chosen relative to when the cost was zero. The black dotted line represents the drug pathway, whereas the grey line represents the natural reward pathway. Both lines seem to decrease as the cost of the reward increases; the grey line decreases significantly more than the black one (slope −0.8 versus 1.5). The drug does not increase as much as the natural reward because the drug's value continues to increase indefinitely, whereas the natural rewards and the cost reach a limit. Consequently, the penalties of substance use become an ever-shrinking fraction of the value as a whole. The increasing values in addiction produce a lock system with a single priority that any consequence, cost, or alternative reward may not override.

This contrasts with what happens in natural behaviors. In the zebra finch experiment, the birds' priorities shifted from singing accuracy to courtship, portraying how the brain can recalibrate and prioritize its most important goal. The most evident indicator that addiction corrupts the RPE mechanism from the Redish paper [17] is the "double dopamine signal" principle. Just as established in section II and again in the Shea analysis, an agent attributes a cue to the arrival of a reward [10]. In the case of drug use, the brain still undergoes this mechanism – it associates a cue with drug delivery. However, due to neuropharmacological mechanisms, the dopamine level does not stabilize and continues to increase as the agent receives the drug. Here, the RPE framework is still maintained. The system is trapped in a permanent state of positive error. Producing a double dopamine signaling: one at the cue and one at the drug delivery. Across all the dysfunctional factors explored - unbound value inflation, insensitivity to cost, and double dopamine signaling - the RPE framework is maintained. The same mechanism is working, making the same calculation, except this time the input is falsified. Addiction does not break but rather exploits RPE. [20]

VI. Parkinson's And Its Suppression Of Dopamine Signaling

The paper has determined so far that RPE is still present in addiction. But what would happen if dopamine neurons were to act oppositely? If we lost our dopamine neurons? Parkinson's disease reveals this exact scenario: a near-total loss of all dopaminergic learning signals [21, 22]. While addiction

overwrites flexible learning by an ever-increasing drug-dominated value, Parkinson's disease forbids the system from learning by diminishing dopamine signals. In Parkinson's disease, the dopamine neurons located in the substantia nigra degenerate. These neurons are the same ones responsible for encoding RPE via phasic patterns [21]. As established in previous sections, dopamine neurons enable the basal ganglia to learn from previous experiences and choose actions. When dopamine neurons no longer exist, the brain experiences learning difficulties, as well as difficulties with decision-making and movement [22].

In 2010, Kravitz et al. used a technique where specific neurons were turned on or off using light: optogenetics [20]. Kravitz et al. selectively expressed a light-sensitive protein - originally found in green algae -, channelrhodopsin-2, in either the direct pathway neurons: D1 neurons (the brain's Go system) or the indirect pathway neurons: D2 neurons (the brain's Stop system) [20]. When the indirect pathway neurons were activated, the mice began moving slowly, with bradykinetic symptoms resembling those of Parkinson's disease. On the contrary, direct neural activation of the pathway led to greater movement and easier action initiation. Figure 3C of the Kravitz et al. 2010 paper [20] shows two side-by-side bar charts. The first one, entitled D1-ChR,2, illustrates the direct pathway (red), whereas the D2-ChR2 represents the indirect pathway (green). The y-axis represents the percentage time the mice spent in each state (frozen, fine movement, ambulation). The x-axis has three sections: Pre, before the laser; Laser, during the laser; and Post, after the illumination. In the D2-ChR2 panel - indirect - we can see that while ambulation significantly decreases, the mice go from spending little time frozen to spending most, if not all, of their time immobile, representing optogenetically induced Parkinson's. Conversely, the D1-ChR2 direct pathway shows freezing during illumination, along with increased ambulation. However, both behaviors appear to return to baseline once the laser is turned off, thereby indicating reversibility. This portrays that the behavioral state is not a consequence of lasting damage but rather a moment-to-moment activation or deactivation of pathways. Referring back to the research question, the real-time presence or absence of pathway activity directly mirrors the normal dopamine signaling that regulates the circuit.

The scientists later genetically engineered mice with Parkinson's disease, removing most, if not all, of their dopamine neurons. The direct pathway was then stimulated through optogenetic activation, rehabilitating the mice's motor behavior [21]. Crucially, demonstrating that the motor circuit itself is not affected in Parkinson's, instead, the signals required to tip the balance towards action are dysregulated. Figure 4E of the Kravitz et al. 2010 paper [20] illustrates this by comparing the y-axis - percentage of time spent freewing, in fine movement, or in ambulation - across conditions (before and after the mice's dopamine was destroyed). The black bars represent the mice in a healthy state. The mice show little to no freezing, normal fine motor movements, and extensive ambulation. Conversely, the red bars show the mice's behavior after the dopamine neurons were removed, before laser illumination.

The graph shows a lot of time spent frozen, with hardly any ability to walk. Once the laser is turned on, the direct pathway is artificially activated. As a result, the mice's behavior recovers almost completely to the baseline. When light is turned off, the mice encounter Parkinson-like symptoms. The striking resemblance between the behavior when light was turned off and the mice's normal, healthy behavior suggests that dopamine loss does not physically damage the motor circuit. Instead, the results show that the circuit is solely missing correct signaling. If damage to the motor circuit were present, artificially

stimulating the direct pathway would have failed to restore normal behavior. It is therefore clear that dopamine does not create movement; it informs the basal ganglia of the actions they should take, depending on their value. In the absence of dopaminergic signals, optogenetics can artificially activate the direct pathway, ultimately replacing dopamine and enabling normal behavior. The neural deficit in Parkinson's can therefore be considered informational instead of structural. Parkinson's is not merely a movement disorder but also a learning disability. Dopamine acts as a context-dependent signal that is strengthened or weakened depending on RPE. When an action creates a better-than-expected outcome (+ RPE), there is a burst in dopamine activation. This dopaminergic firing activates D1 receptors in the direct pathway MSN's. The pathway is consequently strengthened, increasing the likelihood of it happening again. Conversely, when a worse-than-expected outcome occurs (-RPE), neural firing decreases. This provokes the same decrease through D1 receptor activation by D2 receptor in the indirect pathway MSN. Strengthening the suppression of the unsuccessful action and preventing it from happening again. These processes enable the basal ganglia to determine which actions shall and shall not be repeated in the future. In Parkinson's disease, these dopamine signals are absent.

As a result, neither the positive reinforcement learning nor the negative teaching signals are delivered. Positive actions can no longer be reinforced through the direct pathway (D1), and unsuccessful actions can no longer be suppressed and prevented through the indirect pathway (D2). Therefore, Parkinson's not only prevents one from executing an action but also disrupts the reinforcement learning processes by hijacking the brain's ability to learn action-outcome contingencies - the process of learning which action to reproduce through trial and error. Often, a patient may not be able to execute an action, but they know which action they want to take. In Parkinson's, this is not the case. The loss of dopamine signaling in the RPE also causes the patient's internal models to degrade [22]. Parkinson's physical impairments, such as tremor and bradykinesia, mirror internal deficits, including disruption of reinforcement learning, inability to make decisions, and lack of motivation [14].

The paper has so far established that, in the absence of dopamine neurons and, therefore, their signaling, the system is unable to update the values of each action. Therefore, an agent's inaction is not caused by its motor incapacity but rather by its inability to adjust value to an action. The paper has also shown that in Parkinson's disease, the RPE framework is still maintained; what differs between contexts is how the receiving circuit interprets and applies the signal. Overall, Parkinson's disease underscores the broad implications of RPE across contexts, showing that it operates whenever the brain must decide which action to take, including the simplest decisions, such as whether to move or remain still.

LIMITATIONS

Five main limitations can be held against this paper. Firstly, a certain degree of authorial judgment was involved despite the inclusion criteria set, due to the narrative literature review structure. A systematic review would not have such a limitation. As a result, it can be argued that another study selection would have led to a somewhat different interpretation and conclusion. Secondly, the evidence reviewed consisted of experimental animal models and was used in the paper to draw conclusions about human cognition and

behavior. Although the agents used (non-human primates, mice, and songbirds) have played a pivotal role in the understanding of dopaminergic learning, caution is warranted regarding the relationship between the findings and actual human functioning. Thirdly, studies reporting newer, positive, and statistically significant results are more likely to be published than null findings. This is a crucial point regarding publication bias and its effect on the accessible evidence. Fourthly, the RPE research is a rapidly evolving field. This paper encapsulates only some relevant studies (according to the author and the inclusion criteria) at a single point in time. Further discoveries could complicate our understanding by showing, for example, that dopamine signals can convey more than RPE. For example, there are competing ideas in the dopamine field that changes in dopamine signaling are not always related to reward outcomes and instead to general movement patterns such as the speed with which a mouse moves their head [23]. These data complicate narrow perspectives on dopamine function that it can only subserve RPE signaling and invite further inquiry. Lastly, due to the heterogeneity of the studies cited (different species, contexts, and methodologies), a certain caution must be exercised, as the results cannot be directly compared with one another.

CONCLUSION

In conclusion, based on evidence from various experimental papers, one can deduce that reward prediction error is, in fact, maintained as a computational principle across the species and contexts observed in this paper. However, it is not to be considered a monolithic signal that responds to everything in the same way. Instead, RPE exhibits plasticity, enabling it to adapt to the demands of a particular agent in each context. The adaptation occurs through the process of retuning, in which dopamine neurons shift what they evaluate based on the agent's current needs, and circuit routing, in which specific signals are directed to the brain area most appropriate for the situation. Understanding the RPE mechanism is crucial not only to better comprehend the mechanism that drives learning but also to address pathologies caused by dopamine malfunctions that affect millions around the world. These pathologies include addiction, where the dopaminergic system is corrupted, and rather than normally stabilizing after learning, dopamine levels escalate indefinitely. Conversely, Parkinson’s disease portrays the consequences of losing dopamine signaling. Without RPE to guide basal ganglia plasticity, the basal ganglia can no longer update behavior, causing impairments in movement and motor learning.

ACKNOWLEDGEMENTS

Thank you to Jesse H. Goldberg, M.D., Ph.D., mentor from Cornell University, for his guidance in developing this research paper and for inspiring me with his work on dopamine and RPE. Thank you to my mother and my brother for teaching me to be passionate and encouraging me in the pursuit of those passions.

REFERENCES:

[1] Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology*, *191*(3), 391–431. https://doi.org/10.1007/s00213-006-0578-x

[2] Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: "liking", "wanting", and learning. *Current Opinion in Pharmacology*, *9*(1), 65–73. https://doi.org/10.1016/j.coph.2008.12.014

[3] Watson, S. (2024). *Dopamine: the Pathway to Pleasure*. Harvard Health; Harvard Medical School. https://www.health.harvard.edu/mind-and-mood/dopamine-the-pathway-to- pleasure

[4] Franco, R., Reyes-Resina, I., & Navarro, G. (2021). Dopamine in Health and Disease: Much More Than a Neurotransmitter. *Biomedicines*, *9*(2), 109. https://doi.org/10.3390/biomedicines9020109

[5] Alimoradi, Z., Lotfi, A., Lin, C. -Y ., Griffiths, M. D., & Pakpour, A. H. (2022). Estimation of Behavioral Addiction Prevalence During the COVID-19 Pandemic: A Systematic Review and Meta-analysis. *Current Addiction Reports*, *9*. https://doi.org/10.1007/s40429-022-00435-6

[6] F. Minkove, J. (2022). *New Research and Insights into Substance Use Disorder*.Hopkinsmedicine.org.https://www.hopkinsmedicine.org/news/articles/2022/06/new-research-and-insights-into-substance-use-disorder?

[7] *Parkinson's disease: A looming pandemic*. (2017). ScienceDaily. https://www.sciencedaily.com/releases/2017/11/171113122956.htm

[8] Neurosity. (2026, February 14). *Reward Prediction Error: The Dopamine Surprise Signal*. Neurosity.co; Neurosity. https://neurosity.co/guides/reward-prediction-error-dopamine

[9] Gadagkar, V, Puzerey, P. A., Chen, R., Baird-Daniel, E., Farhang, A. R., & Goldberg, J. H. (2016). Dopamine neurons encode performance error in singing birds. *Science*, *354*(6317), 1278–1282. https://doi.org/10.1126/science.aah6837

[10] Xie, Y, Huang, L., Corona, A., Pagliaro, A. H., & Shea, S. D. (2023). A Dopaminergic reward prediction error signal shapes maternal behavior in mice. *Neuron*, *111*(4), 557-570.e7.

https://doi.org/10.1016/j.neuron.2022.11.019

[11] Roeser, A., Gadagkar, V, Das, A., Puzerey, P. A., Kardon, B., & Goldberg, J. H. (2023). Dopaminergic error signals retune to social feedback during courtship. *Nature*, *623*(7986), 375–380. https://doi.org/10.1038/s41586-023-06580-w

[12] Olds and Milner 1954 "POSITIVE REINFORCEMENT PRODUCED BY ELECTRICAL STIMULATION OF SEPTAL AREA AND OTHER REGIONS OF RAT BRAIN"

[13] Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

[14] Schütt, H. H., Kim, D., & Ma, W. J. (2024). Reward prediction error Neurons implement an efficient code for reward. *Nature Neuroscience*, *27*(7), 1333–1339. https://doi.org/10.1038/s41593-024-01671-x

[15] Young, C. B., Sonne, J., & Reddy, V (2023, July 24). *Neuroanatomy, Basal Ganglia*.Nih.gov; StatPearlsPublishing.https://www.ncbi.nlm.nih.gov/books/NBK537141/

[16] Rice, M. E., Patel, J. C., & Cragg, S. J. (2011). Dopamine release in the basal ganglia. *Neuroscience*, *198*, 112–137. https://doi.org/10.1016/j.neuroscience.2011.08.066

[17] Redish, A. D. (2004). Addiction as a Computational Process Gone Awry. *Science*, *306*(5703), 1944–1947. https://doi.org/10.1126/science.1102384

[18] *Why our brains are wired for addiction: What the science says*. (2025). News Center. https://med.stanford.edu/news/insights/2025/08/addiction-science-huma -br ain-ancient-wiring.html

[19] Volkow, N. D., Fowler, J. S., Wang, G. J., Baler, R., & Telang, F. (2009). Imaging dopamine's role in drug abuse and addiction. *Neuropharmacology*, *56*(1), 3–8. https://doi.org/10.1016/j.neuropharm.2008.05.022

[20] Kravitz, A. V, Freeze, B. S., Parker, P. R. L., Kay, K., Thwin, M. T., Deisseroth, K., & Kreitzer, A. C. (2010). Regulation of Parkinsonian motor Behaviors by optogenetic control of basal ganglia circuitry. *Nature*, *466*(7306), 622–626. https://doi.org/10.1038/nature09159

[21] Zhou, Z., Yi, L., Wang, Q., Lim, T. M., & Tan, E. K. (2023). Role of Dopamine in the pathophysiology of Parkinson's disease. *Translational Neurodegeneration*, *12*(1). https://doi.org/10.1186/s40035-023-00378-6

[22] Triarhou, L. C. (2013). *Dopamine and Parkinson's Disease*. National Library of Medicine; Landes Bioscience. https://www.ncbi.nlm.nih.gov/books/NBK6271/

[23] Barter, J. W., Li, S., Lu, D., Bartholomew, R. A., Rossi, M. A., Shoemaker, C. T., Salas-Meza, D., Gaidis, E., & Yin, H. H. (2015). *Beyond reward prediction errors: the role of dopamine in movement kinematics. Frontiers in Integrative Neuroscience, 9. https://doi.org/10.3389/fnint.2015.00039*

KEY TERMS

Dopamine	A neurotransmitter associated with motivation, movement, pleasure, and learning.
Synapse	A tiny gap between two neurons. Also known as a functional junction.
Neurotransmitters	The body's chemical messengers. Delivering messages across gaps between neurons - synapses -.
Reward Prediction Error (RPE)	The difference between expected and actual outcome.
Reinforcement Learning	Process that teaches an agent which actions to repeat or not based on trial and error learning.
Dopaminergic Signaling	The process by which dopamine travels from one neuron across a synapse to bind to a receptor of another neuron, consequently influencing the activity of the brain.
Pup Retrieval	The process during which a mother mouse locates her pup, makes contact with it, and brings it back

	to its nest. This is known as parenting behavior.
Phasic dopamine activity / phasic activation	Refers to a brief and fast dopamine neuron activation occurring at the exact moment when the brain locates an important outcome.
Cue / conditioned stimulus / CS	Refers to a sensory signal which an agent's brain learns to associate with an upcoming outcome.
Dopamine reuptake	The naturally occurring process which the brain removed dopamine left in the gap between neurons.