NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Orbitofrontal Cortex Supports Behavior and Learning Using Inferred but not Cached Values

Joshua L. Jones[1,*], **Guillem R. Esber**[2], **Michael A. McDannald**[2], **Aaron J. Gruber**[3], **Alex Hernandez**[1], **Aaron Mirenzi**[2], and **Geoffrey Schoenbaum**[1,2,*]

[1]Department of Anatomy and Neurobiology, University of Maryland, School of Medicine 20 Penn St. Baltimore, MD 21201, U.S.A

[2]National Institute on Drug Abuse Intramural Research Program, Cellular Neurobiology Research Branch, Behavioral Neurophysiology Research Section, 251 Bayview Blvd., Baltimore, MD 21201, U.S.A

[3]Department of Neuroscience, University of Lethbridge, Alberta, T1K 3M4, Canada

Abstract

Computational and learning theory models propose that behavioral control reflects value that is both cached (computed and stored during previous experience) and inferred (estimated on-the-fly based on knowledge of the causal structure of the environment). The latter is thought to depend on the orbitofrontal cortex. Yet, some accounts propose that the orbitofrontal cortex contributes to behavior by signaling "economic" value, regardless of the associative basis of the information. We found that the orbitofrontal cortex is critical for both value-based behavior and learning when value must be inferred but not when a cached value is sufficient. The orbitofrontal cortex is thus fundamental for accessing model-based representations of the environment to compute value rather than for signaling value per se.

Computational and learning theory accounts have converged on the idea that reward-related behavioral control reflects two types of information (1–3). The first is derived from habits, policies or cached values. These terms reflect underlying associative structures that incorporate a pre-computed value stored during previous experience with the relevant cues. Behaviors based on this sort of information are fast and efficient, but do not take into account changes in the value of the expected reward. This type of information contrasts with the second category, referred to as goal-directed or model-based, in which the value is inferred from knowledge of the associative structure of the environment, including how to obtain the expected reward, its unique form and features, and current value. The associative model is stored, but a pre-computed value is not. Rather the value is computed or inferred on-the-fly when it is needed. While behavior based on inferred value is slower, it can be more adaptive and flexible.

Though evidence suggests that different brain circuits mediate their respective influences (1–3), much of cognitive neuroscience - and particularly neuroeconomics - does not attend to these distinctions. For example, proposals for a common neural currency to allow the comparison of incommensurable stimuli (e.g. apples and oranges) typically do not clearly specify the associative structure underlying the value computation. And since economic value is typically measured through revealed preferences, with no explicit control for the

*To whom correspondence should be addressed: geoffrey.schoenbaum@nih.gov.

source of the underlying value, it would *by default* include both cached and inferred value, at least as defined by computational and learning theory accounts (1–3).

The calculation of economic value is often assigned to the orbitofrontal cortex (OFC), a prefrontal area heavily implicated in value-guided behavior (4–6). Yet, behavioral studies across species implicate this region broadly not in value-guided decisions per se, but rather in behaviors that require a new value to be estimated after little or no direct experience (7–14). Further, the OFC is often involved in a behavior depending upon whether learning is required (10, 15, 16), even when that learning does not involve changes in value (17). These data seem to require the OFC to perform one function, anticipating outcomes, in some settings, while it performs another, calculating economic value, in others. However an alternative hypothesis is that the OFC performs the same function in all settings, specifically contributing to value-guided behavior and learning when value must be inferred or derived from model-based representations. We tested this hypothesis in rats using sensory preconditioning and blocking.

In sensory preconditioning, a subject is taught a pairing between two cues (e.g. white noise and tone) and later learns that one of these cues predicts a biologically meaningful outcome (e.g. food) (18). Thereafter the subject will exhibit a strong conditioned response to both the reward-paired cue and the preconditioned cue. The response to the preconditioned cue differs from the response to the reward-paired cue, in that it cannot be based on a cached value; rather it must reflect the subject's ability to infer value by virtue of a knowledge of the associative structure of the task (see Supplementary Discussion for further details). If the OFC is involved in behavior that requires inferred value, then inactivating it at the time of this test should prevent behavior driven by this preconditioned cue, while leaving unimpaired behavior driven by the reward-paired cue.

Cannulae were implanted bilaterally in the OFC of rats (19 controls; 16 inactivated or OFCi) at coordinates used previously (12, 19) (Fig. 1A and B). After recovery from surgery, these rats were food deprived and then trained in a sensory preconditioning task (Fig. 1; Materials and Methods).

In preconditioning, rats were taught to associate 2 pairs of unrelated auditory cues (A->B and C->D; clicker, white noise, tone, siren; counterbalanced). Food cup responding was measured during presentation of each cue versus baseline as an index of conditioning; the rats responded at baseline levels to all cues (Fig. 1A and B). A 2-factor ANOVA (cue x treatment) comparing the percentage of time spent in the food cup during each cue found no effects (F's < 1.27; p's > 0.29).

In conditioning, rats were taught that one of the preconditioned cues (B) predicted reward. As a control, the other preconditioned cue (D) was presented without reward. Rats learned to discriminate between the rewarded (B) and non-rewarded cue (D), increasing responding across sessions during the former more than the latter (Fig. 1C and D). A 3-factor AVOVA (cue x treatment x session) revealed significant main effects of cue ($F_{(1,33)} = 170.5$, $p < 0.0001$) and session ($F_{(5,165)} = 54.75$, $p < 0.0001$), and a significant cue x session interaction ($F_{(5,165)} = 64.6$, $p < 0.0001$), but no significant main effect nor any interactions with treatment (F's <1.49, p's > 0.19).

In the probe test, we assessed responding to the preconditioned cues (A and C) after infusions of either saline or a GABA agonist cocktail containing baclofen and muscimol. Rats received 3 rewarded presentations of B and D followed by 6 unrewarded presentations of A and C, in a counterbalanced design. Both control and OFCi rats exhibited robust responding to the reward-paired cue (B) (Fig. 1E and F) and not to the cue that was presented without reward (D). An analysis restricted to the first presentation of each cue,

prior to any reward delivery, revealed a significant main effect of cue ($F_{(1,33)} = 53.21$, $p < 0.0001$) and no significant effect or interaction with treatment (F's < 1.9, p's > 0.17). However, only controls showed elevated responding to the preconditioned cue (A) that had been paired with the reward-paired cue (B) (Fig. 1E). Controls responded significantly more to this cue than to the preconditioned cue (C) signaling the non-rewarded cue (D) whereas OFCi rats responded to both preconditioned cues similarly and at a level comparable to the responding shown to the cue signaling non-reward (D) (Fig. 1F). A 2-factor ANOVA (cue x treatment) indicated a significant main effect of cue ($F_{(1,33)} = 14.7$, $p < 0.001$) and a significant interaction between cue and treatment ($F_{(1,33)} = 7.33$, $p < 0.01$). Both groups responded significantly more to B than D, and controls responded significantly more to A than to either C or D (Bonferroni post-hoc correction; p's < 0.05), while OFCi rats responded similarly to these three cues (p's > 0.05).

These results show that the OFC is required when behavior must be based on inferred, but not cached value. However, they do not address how OFC is involved in learning. To test this question, we utilized blocking (20). In blocking, a subject is taught that a cue predicts reward (e.g. tone predicts food); later that same cue is presented together with a new cue (e.g. light-tone), still followed by reward. If this is done, the subject will subsequently show little conditioned responding to the new cue (e.g. the light in our example). The ability of the original cue to predict the reward is said to block learning. The OFC is not required for this type of blocking (17, 21), indicating that the OFC is not necessary for modulating learning based on cached value, but this does not address whether the OFC is necessary for blocking based on inferred value. If the OFC performs the same function during learning, then inactivating the OFC during blocking with the preconditioned cue should result in unblocking.

To test this, we trained a subset of rats from the experiment described above in an inferred value blocking task (Fig. 2; see Supplementary Materials). Rats underwent 2 days of training in which the preconditioned auditory cues (A and C) were presented with novel light cues (X and Y; houselight, flashing cue light; counterbalanced). Both pairs of cues were reinforced with the same reward previously paired with B (AX->sucrose; CY->sucrose). Prior to each session, rats received infusions of saline or the baclofen/muscimol cocktail. Both groups showed a significant increase in responding, and there was no overall difference between the two groups (Fig. 2A). A 3-factor ANOVA (treatment x cue x session) demonstrated a significant effect of session ($F_{(1,19)} = 16.53$, $p<0.001$), but there was no significant main effect nor any interactions with treatment (F's < 1.44, p's > 0.24).

One day after blocking, these rats were presented with several AX and CY reminder trials, followed by unreinforced trials in which cues X and Y were presented. Controls exhibited significantly higher conditioned responding to the control cue (Y) than to the blocked cue (X). Indeed responding to the blocked cue was no greater than baseline (Fig. 2B; $t_{(1,11)} = 0.67$, $p=0.52$). By contrast, OFCi rats showed increased responding to both cues, consistent with an inability to use inferred value to block learning (Fig. 2B). A 2-factor ANOVA (treatment x cue) revealed a significant interaction between cue and treatment ($F_{(1,19)} = 7.70$, $p=0.012$), and controls responded significantly more to Y than to X (Bonferroni post-hoc correction; $p < 0.05$), whereas OFCi rats responded similarly to these two cues ($p > 0.05$).

These findings demonstrate that the OFC is involved in value-based behavior when the value must be inferred from an associative model of the task, but not when the same behavior can be based entirely on a value cached or stored in cues based on past experience. This is consistent with previous results implicating the OFC in changes in conditioned responding after reinforcer devaluation (7, 8, 10, 13, 22, 23). Our results confirm that OFC

is required for knowledge of the associative structure at the time of the decision, rather than due to some idiosyncratic involvement in taste perception, reward learning per se, or devaluation, because in our task we did not alter the reward in any way. By inactivating only at the time of the probe test, we show clearly that the OFC is required for using the previously acquired associative structure. Thus it may be stored elsewhere, but it cannot be applied to guide behavior effectively without the OFC. By including an explicit control for cached value, we show that this deficit is specific for inferred value at the time of decision-making. These data are also consistent with several fMRI studies showing that neural activity in OFC may be particularly well-tuned to reflect model-based information at the time of decision-making (13, 24). In this regard, it is notable that this same function was also required for modulating learning, since the preconditioned cue also failed to function as a blocker after OFC inactivation. This suggests that the OFC plays a general role in signaling inferred value, which might be used by other brain regions for a variety of purposes, rather than a special role in the service of decision-making.

It is remarkable that the inferred value signal evoked by this cue resulted in blocking. Blocking normally utilizes a cue that has been paired directly with reward. Theoretical accounts focus on the fact that the value of the expected reward is already fully predicted and therefore there is no prediction error to drive learning (25–27). However these accounts do not specify the source of this value, and generally it is assumed to be a sort of cached or general value. In fact, temporal difference reinforcement learning, is specific on this point (3). Our experiment shows that inferred value can also modulate learning by serving as a blocking cue, allowing learning to be modulated not only by experienced information, but also by inferred knowledge.

These results also contradict the argument that the OFC is specifically tasked with calculating value in a common currency, devoid of identifying information about identity and source, since the OFC was necessary for value-based behavior only when calculation of that value required a model-based representation of the task. Indeed, the OFC is necessary for behavior and learning when only the identity of the reward is at issue (17, 28), suggesting that the OFC functions as part of a circuit that maintains and utilizes the states and transition functions that comprise model-based control systems, rather than as an area that blindly calculates general or common value. Remarkably, none of these results *require* that value be calculated within OFC at all. Though radical, such speculation is in line with evidence that other prefrontal regions do as well or better than OFC in representing general outcome value (29). Moreover while activity in some OFC neurons correlates with economic value, representations are usually much more specific to elements of task structure (29, 30). The OFC may only be necessary for economic decision making insofar as the value required reflects inferences or judgments analogous to what we have tested in here. Data implicating the OFC in the expression of transitive inference (11) or willingness to pay (14) may reflect such a function, since in each setting, the revealed preferences are expressed after little or no experience with the imagined outcomes. Limited experience, a defining feature of economic decision-making (5), would minimize the contribution of cached values, biasing subjects to rely on model-based information for the values underlying their choices.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References and Notes

1. Daw ND, Niv Y, Dayan P. Nature Neuroscience. 2005; 8:1704.

2. Cardinal RN, Parkinson JA, Hall G, Everitt BJ. Neuroscience and Biobehavioral Reviews. 2002; 26:321. [PubMed: 12034134]

3. Balleine, BW.; Daw, ND.; O'Doherty, JP. Neuroeconomics: Decision Making and the Brain. Glimcher, PW.; Camerer, CF.; Fehr, E.; Poldrack, RA., editors. Elsevier; Amsterdam: 2008. p. 367-388.

4. Montague PR, Berns GS. Neuron. 2002; 36:265. [PubMed: 12383781]

5. Padoa-Schioppa C. Annual Review of Neuroscience. 2011; 34:333.

6. Kim H, Shimojo S, O'Doherty JP. Cerebral Cortex. 2011; 21:769. [PubMed: 20732900]

7. Gallagher M, McMahan RW, Schoenbaum G. Journal of Neuroscience. 1999; 19:6610. [PubMed: 10414988]

8. West EA, DesJardin JT, Gale K, Malkova L. Journal of Neuroscience. 2011; 31:15128. [PubMed: 22016546]

9. Machado CJ, Bachevalier J. European Journal of Neuroscience. 2007; 25:2885. [PubMed: 17561849]

10. Izquierdo AD, Suda RK, Murray EA. Journal of Neuroscience. 2004; 24:7540. [PubMed: 15329401]

11. Camille N, Griffiths CA, Vo K, Fellows LK, Kable JW. Journal of Neuroscience. 2011; 31:7527. [PubMed: 21593337]

12. Takahashi Y, et al. Neuron. 2009; 62:269. [PubMed: 19409271]

13. Gottfried JA, O'Doherty J, Dolan RJ. Science. 2003; 301:1104. [PubMed: 12934011]

14. Plassmann H, O'Doherty J, Rangel A. Journal of Neuroscience. 2007; 27:9984. [PubMed: 17855612]

15. Jones B, Mishkin M. Experimental Neurology. 1972; 36:362. [PubMed: 4626489]

16. Hornak J, et al. Journal of Cognitive Neuroscience. 2004; 16:463. [PubMed: 15072681]

17. McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G. Journal of Neuroscience. 2011; 31:2700. [PubMed: 21325538]

18. Brogden WJ. Journal of Experimental Psychology. 1939; 25:323.

19. Burke KA, Takahashi YK, Correll J, Brown PL, Schoenbaum G. European Journal of Neuroscience. 2009; 30:1941. [PubMed: 19912335]

20. Kamin, LJ. Punishment and Aversive Behavior. Campbell, BA.; Church, RM., editors. Appleton-Century-Crofts; New York: 1969. p. 242-259.

21. Burke KA, Franz TM, Miller DN, Schoenbaum G. Nature. 2008; 454:340. [PubMed: 18563088]

22. Pickens CL, Saddoris MP, Gallagher M, Holland PC. Behavioral Neuroscience. 2005; 119:317. [PubMed: 15727536]

23. Critchley HD, Rolls ET. Journal of Neurophysiology. 1996; 75:1673. [PubMed: 8727405]

24. Hampton AN, Bossaerts P, O'Doherty JP. Journal of Neuroscience. 2006; 26:8360. [PubMed: 16899731]

25. Pearce, JM.; Kaye, H.; Hall, G. Quantitative Analyses of Behavior. Commons, ML.; Herrnstein, RJ.; Wagner, AR., editors. Vol. 3. Ballinger; Cambridge, MA: 1982. p. 241-255.

26. Sutton, RS.; Barto, AG. Reinforcement Learning: An introduction. MIT Press; Cambridge MA: 1998.

27. Rescorla, RA.; Wagner, AR. Classical Conditioning II: Current Research and Theory. Black, AH.; Prokasy, WF., editors. Appleton-Century-Crofts; New York: 1972. p. 64-99.

28. Ostlund SB, Balleine BW. Journal of Neuroscience. 2007; 27:4819. [PubMed: 17475789]

29. Kennerley SW, Behrens TE, Wallis JD. Nature Neuroscience. 2011 AOP.

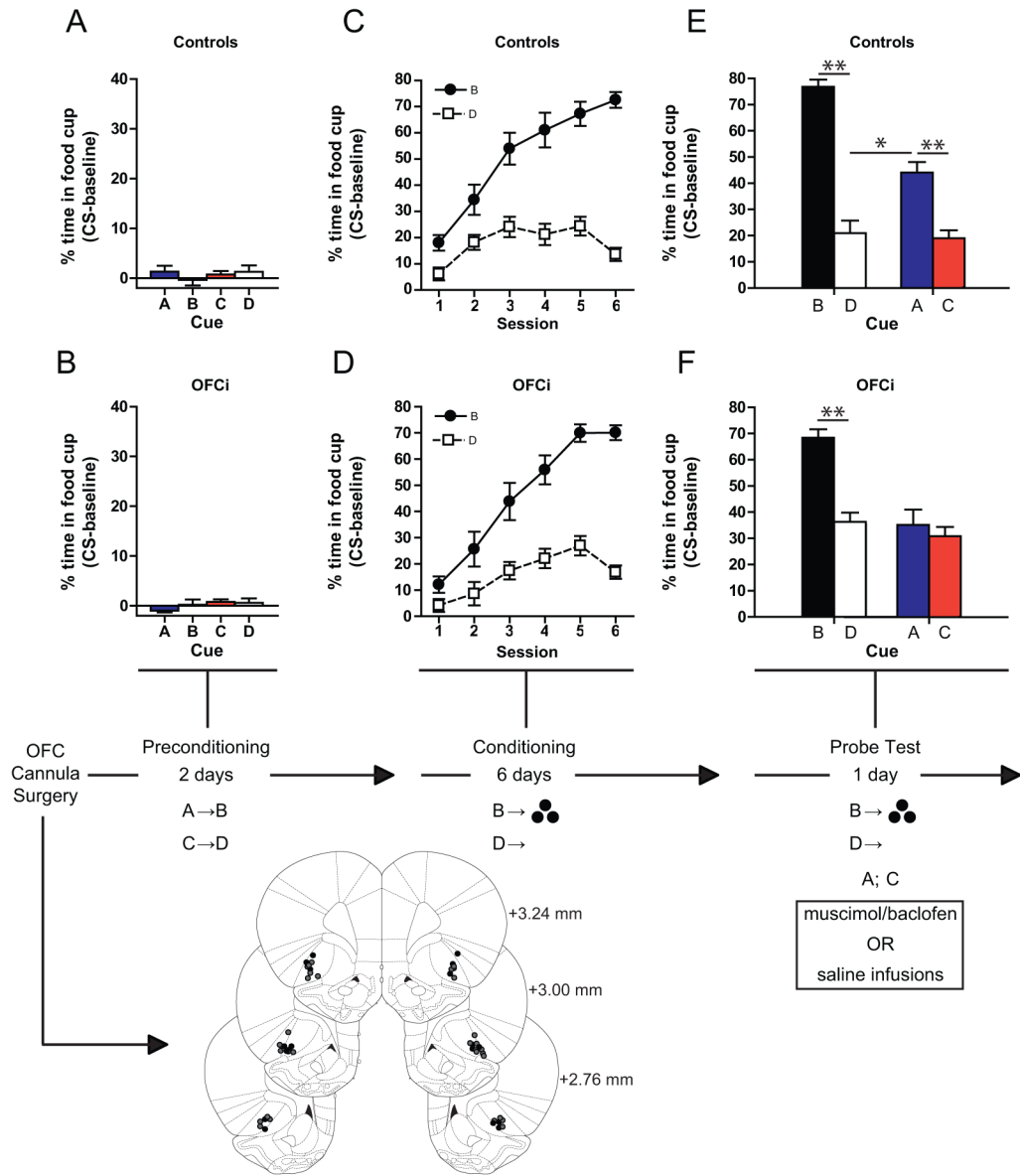30. Padoa-Schioppa C, Assad JA. Nature. 2006; 441:223. [PubMed: 16633341]

Fig. 1.
The OFC is necessary when behavior is based on inferred value. Figures show the percentage of time spent in food cup during presentation of the cues during each of the three phases of training: preconditioning (A,B), conditioning (C,D), and the probe test (E,F). OFC was inactivated only during the probe test. Cannulae positions are shown below, vehicle (Black circles), OFCi (Gray circles). *p < 0.05, **p < 0.01 or better. Error bars = SEM.
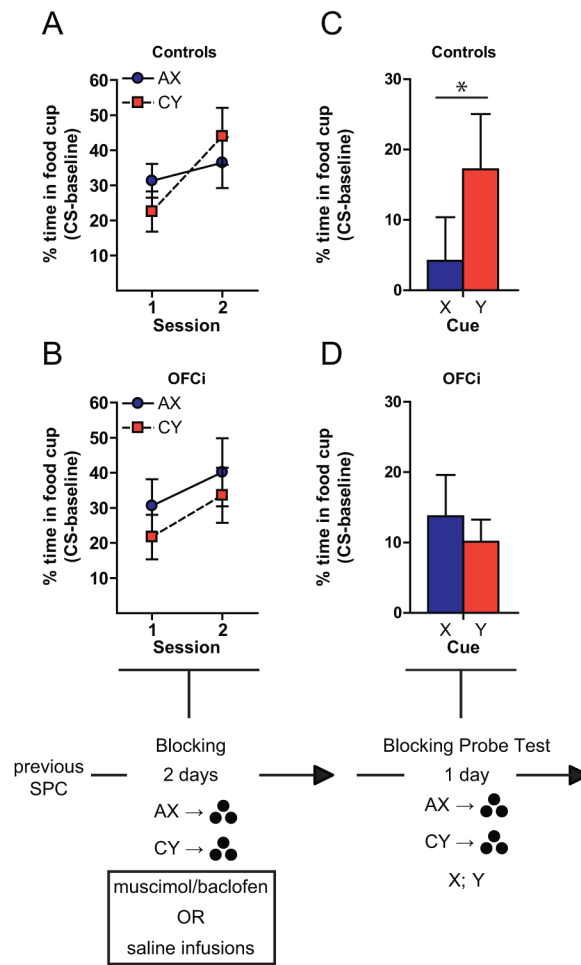
Fig. 2.
The OFC is necessary when learning is based on inferred value. Figures show the percentage of time spent in food cup during presentation of the cues during blocking (A,B) and the subsequent probe test (C,D). OFC was inactivated during blocking. $*p < 0.05$, $**p < 0.01$ or better. Error bars = SEM.