International Core Journal of Engineering Volume 7 Issue 10, 2021
ISSN: 2414-1895 DOI: 10.6919/ICJE.202110_7(10).0059

Effect of Reversed Associative Learning Using Expectation Reward Prediction Error on Neuron Circuits

Jerry Y. Liu*

Canterbury School, New Milford, Connecticut, 06776, United States.

*jliu22@cbury.org

Abstract

Associative learning is a process triggered by a stimulus followed by a reward or punishment. The subject often forms an expectation to future events after receiving a controlled stimulus, and when its prediction is challenged, dopamine is released and new memories formed. In this experiment, we reverse the conventional procedure of "stimulus then reward" and examine the impact of such change in the subject's learning ability and neuron circuit activity. We predicted three different outcomes for the learning ability pattern: no learning, improved learning, and counterintuitive response. The result of this experiment may reinforce, challenge, and redefine pre-established reward prediction error theory. The experiment was established surrounding two keywords: reversed and expectation. Two different odors (O_1 & O_2) are given to two groups of house mice as CS before and after the reward (5-milliliter sugar water), each with 20% chance and 80% chance. A tetrode is placed in the mice ventral tegmental area (VTA) to monitor the activity of dopamine neurons and GABA neurons, which will provide the readers a direct and intuitive visual image connection between the learning ability and dopamine secretion.

Keywords

Dopamine; Expectation; Reversed Associative Learning; Prediction Error; Probability.

1. Introduction

Associative learning is a common learning approach that allows organisms to respond to complex stimuli. [1] The conventional procedure of associative learning known as Pavlovian learning includes a controlled stimulus (CS) followed by a reward [2,3], and the stimulus signal sent to a subject will be compared with its expectation, the difference will result in associative learning [4,5]. This specific associative learning method is known as reward prediction error, a process that's predominantly driven by dopamine. [6,7] Therefore we deduce the learning procedure can also be manipulated by alternating the types of CS and neuron circuits, using psychiatric medications, as they directly affect the dopamine secretion [8].

Abundant experiments focusing on dopamine activity within local neuron circuitry show unexpected rewards trigger far more concentrated and exciting responses compared to expected rewards; in fact, as training moves on, most dopamine bursts are fully suppressed once subjects adopt the pattern. [9-11] As the name suggests: associative learning. It's also shown that most subjects' dopamine burst triggers shortly after the CS instead of the reward once the subjects establish the relationship between CS and reward [12].

Similar behavior was observed among different types of subjects receiving various stimuli and rewards [13-15]; however, many theories claim to result in associative learning; however, a recent experiment on crickets established an affirmative connection between reward prediction error theory

ISSN: 2414-1895 DOI: 10.6919/ICJE.202110_7(10).0059

and associative learning. [16] Beyond proving a theorem, other experiments expanded the research by studying the dopamine firing pattern when many different factors were involved, such as practicing different CS, using drugs to inhibit GABA neurons, changing reward size, and adding a short time gap between CS and reward. [16] A phenomenon in their research demonstrates how subjects respond toward rewards that are released with varied probabilities: In one of their research, the reward was associated with odor A and B, when subjects receive the odor, odor A has a 10% chance to be followed by a reward; while odor B is paired with a 90% chance of reward; meanwhile, the researcher probed neuron activities in the VTA. Record neuron analysis, dopamine identification experiment shows the dopamine fired when they received the odor, but after the reward was released, mice that received odor A released significantly higher dopamine compared to odor B [9]. The burst of dopamine is indicative of mice's ability to acknowledge the rarer reward.

Although Eshel's research primarily focuses on dopamine activity, it exhibits animals' ability to understand the association when there is a precursor to notify an upcoming reward. We see prediction error as a way for the animals to discriminate between the possibility of two different events, and an important precondition of associative learning using reward prediction error is the presence of signaling. Elaborate from that, since the procedure was reversed, no signal will remind the subject before the reward is given, preventing association between and the odor after the reward might as well be a random event for the subject. Based on the idea, a hypothesis can be made: when the procedure of associative learning is reversed, subjects will be confused by the relationship between reward and odor.

2. Materials and Methods Used for Preparation

Fifteen male adults Mus musculus (3 months) that were randomly chosen from a population with no medical history will be used in this experiment. All subjects receive a pre-lab preparation for 48 hours to disturb pre-formed associations between a stimulus and potential dopamine triggers by isolating in separate cells kept at 26°C, all subjects will be sustaining from food to prevent additional odors. (Nutritions will be injected to keep mice flourish) After the preparations, the fifteen mice will be split into two groups: (1) 5 mice will perform traditional associative learning as the control group, and (2) 10 mice will perform the reverse prediction error learning to study animals' learning ability associated with dopamine activity when the cue is absent.

The experiment consists of two sessions: training and testing. Subjects will be placed in new environments that repeat certain events over time in phase 1 and be tested on their memorization in phase 2. The training session for both groups will take place in experiment chamber 1 (75 cm length, 45 cm width, and height), a cell built of frosted glass that contains a dropper, infusion tube on the ceiling, and an elevated glass platform on the bottom. The rest of the floor is covered with wood chips and sawdust. (Fig. 1)

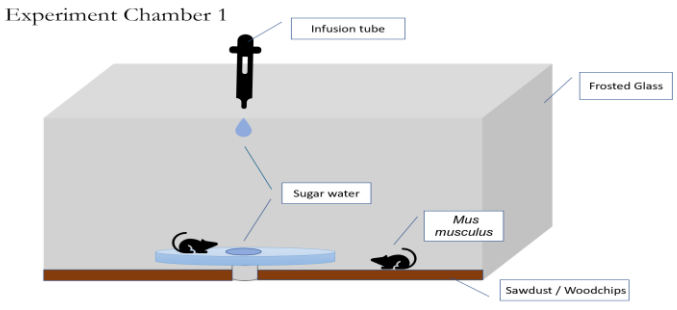

Fig. 1 Experiment chamber 1 provides an isolated environment to form expected associations. The primary equipment in this system - infusion tube and dropper are responsible to provide CS and rewards for the subjects. The remaining designs serve to prevent distractions: frosted glass dims the lights outside the chamber, sawdust and woodchips absorb excess odors and help when cleaning feces; furthermore, the combination of droppers and glass plate reduces redundant stimulus.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202110_7(10).0059

2.1 Traditional associative learning group

After the subjects (n = 5) are examined one last time for any physical injuries and mental stability, they will enter the experiment chamber. Every 10 minutes an odor is released with a chance followed by 5 mL sugar water (O_1 = 10%, O_2 = 90%). We hypothesize the subjects' dopamine burst would be triggered by the CS after a few rounds of training and dopamine level will leap when O_1 is followed by a reward. Eventually, the burst would be fully suppressed. As described above.

2.2 Reverse prediction error learning group

Similar to the traditional associative learning group, this group will receive sugar water every 10 minutes and may be followed by an odor (O_1 = 10%, O_2 = 90%). (Fig. 2) Although we can already infer the firing pattern in the conventional control group, we can only hypothesize this group behaves along with the definition and cannot establish the necessary connection, tetrode should receive negative feedback from the VTA region during CS and positive feedback after the reward.

Both groups' training will be repeated thirty times.

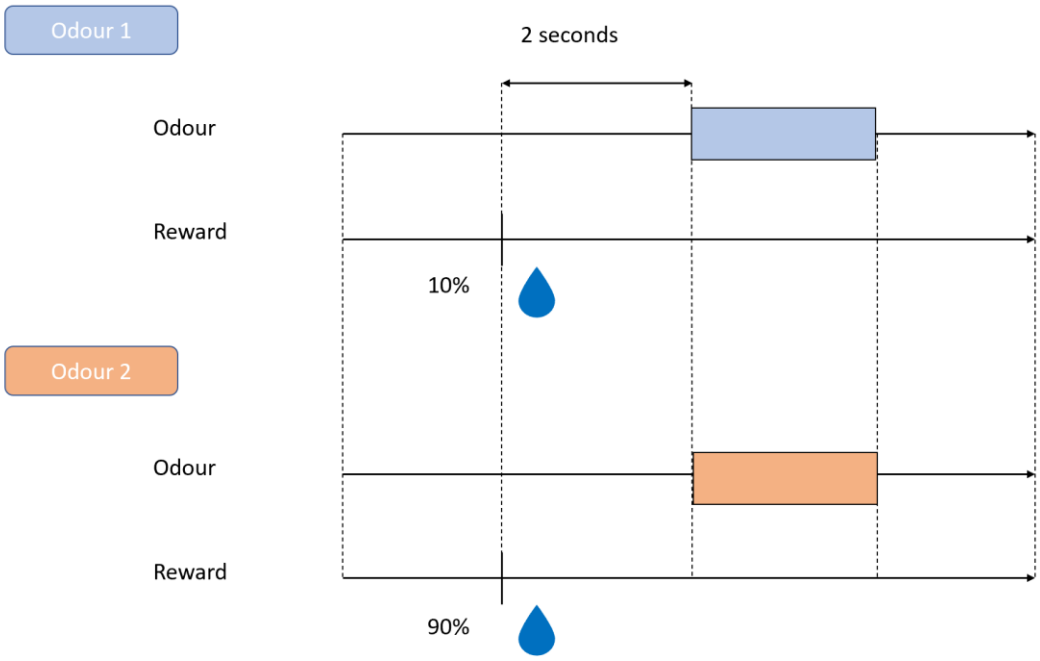

Fig. 2 How the two odors behave during a single trial of reversed practice. After a reward is released, there is a 10% chance odor 1 would be released, 90% chance no odor would be released. Similar for odor 2, but the reward is a lot more likely to be followed by the odor, and only a 10% chance to not receive the CS. A 2 seconds gap between the reward and CS allows subjects' responses toward the reward to be recorded.

3. Material and Methods Used to Test the Hypothesis

Until this step, we have completed all the preparations and training. The result examination takes place in a separate testing chamber 2 (Fig. 3). Before subjects enter the chamber, they should rest for an hour to avoid muscle memories. Later, individual mice would be released consecutively into the chamber and decide between the two odors given earlier during preparation in the center circular lobby. Trained subjects would enter the chamber through the entrance and encounter a divided pathway. The right passage leads to a room with odor 1 and the left leads to odor 2; meanwhile, subjects should be able to distinguish the passage each odor is coordinated with.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202110_7(10).0059

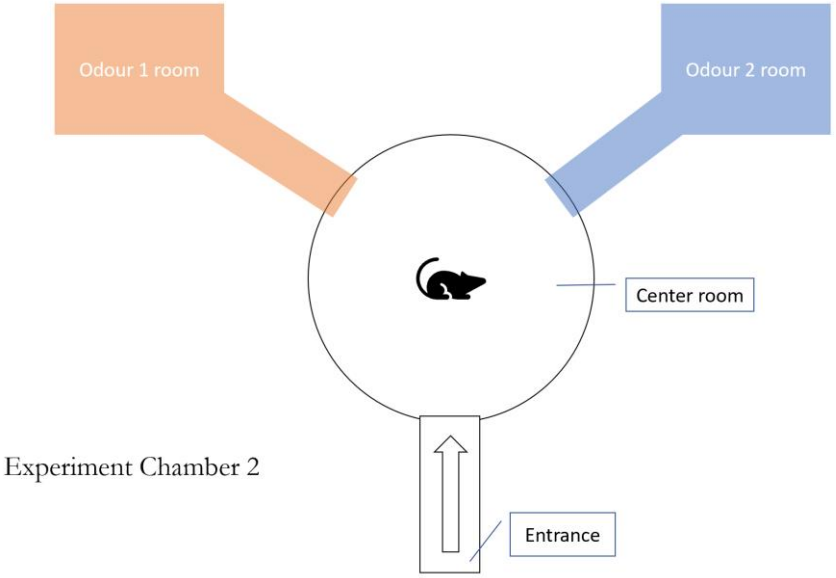

Fig. 3

Fig. 3 The entrance (35 cm length, 15 cm width, 10 cm height) can be directly attached to the individual cells where subjects rest. The entrance leads to a circular lobby (30 cm diameter, 10 cm height) that connects to two scented rooms. The left path and room contain odor 1 and the right path and room contain odor 2. A pipe is attached to the top of the lobby to pump out odors and maintain convection; furthermore, the center room's floor is covered with a layer of wood chips and sawdust to absorb mingled scents.

Based on the previous deduction, we can reason that the destination of the conventional associative group is likely the room on the right with odor 2. Since most subjects should be fully trained to establish an association between odor 2 and rewards. However, a reversed learning group can generate several outcomes each corresponding to a different hypothesis.

3.1 Most subjects ended in the left room

Most subjects decided to head for a room that's not intended to associate with a larger chance of reward, instead, they have chosen a room with an odor that stimulates more dopamine when a rare reward shows up. This phenomenon could be caused by the subject's addiction to dopamine secretion, and lack of interest in the reward. This result shows animals will voluntarily give up material benefits for neural excitement. The outcomes may vary upon different subjects' species.

3.2 Most subjects ended in the right room

Most subjects are capable of associating a CS after the reward with the reward. This shows prediction error is not the only method of establishing associative learning and animals are capable of associating two independent events; however, we are uncertain about the strength of the memory created.

3.3 The subjects were randomly distributed among the two rooms

This outcome will strengthen the definition of current prediction error associative learning theory; most subjects didn't associate the CS with the reward when a cue is absent.

4. Conclusion

The reversed associative learning experiment provides critical information for the current associative learning theory. An unexpected reward followed with a CS is a unique method to examine the flexibility of animals' learning ability; furthermore, the dopamine analysis during stimuli, rewards, and testing process will help us comprehend their decisions driven by neuron movements. Combining data from the macroscopic and microscopic levels will provide a glance at the impact of associative

ISSN: 2414-1895 DOI: 10.6919/ICJE.202110_7(10).0059

learning; through repeated experiments, the study may even yield a different, perhaps a better, learning method. In ideal conditions, we will be able to manipulate the whole associative learning process to expand human's visual, aural, and tactual memories, and accelerate the learning process.

Many more questions were raised from this experiment, despite the obstacles caused by lack of resources, the experiment could also be carried out using different CS, different drugs, and subjects.

References

[1] Handler, Annie et al. "Distinct Dopamine Receptor Pathways Underlie the Temporal Sensitivity of Associative Learning." Cell vol. 178,1 (2019): 60-75.e19. doi:10.1016/j.cell.2019.05.040.

[2] Rehman, Ibraheem, et al. "Classical Conditioning." StatPearls, StatPearls Publishing, 27 August 2020.

[3] Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron. 2005 Jul 7;47(1):129-41. doi: 10.1016/j.neuron.2005.05.020. PMID: 15996553; PMCID: PMC1564381.

[4] Rescorla, Robert A., et al. "Classical conditioning II: current research and theory." (1972): 64-99.

[5] Schultz, W et al. "A neural substrate of prediction and reward." Science (New York, N.Y.) vol. 275,5306 (1997): 1593-9. doi:10.1126/science.275.5306.1593.

[6] Schultz, Wolfram. "Dopamine reward prediction error coding." Dialogues in clinical neuroscience vol. 18,1 (2016): 23-32. doi:10.31887/DCNS.2016.18.1/wschultz.

[7] Steinberg, Elizabeth E et al. "A causal link between prediction errors, dopamine neurons and learning." Nature neuroscience vol. 16,7 (2013): 966-73. doi:10.1038/nn.3413.

[8] Van Zessen, Ruud et al. "Activation of VTA GABA neurons disrupts reward consumption." Neuron vol. 73,6 (2012): 1184-94. doi:10.1016/j.neuron.2012.02.016.

[9] Eshel, Neir et al. "Arithmetic and local circuitry underlying dopamine prediction errors." Nature vol. 525,7568 (2015): 243-6. doi:10.1038/nature14855.

[10] Enomoto, Kazuki et al. "Dopamine neurons learn to encode the long-term value of multiple future rewards." Proceedings of the National Academy of Sciences of the United States of America vol. 108,37 (2011): 15462-7. doi:10.1073/pnas.1014457108.

[11] Schultz, W et al. "Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task." The Journal of neuroscience : the official journal of the Society for Neuroscience vol. 13,3 (1993): 900-13. doi:10.1523/JNEUROSCI.13-03-00900.1993.

[12] Flagel, Shelly B et al. "A selective role for dopamine in stimulus-reward learning." Nature vol. 469,7328 (2011): 53-7. doi:10.1038/nature09588.

[13] Carew, T J et al. "Classical conditioning in a simple withdrawal reflex in Aplysia californica." The Journal of neuroscience : the official journal of the Society for Neuroscience vol. 1,12 (1981): 1426-37. doi:10.1523/JNEUROSCI.01-12-01426.1981.

[14] Dudai, Y et al. "dunce, a mutant of Drosophila deficient in learning." Proceedings of the National Academy of Sciences of the United States of America vol. 73,5 (1976): 1684-8. doi:10.1073/ pnas.73.5. 1684.

[15] Mizunami, Makoto, et al. Invertebrate Learning and Memory: Chapter 41. Olfactory and Visual Learning in Cockroaches and Crickets. Vol. 22. Elsevier Inc. Chapters, 2013.

[16] Terao, Kanta et al. "Critical evidence for the prediction error theory in associative learning." Scientific reports vol. 5 8929. 10 Mar. 2015, doi:10.1038/srep08929.