



# HHS Public Access

Author manuscript

*Curr Opin Behav Sci.* Author manuscript; available in PMC 2022 April 01.

Published in final edited form as:

*Curr Opin Behav Sci.* 2021 April ; 38: 1–7. doi:10.1016/j.cobeha.2020.06.003.

## Latent learning, cognitive maps, and curiosity

Maya Zhe Wang, Benjamin Y. Hayden

Department of Neuroscience, Center for Magnetic Resonance Research, and Center for Neuroengineering University of Minnesota, Minneapolis MN 55455

### Abstract

Curiosity is a desire for information that is not motivated by strategic concerns. Latent learning is not driven by standard reinforcement processes. We propose that curiosity serves the purpose of motivating latent learning. While latent learning is often treated as a passive or incidental process, it normally reflects a strong evolved pressure to actively seek large amounts of information. That information in turn allows curious decision makers to represent the structure of their environment, that is, to form cognitive maps. These cognitive maps then drive adaptive flexible behavior. Based on recent data, we propose that orbitofrontal cortex (OFC) and dorsal anterior cingulate cortex (dACC) play complementary roles in curiosity-driven learning. Specifically, we propose that (1) OFC tracks intrinsic value of information and incorporates new information into a cognitive map; and (2) dACC tracks the environmental demands and information availability to then use the cognitive map from OFC to guide behavior.

### Keywords

curiosity; latent learning; cognitive maps; information-seeking; humanlike curiosity

### Main Text

The natural environment offers a plethora of rewards to most foragers but acquiring these rewards requires knowledge [1,2,3,4]. For example, red knots (arctic shorebirds), feed on bivalves that are patchily distributed and buried in the mud [5]. Notably, the locations of these prey cannot be guessed based on visual inspection, but can be inferred based on a rich knowledge of likely patch structure and distribution of other prey. When foraging, the birds reside in patches longer than predicted by simple foraging models; their overstaying can be explained by modified models that include a bonus for the information that the extra

---

**Corresponding author:** Maya Zhe Wang, Department of Neuroscience and Center for Magnetic Resonance Research University of Minnesota, Minneapolis MN 55455 mayawangz@gmail.com.

Competing interests

The authors have no competing interests to declare.

COI

The authors declare no conflicts of interest

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

residence time provides. Knots are typical of many natural decision makers, which are constantly starved for information. Notably, this is an area where laboratory experiments tend to differ most starkly from natural decision-making contexts; in the lab relevant information is typically made available and, if obscured, simple.

Curiosity, which we can define as a drive for non-strategic information, is a major driver of learning and determinative of the success of development in humans and other animals [6,7,8,9,10,11]. Its features are preserved across species and over the lifespan. It appears to be associated with at least somewhat discrete neural circuits [12,13,14].

### Latent learning

Classical concepts of learning held that all learning is driven by reinforcement contingencies. These ideas are fundamental to the “Law of Effect” and to Hebbian learning [15,16]. That work, in elaborated form, is central to reinforcement learning, one of the most successful psychological theories and the basis of a generation of systems neuroscience, and to much of machine learning.

However robust learning can occur in the absence of reward [17,18]. This idea poses a challenge to simple reinforcement learning models, which Tolman termed the “*stimulus response school*.” In a classic latent learning setup, a rat is released into a large maze with no reward. Naive rats typically amble around the maze, ostensibly with no purpose. Later, the experimenters introduce a reward to a specific location in the maze. The rats with maze exposure learned to locate the reward much more quickly than ones who were naive to that maze. The rats learned the maze structure – and formed a cognitive map – *latently*.

### Curiosity and cognitive maps

Any forager placed within a complex natural environment must naturally trade off between the costs and benefits of exploration. In addition to the metabolic costs of locomotion, sensory processing, and learning, active exploration carries opportunity costs: that time could be better spent searching for food, courting and reproducing, or avoiding predators. For example, in the case of the knot, the delay in travel time imposes an opportunity cost in the form of foregone large rewards at new patches. Even motivational processes driven by distal reward seeking must necessarily discount future rewards and uncertain rewards, and the benefits of exploration are unavoidably delayed beyond the temporal horizon and, individually, infinitesimally unlikely. So reward-maximizing calculation is unlikely to sufficiently motivate search. Instead, evolution must endow the decision maker with intrinsic motivation to learn and ultimately to map its environment [8,19,20,21].

Indeed, curiosity would seem to go hand in hand with the learning of cognitive maps. Cognitive maps refer to detailed mental representations of the relationship between various elements in the world and their sequelae [22,23,24]. Having a cognitive map allows a decision maker to not just guess what will happen but also to deal with unexpected changes in our environment. The classic idea about cognitive maps - also attributable to Tolman - is that they allow us to respond flexibly when contingencies change (e.g. when the layout of a maze changes, [18,25]). That kind of flexibility is very difficult to implement with basic

reinforcement learning processes [18,22,23,24,25]. Instead, it requires a sophisticated representation of the structure of the world.

Critically, cognitive maps typically require a rich representation of the world- they require a level of detail that is not normally available from reinforcement learning processes. That detailed representation of the linkages between adjacent spaces allows for vicarious travel along those linkages. Because it is so much more detailed, it requires orders of magnitude more information than standard reward-motivated reinforcement learning can give. Getting that information cannot occur if it needs extrinsic rewards - those rewards simply are not available in the environment.

We propose, therefore, that latent learning is motivated and enabled by curiosity. However, Tolman conceived of latent learning as a fundamentally passive process, one that took place during apparently purposeless exploration - almost as if by accident. We propose, instead, that latent learning in practice tends to be more actively driven. However, this purposive exploration must be driven by the evolutionary advantage brought by curiosity and ultimately by the extreme information gap experienced by foragers in the natural world.

### The analogy to artificial intelligence

The problems faced by a naturalistic decision-maker or forager are similar in many ways to the problems faced by artificially intelligence (AI). When AI performs classic Atari games, it uses straightforward RL principles [26]. But those games, especially the ones that AI is good at differ from natural situations. The real world – and some games like Pitfall and Montezuma’s Revenge - are what is known as hard-exploration problems [27,28,29]. Rewards are sparse (they require dozens or hundreds of correct actions), so gradient descent procedures are nearly useless. For example, in Pitfall, the first opportunity to gain any points comes after ~60 seconds of perfect play involving dozens of precisely timed moves. Moreover, rewards are often deceptive (they result in highly suboptimal local minima, so getting a small reward promotes adherence to a suboptimal strategy). RL agents that do well at relatively naturalistic hard-exploration games tend to have deliberate hard-coded exploration bonuses [28,29].

The AI domain provides a good illustration of how cognitive maps can be crucial for the success of curiosity. The optimal search strategy in sparse (natural) environments is typically to identify a locally promising region and then perform strategic explorations from that spot to identify subsequent ones [30]. That exploration will not be random, but will take place along identified high-value destinations. AI agents suffer from the problem of *detachment*, that is, when they explore the environment, they leave the relatively high-reward areas of space to explore lower-reward ones [28]. Most such areas are likely to be dead ends, and, when a dead end is detected, the agent ought to return to the high reward area and pursue other promising paths. However, the basic curiosity-based approach, which gives intrinsic rewards for novelty, repels the agent from returning to the promising region of space, precisely because it’s the most familiar and least intrinsically rewarded (it’s also not extrinsically rewarding, because any extrinsic reward has been consumed on the path there, and does not replenish in the meantime). This in turn requires making some kind of internal

map of space so that the agent can return to the locus of high potential reward and explore more efficiently than a wholly random path.

A closely related problem that AI agents - and real-world agents as well - face, is the problem of *derailment* [28]. To explore a space efficiently, an agent must be able to return to promising states and use those as a starting point for efficient exploration. From there, the agent must engage in random search. However, in real environments, returning to a promising state may require a very precise sequence of actions that cannot be deviated from - so stochasticity must be controlled until that state is achieved, at which case it must begin again in earnest. As such, a stochastic search must be carefully controlled depending on one's place in the larger environment - which requires basic mapping functions, and cannot be done with simple RL-type learning. Moreover, important factors governing the exploration process, such as detecting an information gap, deriving the value of information itself, and directing exploration towards potential sites that might be low in external reward but high in information/entropy, simply cannot be supported by only experienced reward history. The key to achieve this, again, is to have a mental map, or internal model, of what is available, and what is novel and potentially offer high information content (high entropy).

### Operational definition of curiosity

Developing these ideas about curiosity, latent learning, and cognitive maps holds great potential in neuroscience. However, it faces several problems from the get-go. We and others have defined curiosity as a motivation to seek information that lacks instrumental or strategic benefit [6,7,8,10]. By this definition, many explorative and playing behaviors qualify as a demonstration of curiosity [9,31]; even risk-seeking and other learning behaviors may be explained that way [32,33,34,35]. But this definition is vague and does not readily lend itself to many laboratory contexts. In an effort to remedy these drawbacks, we developed a conservative operational definition that combines three criteria: (1) a curious research subject is willing to sacrifice primary reward in order to obtain information; (2) the amount of reward a subject is willing to pay scales with the amount of potentially available additional information; and (3) additionally gained information provides no obvious instrumental or strategic benefit.

We devised a more complex task [37] that would circumvent published criticisms [36] of the observing task. This task is based on the observation that monkeys seek counterfactual information - information about what would have happened had they chosen differently [38,39]. In the *counterfactual curiosity task*, monkeys choose between two risky offers. During testing, monkeys are sometimes given the opportunity to choose an option that will provide valid information about the outcome that would have occurred had they chosen the other option. Monkeys are willing to pay to choose this option, indicating that they are curious about counterfactual outcomes. Moreover, monkeys paid more for options that provide more counterfactual information. We speculate that this curiosity-driven information-seeking helps monkeys to develop a sophisticated cognitive map of their task environments [37].

## Functional neuroanatomy of curiosity in the frontal lobes

Our ultimate goal is to understand the neural circuitry underlying curiosity-driven choice. Here we summarize the tentative picture, with a focus on two prefrontal regions, the orbitofrontal cortex (OFC) and the dorsal anterior cingulate cortex (dACC). Both regions are implicated in neuroimaging studies of curiosity [40,41,42]. The neuroanatomy of curiosity is more complex and includes other areas such as hippocampal areas [12,43,44] and basal ganglia [33]. But we would like to highlight OFC and dACC for their potential involvement that bridges curiosity, latent learning, and cognitive maps.

**Orbitofrontal cortex:** We propose that OFC serves to (i) track the intrinsic value of information, (ii) maintain a cognitive map of state space, and (iii) update that map when new information is gained. The clear role of OFC in cognitive mapping has been one of the major intellectual advances of the past decade, and is demonstrated in rodents, monkeys, and humans [23,24,45,46]. This theory, for example, integrates economic findings in OFC with evidence that it carries non-economic variables [46,48,49,50]. This extends to curiosity [51]. In that study, we found that OFC neurons encode the value of information and (confirming much previous work) the value of offers. Critically, OFC uses distinct codes for informational value and more standard juice value. This distinction presumably reflects the role of that information in updating the cognitive map -- even though this information may not offer immediate strategic benefit. In other words, OFC doesn't use a single coherent value code across contexts, but rather, represents task-relevant information in multiple formats, as would be expected in a map rather than a simple reinforcement learning situation. Of course, OFC does not achieve this alone. Studies using similar paradigms show that information is signaled by other systems, including the midbrain dopamine system (e.g. [52,53,54]).

**Dorsal anterior cingulate cortex:** We propose that dACC plays a distinct and complementary role to OFC. Specifically, it appears to track both information delivery and level and task demands for use by OFC in updating the cognitive map and applying it to instrumental use. This idea is motivated by the observation that dACC tracks informativeness, counterfactual information [33,38], environmental demands [55,56], as well as various economic variables (e.g. [57,58,59,60]). It is further motivated by observations about the relative hierarchical positions of the two regions and their relative contributions to choice [61,62,63]. Perhaps most relevantly, in a recent study, White et., al. [33] trained monkeys to associate juice rewards with various reward probabilities with different fractals. Single units in dACC showed increased firing rates to increased uncertainty, and thus to higher expectation of information (when the uncertainty resolved). Moreover, dACC firing rates ramped up to the anticipation of the information that came with the resolution of the uncertainty. In other words, dACC neurons did not simply encode different levels of uncertainty which remained at a constant level for each trial; nor did they ramp up firing rates in anticipation to reward delivery (see also our own results, which paint a somewhat similar picture, [64]).

## Conclusion and future directions

Curiosity has long been treated mystically, as if it is impenetrable to scholarly study. Even when treated as a regular psychological phenomenon, curiosity is often studied in an ad hoc manner. More recent work has made great progress in developing formal approaches to understand the phenomenon systematically study its neural substrates. That formal approach, aided by remarkable progress in AI, has in turn allowed neuroscientists to tentatively start to understand the circuitry of curiosity. That work in turn will likely be critical for understanding naturalistic decision-making, which is marked by the need to make quick decisions with orders of magnitude less information than would be ideal.

## Acknowledgements

We thank Ethan Bromberg-Martin for helpful discussions.

### Funding statement

This research was supported by a National Institute on Drug Abuse Grant R01 DA038106 (to BYH).

## REFERENCES

1. Calhoun AJ, & Hayden BY (2015). The foraging brain. *Current Opinion in Behavioral Sciences*, 5, 24–31.
2. Gottlieb J, & Oudeyer PY (2018). Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience*, 19(12), 758–770. [PubMed: 30397322]
3. Oudeyer PY, & Smith LB (2016). How evolution may work through curiosity-driven developmental process. *Topics in Cognitive Science*, 8(2), 492–502. [PubMed: 26969919]
4. Mobbs D, Trimmer PC, Blumstein DT, & Dayan P. (2018). Foraging for foundations in decision neuroscience: Insights from ethology. *Nature Reviews Neuroscience*, 19(7), 419–427. [PubMed: 29752468] a. This integrative review describes research at the intersection of neuroscience and foraging theory, with an emphasis on how understanding foraging theory – and behind it, evolution, can help neuroscientific theories.
5. Van Gils JA, Schenk IW, Bos O, & Piersma T. (2003). Incompletely informed shorebirds that face a digestive constraint maximize net energy gain when exploiting patches. *The American Naturalist*, 161(5), 777–793.
6. Berlyne DE (1966). Curiosity and exploration. *Science*, 153(3731), 25–33. [PubMed: 5328120]
7. Oudeyer PY, Gottlieb J, & Lopes M. (2016). Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. In *Progress in brain research* (Vol. 229, pp. 257–284). Elsevier. [PubMed: 27926442]
8. Kidd C, & Hayden BY (2015). The psychology and neuroscience of curiosity. *Neuron*, 88(3), 449–460. [PubMed: 26539887]
9. Byrne RW (2013). Animal curiosity. *Current Biology*, 23(11), R469–R470. [PubMed: 23743408]
10. Loewenstein G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological bulletin*, 116(1), 75.
11. Golman R, & Loewenstein G. (2018). Information gaps: A theory of preferences regarding the presence and absence of information. *Decision*, 5(3), 143.a. This important paper integrates much of the psychology and even philosophy of curiosity and highlights the key importance of information gaps – an awareness of a specific lack of knowledge – in motivating curiosity.
12. Gruber MJ, Gelman BD, & Ranganath C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron*, 84(2), 486–496. [PubMed: 25284006]
13. Cervera RL, Wang MZ, & Hayden B. (2020). Curiosity from the Perspective of Systems Neuroscience. *Current Opinion in Behavioral Science*. In press.

14. Lau JKL, Ozono H, Kuratomi K, Komiya A, & Murayama K. (2020). Shared striatal activity in decisions to satisfy curiosity and hunger at the risk of electric shocks. *Nature Human Behaviour*, 1–13. This paper goes further than any other to demonstrate that curiosity works – both neurally and psychologically – much the same way that hunger for food does.
15. Thorndike EL (1927). The law of effect. *The American journal of psychology*, 39(1/4), 212–222.
16. Walton ME, Behrens TE, Buckley MJ, Rudebeck PH, & Rushworth MF (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron*, 65(6), 927–939. [PubMed: 20346766]
17. Blodgett HC (1929). The effect of the introduction of reward upon the maze performance of rats. University of California publications in psychology.
18. Tolman EC (1948). Cognitive maps in rats and men. *Psychological review*, 55(4), 189. [PubMed: 18870876]
19. Bennett D, Bode S, Brydevall M, Warren H, & Murawski C. (2016). Intrinsic valuation of information in decision making under uncertainty. *PLoS computational biology*, 12(7).
20. Grant S, Kajii A, & Polak B. (1998). Intrinsic preference for information. *Journal of Economic Theory*, 83(2), 233–259.
21. Tversky A, & Edwards W. (1966). Information versus reward in binary choices. *Journal of Experimental Psychology*, 71(5), 680. [PubMed: 5939707]
22. Behrens TE, Muller TH, Whittington JC, Mark S, Baram AB, Stachenfeld KL, & Kurth-Nelson Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron*, 100(2), 490–509. [PubMed: 30359611]
23. Wikenheiser AM, & Schoenbaum G. (2016). Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. *Nature Reviews Neuroscience*, 17(8), 513–523. [PubMed: 27256552] a. This masterful review integrates important theories about the cognitive map in OFC with the related work on mapping functions in the hippocampus.
24. Wilson RC, Takahashi YK, Schoenbaum G, & Niv Y. (2014). Orbitofrontal cortex as a cognitive map of task space. *Neuron*, 81(2), 267–279. [PubMed: 24462094]
25. Schoenbaum G, & Roesch M. (2005). Orbitofrontal cortex, associative learning, and expectancies. *Neuron*, 47(5), 633–636. [PubMed: 16129393]
26. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, ... & Petersen S. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533. [PubMed: 25719670]
27. Bellemare M, Srinivasan S, Ostrovski G, Schaul T, Saxton D, & Munos R. (2016). Unifying count-based exploration and intrinsic motivation. In *Advances in neural information processing systems* (pp. 1471–1479).
28. Ecoffet A, Huizinga J, Lehman J, Stanley KO, & Clune J. (2019). Go-explore: a new approach for hard-exploration problems. arXiv preprint arXiv:1901.10995. a. This work describes a new machine learning approach in which AI is endowed with curiosity as well as features that have cognitive map-like aspects, which allow them to perform better on hard-explore problems than earlier ones.
29. Ecoffet A, Huizinga J, Lehman J, Stanley KO, & Clune J. (2018). Montezuma’s revenge solved by go-explore, a new algorithm for hard-exploration problems (sets records on pitfall, too). *Uber Engineering Blog*, Nov.
30. Ecoffet A, Huizinga J, Lehman J, Stanley KO, & Clune J. (2020). First return then explore. arXiv preprint arXiv:2004.12919.
31. Glickman SE, & Sroges RW (1966). Curiosity in zoo animals. *Behaviour*, 26(1–2), 151–187. [PubMed: 5902289]
32. Bromberg-Martin ES, & Sharot T. (2020). The Value of Beliefs. *Neuron*, 106(4), 561–565. [PubMed: 32437655]
33. White JK, Bromberg-Martin ES, Heilbronner SR, Zhang K, Pai J, Haber SN, & Monosov IE (2019). A neural network for information seeking. *Nature communications*, 10(1), 1–19.
34. Farashahi S, Azab H, Hayden B, & Soltani A. (2018). On the flexibility of basic risk attitudes in monkeys. *Journal of Neuroscience*, 38(18), 4383–4398. [PubMed: 29626169]

35. Pirrone A, Azab H, Hayden BY, Stafford T, & Marshall JA (2018). Evidence for the speed–value trade-off: Human and monkey decision making is magnitude sensitive. *Decision*, 5(2), 129. [PubMed: 29682592]
36. Beierholm UR, & Dayan P. (2010). Pavlovian-instrumental interaction in ‘observing behavior’. *PLoS computational biology*, 6(9).
37. Wang MZ, & Hayden BY (2019). Monkeys are curious about counterfactual outcomes. *Cognition*, 189, 1–10. [PubMed: 30889493]
38. Hayden BY, Pearson JM, & Platt ML (2009). Fictive reward signals in the anterior cingulate cortex. *science*, 324(5929), 948–950. [PubMed: 19443783]
39. Heilbrunner SR, & Hayden BY (2016). The description-experience gap in risky choice in nonhuman primates. *Psychonomic bulletin & review*, 23(2), 593–600. [PubMed: 26286883]
40. Charpentier CJ, Bromberg-Martin ES, & Sharot T. (2018). Valuation of knowledge and ignorance in mesolimbic reward circuitry. *Proceedings of the National Academy of Sciences*, 115(31), E7255–E7264.
41. Kobayashi K, & Hsu M. (2019). Common neural code for reward and information value. *Proceedings of the National Academy of Sciences*, 116(26), 13061–13066.
42. van Lieshout LL, Vandenbroucke AR, Müller NC, Cools R, & de Lange FP (2018). Induction and relief of curiosity elicit parietal and frontal activity. *Journal of Neuroscience*, 38(10), 2579–2588. [PubMed: 29439166] a. This study provides a clear and robust link between neural activity and activity in specific brain regions using an innovative and well-designed task.
43. Kang MJ, Hsu M, Krajbich IM, Loewenstein G, McClure SM, Wang JTY, & Camerer CF (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological science*, 20(8), 963–973. [PubMed: 19619181]
44. Jepma M, Verdonschot RG, Van Steenbergen H, Rombouts SA, & Nieuwenhuis S. (2012). Neural mechanisms underlying the induction and relief of perceptual curiosity. *Frontiers in behavioral neuroscience*, 6, 5. [PubMed: 22347853]
45. Schuck NW, Cai MB, Wilson RC, & Niv Y. (2016). Human orbitofrontal cortex represents a cognitive map of state space. *Neuron*, 91(6), 1402–1412. [PubMed: 27657452] a. This work describes evidence in favor of a novel reconceptualization of orbitofrontal cortex as home of a cognitive map. That work takes decision-making out of the world of stimulus response and gives insight into flexibility, learning, and complex thought.
46. Wang MZ, & Hayden BY (2017). Reactivation of associative structure specific outcome responses during prospective evaluation in reward-based choices. *Nature communications*, 8(1), 1–13.
47. Roesch MR, Taylor AR, & Schoenbaum G. (2006). Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron*, 51(4), 509–520. [PubMed: 16908415]
48. Feierstein CE, Quirk MC, Uchida N, Sosulski DL, & Mainen ZF (2006). Representation of spatial goals in rat orbitofrontal cortex. *Neuron*, 51(4), 495–507. [PubMed: 16908414]
49. Wallis JD, Anderson KC, & Miller EK (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature*, 411(6840), 953–956. [PubMed: 11418860]
50. Sleezer BJ, Castagno MD, & Hayden BY (2016). Rule encoding in orbitofrontal cortex and striatum guides selection. *Journal of Neuroscience*, 36(44), 11223–11237. [PubMed: 27807165]
51. Blanchard TC, Hayden BY, & Bromberg-Martin ES (2015). Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron*, 85(3), 602–614. [PubMed: 25619657]
52. Bromberg-Martin ES, & Hikosaka O. (2011). Lateral habenula neurons signal errors in the prediction of reward information. *Nature neuroscience*, 14(9), 1209. [PubMed: 21857659]
53. Bromberg-Martin ES, & Hikosaka O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron*, 63(1), 119–126. [PubMed: 19607797] a. This is the original paper that used the observing task to probe the neural basis of information-seeking. It is important for generalizing theories of the cognitive role of dopamine away from simple kinds of reward.
54. Guru A, Seo C, Post RJ, Kullakanda DS, Schaffer JA, & Warden MR (2020). Ramping activity in midbrain dopamine neurons signifies the use of a cognitive map. *bioRxiv*.



55. Hayden BY, Pearson JM, & Platt ML (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature neuroscience*, 14(7), 933. [PubMed: 21642973]
56. Kolling N, Behrens TE, Mars RB, & Rushworth MF (2012). Neural mechanisms of foraging. *Science*, 336(6077), 95–98. [PubMed: 22491854]
57. Bush G, Vogt BA, Holmes J, Dale AM, Greve D, Jenike MA, & Rosen BR (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proceedings of the National Academy of Sciences*, 99(1), 523–528.
58. Seo H, & Lee D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *Journal of neuroscience*, 27(31), 8366–8377. [PubMed: 17670983]
59. Azab H, & Hayden BY (2017). Correlates of decisional dynamics in the dorsal anterior cingulate cortex. *PLoS biology*, 15(11), e2003091.
60. Azab H, & Hayden BY (2018). Correlates of economic decisions in the dorsal and subgenual anterior cingulate cortices. *European Journal of Neuroscience*, 47(8), 979–993.
61. Rushworth MF, Noonan MP, Boorman ED, Walton ME, & Behrens TE (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron*, 70(6), 1054–1069. [PubMed: 21689594]
62. Hunt LT, Malalasekera WN, de Berker AO, Miranda B, Farmer SF, Behrens TE, & Kennerley SW (2018). Triple dissociation of attention and decision computations across prefrontal cortex. *Nature neuroscience*, 21(10), 1471–1481. [PubMed: 30258238] a. In this careful study, the authors disambiguate three of the usual suspects regions of economic choice and show that their functions can be triply dissociated.
63. Yoo SBM, & Hayden BY (2018). Economic choice as an untangling of options into actions. *Neuron*, 99(3), 434–447. [PubMed: 30092213]
64. Wang MZ, & Hayden BY (2020). Curiosity is associated with enhanced tonic firing in dorsal anterior cingulate cortex. *bioRxiv*. doi.org/10.1101/2020.05.25.115139a. In this study, we use a version of an observing task to compare activity in dACC and OFC. We find results suggestive of encoding of demand for information, or information gap, in dACC, but not in OFC.

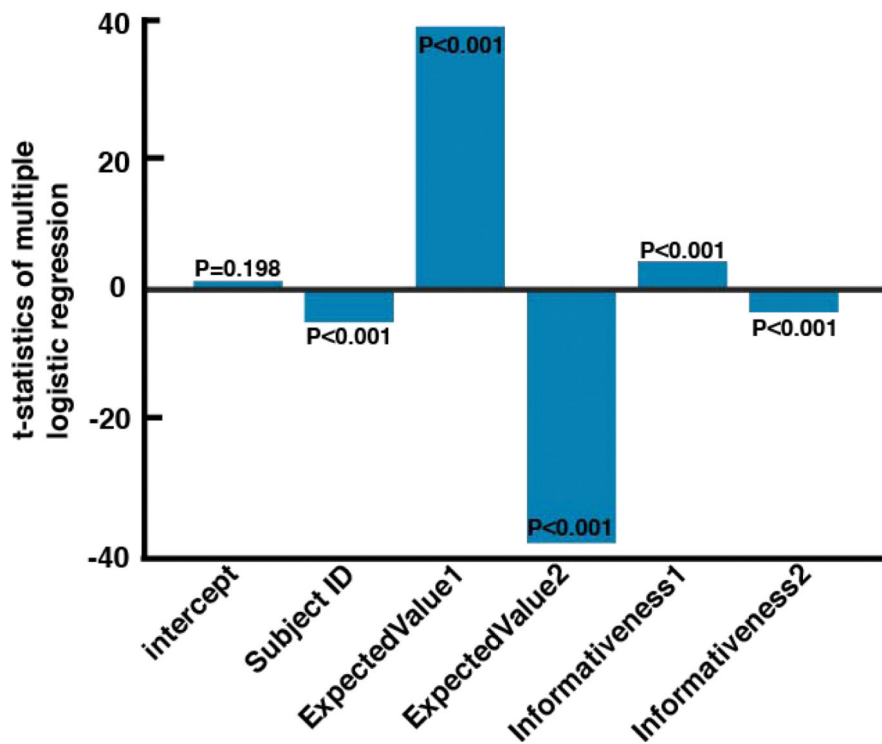
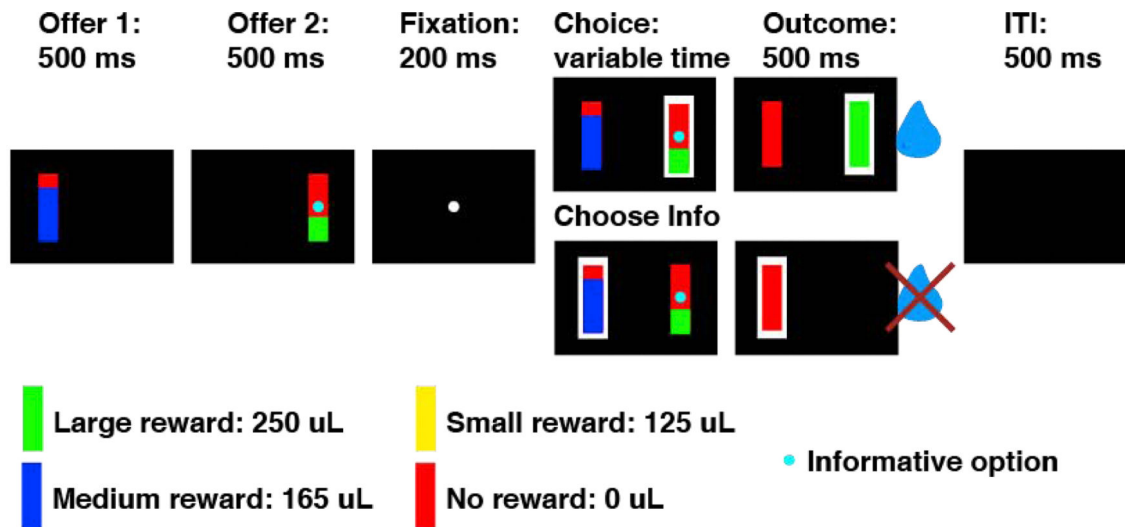


**Figure 1.**

A virtual maze our lab has used for monkeys based on the classic alley maze of Tolman. Tolman and his graduate students placed rats in mazes like this one and found that they would explore the maze unrewarded and would demonstrably learn the features of the structure of the maze in the absence of rewards, a result that is difficult to explain using then-dominant simple stimulus-response learning theories. Tolman proposed that the rats generated a cognitive map that instantiated features of the maze and could be consulted to drive flexible behavior.

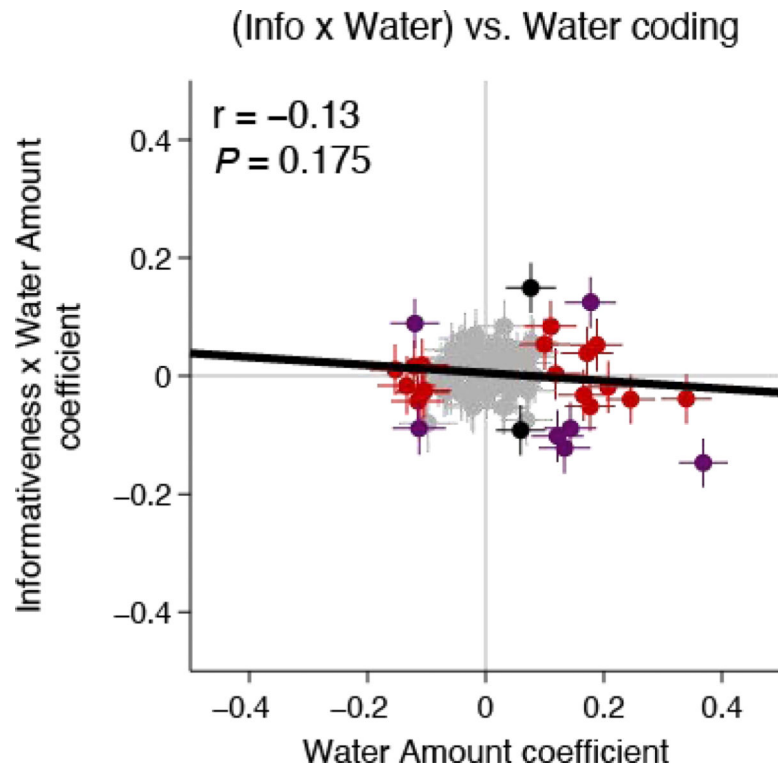


**Figure 2.** Monkey in a tree, illustrating the problem of *derailment* in curiosity research. The monkey must learn foraging strategy through trial and error, which requires a highly variable exploration of the environment. But getting to the end of a branch is somewhat risky and requires suppressing stochastic variability. To successfully deploy curiosity the monkey must have a cognitive map of where variability is good and where it is bad.



**Figure 3.**

In our curiosity task, subjects could choose between risky options for juice rewards. In some trials, they could also gain information about what would have occurred had they chosen differently. By analyzing preference curves on such trials, we could quantify their subjective value of counterfactual information. We found a small but significant positive valuation of counterfactual information in both subjects tested (Wang and Hayden, 2019).



**Figure 4.** Responses of an ensemble of OFC neurons to offers varying in informational value and reward value (Blanchard et al., 2015). We find that individual neurons encode both variables (horizontal and vertical axes indicate tuning coefficients for the two dimensions respectively). However, those codes are themselves uncorrelated, as indicated by the lack of a significant slope between the two dimensions.