



# Neural Mechanisms Underlying Expectation-Guided Decision-Making

Thorsten Kahnt\*

National Institute on Drug Abuse Intramural Research Program, National Institutes of Health, Baltimore, MD, United States

**Keywords:** decision-making, expectation, outcome-guided, inference, orbitofrontal cortex, model-based

## INTRODUCTION

Many decisions are guided by expectations about their outcomes. For instance, we may decide to visit a restaurant because we anticipate the food to be outstanding. How these expectations are represented in the brain, and how they allow us to make adaptive choices are important questions for understanding the neural basis of behavior.

Work across species has revealed brain areas that signal expected rewards (Haber and Knutson, 2010; Kahnt, 2018). This work typically focuses on neural correlates of the value of choice options (Padoa-Schioppa, 2011), that is, how desirable an option is. Activity in many brain areas, including the striatum, ventromedial prefrontal cortex and orbitofrontal cortex (OFC), is correlated with expected value. However, expected outcomes are more than their value—they have a specific identity. Even though we may equally desire pizza funghi and spaghetti arrabiata, they are not the same, and representing expectations about the identity of outcomes is important for adaptive decision-making.

In this opinion, I will summarize recent work from my lab that has shown how the lateral OFC represents expectations about specific outcomes, how these expectations are learned, and how they can be used for adaptive decision-making. Finally, I will summarize evidence that disrupting activity in OFC networks that represent specific outcome expectations impairs adaptive behavior. Together, these findings support the view that the OFC contributes to expectation-guided decision-making by enabling us to simulate the consequences of our choices.

## NEURAL REPRESENTATIONS OF OUTCOME EXPECTATIONS

Recent studies have shown that the OFC represents not only expectations about the value of future outcomes but also their identity (Howard and Kahnt, 2021). For instance, in one study, we used food odors as specific rewards and selected one sweet and one savory odor for each subject that were matched in rated pleasantness (i.e., value) (Howard et al., 2015). We then lowered the concentration of the food odors to create a set of low-intensity odors, which were rated as less pleasant than the high-intensity odors. The four food odors were then paired with different visual stimuli, such that each odor was reliably predicted by a different symbol. Finally, subjects were presented with these symbols while undergoing functional magnetic resonance imaging (fMRI). Multi-voxel pattern analysis (Kahnt, 2018) to the fMRI responses evoked by the symbols revealed that activity patterns in the lateral OFC, anterior cingulate cortex, and hippocampus differentiated between the two expected food odors, whereas activity patterns in the medial OFC represented the value of the odors, independent of their identity. These findings are in line with other work from our lab (Howard and Kahnt, 2017) as well as with studies showing that activity patterns in the lateral OFC represent values that are tied to specific reward categories, whereas activity in the medial OFC is independent of reward category (McNamee et al., 2013).

## OPEN ACCESS

### Edited by:

Rüdiger J. Seitz,  
Heinrich Heine University of  
Düsseldorf, Germany

### Reviewed by:

Serge H. Ahmed,  
Centre National de la Recherche  
Scientifique (CNRS), France

### \*Correspondence:

Thorsten Kahnt  
thorsten.kahnt@nih.gov

### Specialty section:

This article was submitted to  
Emotion Regulation and Processing,  
a section of the journal  
Frontiers in Behavioral Neuroscience

**Received:** 13 May 2022

**Accepted:** 14 June 2022

**Published:** 01 July 2022

### Citation:

Kahnt T (2022) Neural Mechanisms  
Underlying Expectation-Guided  
Decision-Making.  
*Front. Behav. Neurosci.* 16:943419.  
doi: 10.3389/fnbeh.2022.943419

## LEARNING OF OUTCOME EXPECTATIONS

Outcome expectations are based on associations between predictive stimuli and rewards, and these associations need to be learned and updated through experience. Work in non-human primates has shown that dopamine neurons in the midbrain contribute to learning the value of rewards by signaling reward prediction errors, or the difference between received and expected rewards (Schultz et al., 1997). We hypothesized that midbrain activity encodes a similar signal for identity prediction errors, which may be used for learning reward identity expectations.

In one experiment, hungry subjects were presented with visual symbols that predicted one of two preference-matched food odors (e.g., strawberry or potato chips) in either low or high intensity (Howard and Kahnt, 2018). As in previous studies, subjects reported a higher preference for the high-intensity odors, but there was no preference difference between the sweet and savory food odors. After a number of trials of receiving the predicted odor, either the identity (e.g., subjects expected strawberry but received equally-preferred potato chips) or the intensity (e.g., subjects expected potato chips in low intensity but received the preferred high-intensity odor) of the odor was unexpectedly changed. fMRI activity in the midbrain showed signatures of value-based prediction errors, increasing when subjects received the more preferred high-intensity odor after expecting the less preferred low-intensity odor. However, activity in the same midbrain region also increased when subjects received strawberry after expecting potato chips, in line with the signaling of value-neutral identity prediction errors. Importantly, value- and identity-based prediction errors were found in the same part of the midbrain and were correlated, suggesting that they may originate from the same neural population. Similar findings have been observed in a study that recorded activity from dopamine neurons in rats (Takahashi et al., 2017), as well as in other human imaging studies (Boorman et al., 2016; Schwartenbeck et al., 2016; Suarez et al., 2019).

A question that follows is whether midbrain identity prediction errors actively shape identity learning in downstream areas, or whether they merely act as a permissive gating (i.e., salience) signal to direct attention and boost learning (Bromberg-Martin et al., 2010). We addressed this question, reasoning that if identity prediction errors conveyed salience information without providing specific information, there should be no difference between the midbrain response to reward B when A was predicted and the midbrain response to reward A when B was predicted. In contrast, if identity prediction errors actively shape learning in downstream targets, they should contain specific information such that midbrain responses differ between these two cases. In line with the latter idea, we found that midbrain fMRI patterns in humans and dopamine ensemble responses in rats contain information about the specific identity of the error (Stalnaker et al., 2019), suggesting they could directly update identity expectations in downstream areas, such as OFC.

Indeed, we found that the magnitude of identity prediction error response in the midbrain was correlated with how much identity expectations in the lateral OFC changed after an

identity error (Howard and Kahnt, 2018). This suggests that identity expectations in the lateral OFC are updated through a mechanism that involves identity prediction errors in the dopaminergic midbrain.

## USING EXPECTATIONS FOR INFERENCE

In many cases, we can learn the expected value of choice options through direct experience. For instance, we can learn the value of an item on a restaurant menu by ordering it. However, for many other decisions in life, we simply have not had the opportunity to directly learn values in this way. This especially applies to decisions that are less frequently or only indirectly experienced, like deciding to try out a new restaurant or whether to visit a new country. Also, the values we have learned from previous decisions may have changed since we last made that choice, and using these old values would lead to maladaptive decisions. In these situations, value expectations need to be computed by mentally simulating or inferring the value of the option based on incomplete information. Specific outcome expectations allow us to do this because they are part of a model of the relevant environment which we can use to simulate the consequences of our actions.

Such simulations can be studied in the devaluation task. In a typical experiment, subjects first learn to associate different sensory cues with different foods, e.g., M&Ms and peanuts (Rudebeck et al., 2013; Murray et al., 2015; Reber et al., 2017). After one of the rewards is devalued by feeding the food to satiety, subjects can make choices between the sensory cues. To access the current value of the choice option, subjects must simulate what outcome they will receive by making a particular choice and infer its current value. This allows them to avoid selecting the cue that predicts the devalued outcome. In contrast, if they use the previously learned value, they will make choices that result in both the valued and the devalued outcome.

We have used transcranial magnetic stimulation (TMS) to test whether outcome identity expectations represented in the lateral OFC are necessary for adaptive responding in the devaluation task (Howard et al., 2020). Hungry participants first learned associations between visual symbols and sweet or savory food odors and were then allowed to make choices between these symbols. Stimulation coordinates in the lateral PFC were selected for each participant based on resting-state fMRI connectivity with lateral OFC. After a session of continuous theta burst stimulation (cTBS), which has inhibitory after-effects lasting for 50–60 min (Huang et al., 2005), or sham stimulation, subjects ate a meal that was matched to either the sweet or the savory food odor. After this devaluation procedure, subjects could again make choices between the cues. Targeting the lateral OFC with cTBS had profound effects on subjects' choices after the meal. Whereas, subjects in the sham group adaptively stopped selecting symbols that predicted the devalued odor, subjects in the cTBS group continued to select these stimuli. This shows that OFC activity is required for using specific outcome expectations for making inferences about the current value of choice options.

A different type of inference can be probed in the sensory preconditioning task (Brogden, 1939; Hoffeld et al., 1960). In this task, subjects first learn associations between sensory stimuli A and B, and C and D ( $A \rightarrow B$ ,  $C \rightarrow D$ ). Then, the second cue of each pair (B and D) is paired with either a reward or no reward ( $B \rightarrow$  reward,  $D \rightarrow$  no reward). Finally, responses to all stimuli (A, B, C, and D) are probed. Humans and other animals show stronger responding to stimulus A compared to stimulus C in this final test (Sadacca et al., 2016; Sharpe et al., 2017; Wang et al., 2020b). This pattern of responding is compatible with the idea that subjects mentally step through the associations  $A \rightarrow B$  and  $B \rightarrow$  reward to infer that  $A \rightarrow$  reward.

Activity in the OFC correlates with learning of the stimulus-stimulus associations during the initial learning phase (Sadacca et al., 2018; Wang et al., 2020b), suggesting that the OFC represents the associative structure of the task. In other words, stimulus-stimulus associations appear to be represented in the same way as associations between a sensory stimulus and a food reward. Moreover, OFC is critical for using these associations to perform mental simulations. Pharmacological inactivation of the lateral OFC in rats (Jones et al., 2012) as well as cTBS targeting the lateral OFC network in humans before the final phase of the sensory preconditioning task impairs responding to cue A, without affecting responding to cue B (for which subjects had directly learned the stimulus-outcome associations) (Wang et al., 2020a). Thus, just like neural representations of specific outcome expectations, representations of stimulus-stimulus associations in the lateral OFC network are critical for making mental simulations required for adaptive decision-making.

## DISCUSSION

The work described above outlines the neural mechanisms underlying expectation-guided decision-making. In brief, the OFC represents expectations about specific outcomes, and these expectations are learned through an error-based mechanism that involves the dopaminergic midbrain. The same networks that represent outcome expectations also represent expectations about future events, even if they do not possess any value. Of note, while we often make decisions between options with outcomes that belong to very different categories, our experiments used

outcomes from the same reward category (i.e., food). This can be considered a stronger test of the outcome-specific coding hypothesis, because differences in neural responses to different reward categories may not only reflect outcome-specific coding but also different preparatory or consummatory reward responses. Thus, results from within category experiments are likely to generalize to across category settings. Indeed, previous work on neural representations of different reward categories has revealed comparable findings (Levy and Glimcher, 2011; McNamee et al., 2013; Gross et al., 2014).

Neural representations of specific outcomes enable us to perform mental simulations that are required for adaptive behavior in novel situations or when the value of an outcome has changed since we last made that decision. In other words, these representations allow us to flexibly assign value or meaning to expected outcomes in order to guide our decisions. Together, the findings discussed here are compatible with the view that the OFC network contributes to decision-making by representing a model of the environment, which enables us to make flexible inferences about the outcomes of our decisions.

## AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

## FUNDING

This study was funded by Dr. Rüdiger Seitz, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. The author was supported by the Intramural Research Program at the National Institute on Drug Abuse, and the expressed opinions are the author's own and do not reflect the view of the NIH/DHHS. Siemens Healthineers was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

## ACKNOWLEDGMENTS

The author thanks the members of the Kahnt Lab who have performed the work summarized here.

## REFERENCES

- Boorman, E. D., Rajendran, V. G., O'Reilly, J. X., and Behrens, T. E. (2016). Two anatomically and computationally distinct learning signals predict changes to stimulus-outcome associations in hippocampus. *Neuron* 89, 1343–1354. doi: 10.1016/j.neuron.2016.02.014
- Brogden, W. J. (1939). Sensory pre-conditioning. *J. Exp. Psychol.* 25, 323–332. doi: 10.1037/h0058944
- Bromberg-Martin, E. S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68, 815–834. doi: 10.1016/j.neuron.2010.11.022
- Gross, J., Woelbert, E., Zimmermann, J., Okamoto-Barth, S., Riedel, A., and Goebel, R. (2014). Value signals in the prefrontal cortex predict individual preferences across reward categories. *J. Neurosci.* 34, 7580–7586. doi: 10.1523/JNEUROSCI.5082-13.2014
- Haber, S. N., and Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35, 4–26. doi: 10.1038/npp.2009.129
- Hoffeld, D. R., Kendall, S. B., Thompson, R. F., and Brogden, W. J. (1960). Effect of amount of preconditioning training upon the magnitude of sensory preconditioning. *J. Exp. Psychol.* 59, 198–204. doi: 10.1037/h0048857
- Howard, J. D., Gottfried, J. A., Tobler, P. N., and Kahnt, T. (2015). Identity-specific coding of future rewards in the human orbitofrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5195–5200. doi: 10.1073/pnas.1503550112
- Howard, J. D., and Kahnt, T. (2017). Identity-Specific reward representations in orbitofrontal cortex are modulated by selective devaluation. *J. Neurosci.* 37, 2627–2638. doi: 10.1523/JNEUROSCI.3473-16.2017
- Howard, J. D., and Kahnt, T. (2018). Identity prediction errors in the human midbrain update reward-identity expectations in the orbitofrontal cortex. *Nat. Commun.* 9, 1611. doi: 10.1038/s41467-018-04055-5

- Howard, J. D., and Kahnt, T. (2021). To be specific: the role of orbitofrontal cortex in signaling reward identity. *Behav. Neurosci.* 135, 210–217. doi: 10.1037/bne0000455
- Howard, J. D., Reynolds, R., Smith, D. E., Voss, J. L., Schoenbaum, G., and Kahnt, T. (2020). Targeted stimulation of human orbitofrontal networks disrupts outcome-guided behavior. *Curr. Biol.* 30, 490–498.e4. doi: 10.1016/j.cub.2019.12.007
- Huang, Y. Z., Edwards, M. J., Rounis, E., Bhatia, K. P., and Rothwell, J. C. (2005). Theta burst stimulation of the human motor cortex. *Neuron* 45, 201–206. doi: 10.1016/j.neuron.2004.12.033
- Jones, J. L., Esber, G. R., Mcdannald, M. A., Gruber, A. J., Hernandez, A., Mireni, A., et al. (2012). Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science* 338, 953–956. doi: 10.1126/science.1227489
- Kahnt, T. (2018). A decade of decoding reward-related fMRI signals and where we go from here. *Neuroimage* 180, 324–333. doi: 10.1016/j.neuroimage.2017.03.067
- Levy, D. J., and Glimcher, P. W. (2011). Comparing apples oranges: using reward-specific reward-general subjective value representation in the brain. *J. Neurosci.* 31, 14693–14707. doi: 10.1523/JNEUROSCI.2218-11.2011
- McNamee, D., Rangel, A., and O’Doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nat. Neurosci.* 16, 479–485. doi: 10.1038/nn.3337
- Murray, E. A., Moylan, E. J., Saleem, K. S., Basile, B. M., and Turchi, J. (2015). Specialized areas for value updating and goal selection in the primate orbitofrontal cortex. *Elife* 4, e009. doi: 10.7554/eLife.11695.009
- Padoa-Schioppa, C. (2011). Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* 34, 333–359. doi: 10.1146/annurev-neuro-061010-113648
- Reber, J., Feinseth, J. S., O’Doherty, J. P., Liljeholm, M., Adolphs, R., and Tranel, D. (2017). Selective impairment of goal-directed decision-making following lesions to the human ventromedial prefrontal cortex. *Brain* 140, 1743–1756. doi: 10.1093/brain/awx105
- Rudebeck, P. H., Saunders, R. C., Prescott, A. T., Chau, L. S., and Murray, E. A. (2013). Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nat Neurosci* 16, 1140–1145. doi: 10.1038/nn.3440
- Sadacca, B. F., Jones, J. L., and Schoenbaum, G. (2016). Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. *Elife* 5, e010. doi: 10.7554/eLife.13665.010
- Sadacca, B. F., Wied, H. M., Lopatina, N., Saini, G. K., Nemirovsky, D., and Schoenbaum, G. (2018). Orbitofrontal neurons signal sensory associations underlying model-based inference in a sensory preconditioning task. *Elife* 7, e30373. doi: 10.7554/eLife.30373
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593
- Schwartenbeck, P., Fitzgerald, T. H. B., and Dolan, R. (2016). Neural signals encoding shifts in beliefs. *Neuroimage* 125, 578–586. doi: 10.1016/j.neuroimage.2015.10.067
- Sharpe, M. J., Batchelor, H. M., and Schoenbaum, G. (2017). Preconditioned cues have no value. *Elife* 6, e006. doi: 10.7554/eLife.28362.006
- Stalnaker, T. A., Howard, J. D., Takahashi, Y. K., Gershman, S. J., Kahnt, T., and Schoenbaum, G. (2019). Dopamine neuron ensembles signal the content of sensory prediction errors. *Elife* 8, e010. doi: 10.7554/eLife.49315.010
- Suarez, J. A., Howard, J. D., Schoenbaum, G., and Kahnt, T. (2019). Sensory prediction errors in the human midbrain signal identity violations independent of perceptual distance. *Elife* 8, e032. doi: 10.7554/eLife.43962.032
- Takahashi, Y. K., Batchelor, H. M., Liu, B., Khanna, A., Morales, M., and Schoenbaum, G. (2017). Dopamine neurons respond to errors in the prediction of sensory features of expected rewards. *Neuron* 95, 1395–1405.e3. doi: 10.1016/j.neuron.2017.08.025
- Wang, F., Howard, J. D., Voss, J. L., Schoenbaum, G., and Kahnt, T. (2020a). Targeted stimulation of an orbitofrontal network disrupts decisions based on inferred, not experienced outcomes. *J. Neurosci.* 40, 8726–8733. doi: 10.1523/JNEUROSCI.1680-20.2020
- Wang, F., Schoenbaum, G., and Kahnt, T. (2020b). Interactions between human orbitofrontal cortex and hippocampus support model-based inference. *PLoS Biol.* 18, e3000578. doi: 10.1371/journal.pbio.3000578

**Conflict of Interest:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Kahnt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.