Understanding Credit Assignment: a Two-Step Approach

**Introduction**

There are many steps between a choice and its outcome, making it difficult to find a true cause. This problem, called credit assignment, has been prevalent in the field of neuroeconomics. There are three main parts to being able to precisely assign credit: value, memory, and causal attribution.[1](https://www.zotero.org/google-docs/?5mRy1n) For a choice to be properly credited, it needs to have a value that is accurate to the expected reward[2](https://www.zotero.org/google-docs/?vqkC5w), a unique memory or representation for this value to be applied to[3](https://www.zotero.org/google-docs/?UBrPhU), and a way to pair the value and representation to create a choice-outcome contingency.[4](https://www.zotero.org/google-docs/?OMhh8e) The ventromedial prefrontal cortex (vmPFC) is integral to creating a neural basis for subjective value.[5](https://www.zotero.org/google-docs/?RfdLjZ) When viewing a stimulus or reward, it encodes a unique value signal.[5–7](https://www.zotero.org/google-docs/?jfZ22Q) This signal is relative to the expected value of the reward a choice can lead to, which is informed by understanding the probabilities of the environment.[8](https://www.zotero.org/google-docs/?1ciT9J) Since there is bias in this calculation, there can be substantial variance, but it is known that this signal consistently varies across individuals in certain conditions.[7](https://www.zotero.org/google-docs/?HrXoap) If the reward is presented immediately after the choice, the signal will be strongest, but if there is a delay in getting the reward, the signal will weaken.[7](https://www.zotero.org/google-docs/?B5vcLA) For this value to be useful in credit assignment, it must have a representation to be related to.

To have a representation of a stimulus related to a choice, our brain relies on the hippocampus (HC), a key area for memory.[9](https://www.zotero.org/google-docs/?AhPLDX) The HC creates a unique pattern of brain activity that can be decoded when viewing a particular stimulus and when a reward is given, which can be considered a representation of that stimulus.[10](https://www.zotero.org/google-docs/?9uomwy) There is a similar mechanism seen in memory recall, where the HC will replay a sequence of familiar activity to inform a specific decision.[9,11](https://www.zotero.org/google-docs/?Np0Wxq) This neural replay has been shown to improve decision-making performance.[11](https://www.zotero.org/google-docs/?TUFmpT) It is also considered an efficient method to store memory through generalizing sequences of information to apply to similarly structured assessments.[9](https://www.zotero.org/google-docs/?mYtLS6) With a representation of a choice and various values, they have to be paired for it to be properly credited.

On the other hand, the medial prefrontal cortex (mPFC) creates the contingency between value and choice.[4](https://www.zotero.org/google-docs/?qrEB7s) When a reward is presented, it creates a choice-outcome contingency with the choice that led to the reward and assigns the value of the reward to that choice.[4](https://www.zotero.org/google-docs/?gV0ZTi) In decision-making tasks, rewards are presented many times throughout the course of each session, so these contingencies have to remain flexible to change with new information.[3](https://www.zotero.org/google-docs/?nXZ7Nd) Every time new information is presented, like a different reward, the difference in value will be incorporated into an updated choice-outcome contingency.[3](https://www.zotero.org/google-docs/?qTKbWI) The mPFC uses the same mechanism for unchosen options, using an inferred value to create the contingency with the alternative stimulus.[3](https://www.zotero.org/google-docs/?v4zrVM) The combined action of the vmPFC, HC, and mPFC provide a basic mechanism for precise credit assignment. There are clear representations of value, the various stimuli, and a mechanism by which these can be related and updated.

The difficulty in applying these findings to real-life situations comes from the simplicity of the tasks the results come from. Many tasks use different variants of a one-step task structure, where the subject chooses between one pair of stimuli and then a reward is presented.[3,12](https://www.zotero.org/google-docs/?KWeiGh) This has been extremely useful for understanding the basic mechanisms behind credit assignment, but without additional steps between a choice and a reward, it lacks the depth and complexity of real-life decisions. The goal of my research is to increase the complexity of these tasks to determine how these neural mechanisms work when there are two steps before a reward is presented. Although two-step decision-making is still far from what we experience in life, it’s a crucial step toward a deeper understanding. Since neural replay of sequences of information has been seen in memory recall[9](https://www.zotero.org/google-docs/?XQxQaI), I hypothesize that the brain treats the decision-making process as a chain of dependent events that lead to a reward, instead of a series of independent decisions. Given that the brain optimizes for efficient information storage[9](https://www.zotero.org/google-docs/?ZNy6Rs), it would be reasonable for the sequence of decisions to be used for a representation that led to a reward, which then will have its own value and choice-outcome contingency. Such a pattern would yield precise credit assignment for a series of dependent decisions.

**Methods**

I will recruit 30-35 participants with no history of addiction through the SONA recruitment pool. They will be UC Davis undergraduate students, aged 18-23. They will undergo training at the Center for the Mind and Brain, where they will be instructed on how to properly complete the task. If they perform adequately, they will be invited to complete the same task while inside an fMRI machine for two separate sessions. They will be compensated $100 for completion of all three sessions.

The task will involve a two-step decision-making sequence before reaching a reward. The first stage will involve an option between two distinct doors, one red and one green. Each door will lead to a unique room, which will contain another pair of doors to choose from. Each door in each room will have a unique color to keep them distinguishable. After the door in the room is chosen, a reward ranging from 1-4 will be displayed. The goal for the subject is to collect as many high (4) rewards as possible. To maintain flexibility in their sequences of decisions, each door in the first stage has an 80% chance of leading to the proper room and a 20% chance of leading to the alternate room. The location of each reward will also change randomly throughout the session. Each session will contain 180 trials, leading to 180 rewards, which will be summed up and displayed at the end of the session.

After preprocessing the imaging data using fmriprep, we will be measuring their blood-oxygen-level-dependent (BOLD) signal in the vmPFC, HC, and mPFC. For the behavioral data, we will measure the sum of rewards across the 180 trials per session. To test my hypothesis, I will be using t-tests to compare the BOLD signal against the resting-state brain activity. I will also ensure the participants are completing the task properly by using a t-test to compare each subject's sum of scores to a model that made each decision randomly.

**Expected Results**

Using this task, I will find how the strength of the value signal in the vmPFC changes, how stimuli are represented at feedback in the HC, and how the mPFC creates accurate choice-outcome contingencies. I expect the following: the HC will reinstate the entire sequence of dependent decisions at feedback, the mPFC will use this sequence to create choice-outcome contingencies, and the strength of the value signal in the vmPFC will increase as the participant gets closer to the reward. As for the behavioral results, I expect each participant’s score to be significantly greater than the model’s score.

**Intellectual Merit**

By uncovering the mechanisms of two-step credit assignment, I am adding a layer of complexity that provides promising insight into how our brain treats increasingly complex decisions. With the help of neuroimaging and my experience in this field, there is potential to find a connection between our memory mechanisms and decision-making mechanisms. Additionally, by understanding these processes in healthy subjects, I have the potential to contribute to an understanding of disruptions in the decision-making process, such as in substance addiction.

**Broader Impacts**

I will create a podcast series that explains neuroeconomics in a simple, digestible manner. Within each episode, I will have a central topic, such as credit assignment, where I will break down the current landscape of what we know and how we know it. To increase public interaction, there will be a Q&A section at the end, where I will answer questions from the listeners. Science can feel very distant and unapproachable to the general public, so by providing accessible information and encouraging participation, I will help bridge this gap.

Additionally, I have reached out to the Drug Abuse Resistance Education (DARE) association and have scheduled to visit a group of middle schools in a local underprivileged area. During these visits, I will work with the DARE team to educate the at-risk youth on how we make decisions and how drugs can negatively affect these processes. In providing this education, I will be empowering the youth to make informed decisions, exposing them to rigorous science at a young age, and demonstrating how science can directly help their lives. This can lead to a group of young students who feel empowered by science, instead of ostracized. I will also provide a workshop to the DARE workers on this topic to help them further develop their education efforts and create a long-lasting impact in their future presentations.

[1. Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574. 10.1146/annurev.neuro.29.051605.113038.](https://www.zotero.org/google-docs/?DNHs0F)

[2. Lee, D., Seo, H., and Jung, M.W. (2012). Neural Basis of Reinforcement Learning and Decision Making. Annu. Rev. Neurosci. *35*, 287–308. 10.1146/annurev-neuro-062111-150512.](https://www.zotero.org/google-docs/?DNHs0F)

[3. Witkowski, P.P., Park, S.A., and Boorman, E.D. (2022). Neural mechanisms of credit assignment for inferred relationships in a structured world. Neuron *110*, 2680-2690.e9. 10.1016/j.neuron.2022.05.021.](https://www.zotero.org/google-docs/?DNHs0F)

[4. Broche-Pérez, Y., Herrera Jiménez, L.F., and Omar-Martínez, E. (2016). Neural substrates of decision-making. Neurol. Engl. Ed. *31*, 319–325. 10.1016/j.nrleng.2015.03.009.](https://www.zotero.org/google-docs/?DNHs0F)

[5. Boorman, E.D., Behrens, T.E.J., Woolrich, M.W., and Rushworth, M.F.S. (2009). How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action. Neuron *62*, 733–743. 10.1016/j.neuron.2009.05.014.](https://www.zotero.org/google-docs/?DNHs0F)

[6. Bartra, O., McGuire, J.T., and Kable, J.W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. NeuroImage *76*, 412–427. 10.1016/j.neuroimage.2013.02.063.](https://www.zotero.org/google-docs/?DNHs0F)

[7. Rushworth, M.F.S., Noonan, M.P., Boorman, E.D., Walton, M.E., and Behrens, T.E. (2011). Frontal Cortex and Reward-Guided Learning and Decision-Making. Neuron *70*, 1054–1069. 10.1016/j.neuron.2011.05.014.](https://www.zotero.org/google-docs/?DNHs0F)

[8. Platt, M.L., and Huettel, S.A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. Nat. Neurosci. *11*, 398–403. 10.1038/nn2062.](https://www.zotero.org/google-docs/?DNHs0F)

[9. Huang, Q., and Luo, H. (2023). Shared structure facilitates working memory of multiple sequences via neural replay. Preprint at bioRxiv, 10.1101/2023.07.18.549616 10.1101/2023.07.18.549616.](https://www.zotero.org/google-docs/?DNHs0F)

[10. Mack, Michael.L., and Preston, Alison.R. (2016). Decisions about the past are guided by reinstatement of specific memories in the hippocampus and perirhinal cortex. NeuroImage *127*, 144–157. 10.1016/j.neuroimage.2015.12.015.](https://www.zotero.org/google-docs/?DNHs0F)

[11. Creswell, J.D., Bursley, J.K., and Satpute, A.B. (2013). Neural reactivation links unconscious thought to decision-making performance. Soc. Cogn. Affect. Neurosci. *8*, 863–869. 10.1093/scan/nst004.](https://www.zotero.org/google-docs/?DNHs0F)

[12. Boorman, E.D., Witkowski, P.P., Zhang, Y., and Park, S.A. (2021). The orbital frontal cortex, task structure, and inference. Behav. Neurosci. *135*, 291–300. 10.1037/bne0000465.](https://www.zotero.org/google-docs/?DNHs0F)