



Foundations of human reasoning in the prefrontal cortex

Maël Donoso *et al.*

Science **344**, 1481 (2014);

DOI: 10.1126/science.1252254

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of July 2, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/344/6191/1481.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2014/05/28/science.1252254.DC1.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/content/344/6191/1481.full.html#related>

This article **cites 33 articles**, 7 of which can be accessed free:

<http://www.sciencemag.org/content/344/6191/1481.full.html#ref-list-1>

This article has been **cited by** 1 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/344/6191/1481.full.html#related-urls>

This article appears in the following **subject collections**:

Neuroscience

<http://www.sciencemag.org/cgi/collection/neuroscience>

RESEARCH ARTICLE

HUMAN COGNITION

Foundations of human reasoning in the prefrontal cortex

Maël Donoso,^{1,2,3} Anne G. E. Collins,^{2,4} Etienne Koechlin^{1,2,3*}

The prefrontal cortex (PFC) subserves reasoning in the service of adaptive behavior. Little is known, however, about the architecture of reasoning processes in the PFC. Using computational modeling and neuroimaging, we show here that the human PFC has two concurrent inferential tracks: (i) one from ventromedial to dorsomedial PFC regions that makes probabilistic inferences about the reliability of the ongoing behavioral strategy and arbitrates between adjusting this strategy versus exploring new ones from long-term memory, and (ii) another from polar to lateral PFC regions that makes probabilistic inferences about the reliability of two or three alternative strategies and arbitrates between exploring new strategies versus exploiting these alternative ones. The two tracks interact and, along with the striatum, realize hypothesis testing for accepting versus rejecting newly created strategies.

Human reasoning subserves adaptive behavior and has evolved facing the uncertainty of everyday environments. In such situations, probabilistic inferential processes (i.e., Bayesian inferences) make optimal use of available information for making decisions. Human reasoning involves Bayesian inferences accounting for human responses that often deviate from formal logic (1). Bayesian inferences also operate in the prefrontal cortex (PFC) and guide behavioral choices (2, 3). Everyday environments, however, are changing and open-ended, so that the range of uncertain situations and associated behavioral strategies (i.e., internal maps linking stimuli, actions, and expected outcomes) becomes potentially infinite. In such environments, probabilistic inferences involve Dirichlet process mixtures (4–7) and rapidly yield intractable computations. This computational complexity problem constitutes a fundamental constraint on the evolution of higher cognitive functions and raises the issue of the actual nature of inferential processes implemented in the PFC.

A model of reasoning processes in the human PFC

To address this issue, we proposed a model (8) that describes human reasoning, as it guides behavior, as a computationally tractable, online algorithm approximating Dirichlet process mixtures (9). The algorithm combines forward Bayesian inferences operating over a few concurrent behavioral strategies stored in long-term memory with

hypothesis testing for possibly updating this inferential buffer with new strategies formed from long-term memory. The algorithm notably serves to arbitrate between (i) staying with the ongoing behavioral strategy and possibly learning external contingencies, (ii) switching to other learned strategies, and (iii) forming new behavioral strategies.

For integrating online Bayesian inferences and hypothesis testing, the algorithm's key feature is inferring the absolute reliability of every

monitored strategy: namely, the posterior probability that the current situation matches the situation the strategy has learned, given both action outcomes (and possibly contextual cues), and the possibility that no match occurs with any monitored strategies. To estimate these probabilities, the model assumes that, in the latter case, action outcomes expected from the monitored strategies are equiprobable (9). Thus, every monitored strategy may appear as being either reliable (i.e., more likely matching than not matching the current situation) or unreliable (the converse). When a strategy is reliable, the others are necessarily unreliable, so that the algorithm is an exploitation state (Fig. 1): The reliable strategy is the actor, namely, the unique strategy for selecting and learning the actions that maximize rewards (typically through reinforcement learning), whereas the other monitored strategies are treated as counterfactual. When all monitored strategies become unreliable, the algorithm then switches into an exploration state corresponding to hypothesis testing: A new strategy is formed as a weighted mixture of strategies stored in long-term memory, then probed and monitored as actor (9). If the strategy is a priori unreliable, this probe actor learns, so that the algorithm may subsequently return to the exploitation state in two ways. Either one counterfactual strategy becomes reliable, while the probe actor remains unreliable: The former is then retrieved as actor, and the latter is rejected (disbanded). Or the probe actor becomes reliable, while counterfactual strategies remain unreliable. The probe actor is then confirmed: It remains the actor, the new strategy is simply consolidated into long-term memory, and the repertoire

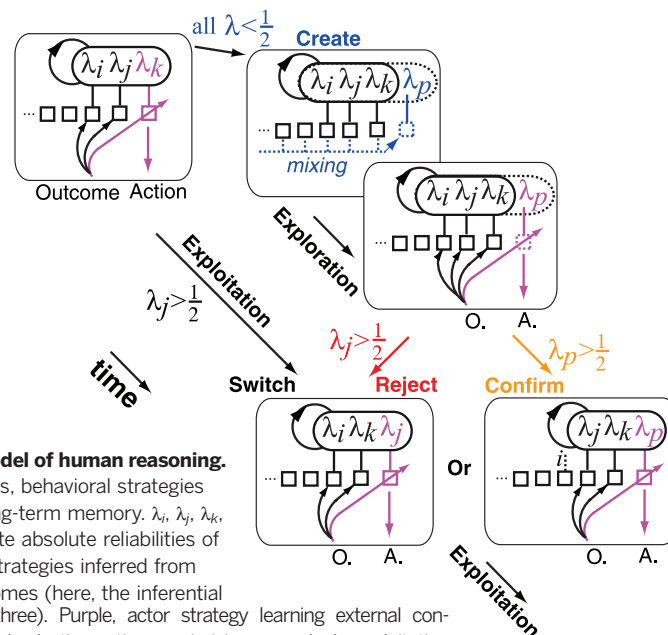


Fig. 1. A model of human reasoning.

Solid squares, behavioral strategies stored in long-term memory. λ_i , λ_j , λ_k , and λ_p denote absolute reliabilities of monitored strategies inferred from action outcomes (here, the inferential capacity is three). Purple, actor strategy learning external contingencies and selecting action maximizing rewards. In exploitation periods, the actor is reliable (i.e., $\lambda_{actor} > 1 - \lambda_{actor}$ or $\lambda_{actor} > \frac{1}{2}$) and the others are necessarily unreliable (because $\sum \lambda \leq 1$). Otherwise, the system switches into exploration (all $\lambda < \frac{1}{2}$) and creates a probe actor (p) from mixing strategies stored in long-term memory (blue). Exploration periods terminate when either one counterfactual strategy (j) or probe actor (p) becomes reliable: The probe actor is then rejected (red) or confirmed (orange). See text for details.

¹INSERM, Laboratoire de Neurosciences Cognitives (U960), 29 rue d'Ulm, 75005 Paris, France. ²Département d'Etudes Cognitives (DEC), Ecole Normale Supérieure, 29 rue d'Ulm, 75005 Paris, France. ³Centre de Neuro-imagerie de Recherche (CENIR), Université Pierre et Marie Curie, 4 place Jussieu, 75005 Paris, France. ⁴Brown University, Providence, RI 02912, USA.

*Corresponding author. E-mail: etienne.koechlin@upmc.fr

of stored strategies is expanded. In case the inferential buffer has further reached its capacity limit, the counterfactual strategy used the least recently as actor is then discarded from the buffer (but remains stored in long-term memory).

Consistent with the capacity limit of human working memory (10), human decisions are best predicted when the inferential buffer is limited to two or three concurrent counterfactual strategies (8). We then hypothesized that the human PFC implements this algorithm. We expected anterior PFC regions to form the inferential buffer (3, 11–13) and more posterior PFC regions in association with basal ganglia to drive actor learning, selection, and creation on the basis of hypothesis testing (14–18). The model predicts that anterior PFC regions concurrently infer the absolute reliability of actor and counterfactual strategies that the algorithm builds online. More posterior PFC regions then detect when, in the inferential buffer, actor strategies become unreliable for creating probe actors, as well as when counterfactual strategies become reliable for retrieving them as actor (and possibly rejecting probe actors). In basal ganglia, the ventral striatum subserves reinforcement learning (16, 19, 20) and is predicted to detect when, in the inferential buffer, probe actors become reliable for confirming them in long-term memory (21).

Behavioral paradigm

To test these predictions, we used functional magnetic resonance imaging (fMRI) and scanned 40 healthy participants, while they were responding to successively presented digits and searching for three-digit combinations by trial and error (fig. S1) (9). Feedbacks were noisy, and combinations changed episodically. Unbeknownst to them, participants performed two distinct sessions. In the open session, every episode corresponded to new combinations, whereas in the recurrent session, only three combinations re-occurred unpredictably across episodes. The protocol thus induced participants to reason from feedbacks whether they had to persevere with the same combination and possibly adjust it, reuse previously learned ones, or learn by searching for new combinations.

In every trial, participants' responses were either correct, perseverative (incorrect in the current episode but correct in the preceding episode), or exploratory (neither correct nor perseverative). Overall, participants performed much below the statistical optimum (8). In both conditions, correct response rates increased from ~5% at episode onsets to a plateau at ~85% about 25 trials later (chance level: 25%) (Fig. 2, left). Exploratory response rates increased from ~10% at episode onsets, peaked at ~40% five trials later, and then returned to ~10% (chance level: 50%). Correct responses increased and exploratory responses vanished faster in the recurrent episodes than in the open episodes (both F values > 21.8 , P values < 0.0001). In the first trials of recurrent episodes, furthermore, a positive feedback caused the production of correct responses in the next trial, even when the two successively presented digits differed:

The statistical dependence between two successive correct responses increased in the first trials of recurrent compared with open episodes (trials 1 and 2: T values > 2.25 , $P < 0.03$) (Fig. 2, bottom), while remaining similar in both conditions on the following trials. In these first recurrent trials, accordingly, participants used feedbacks to retrieve previously learned combinations rather than re-collecting each digit-response association separately. Participants consequently built and stored multiple combinations and monitored feedbacks for either retrieving these combinations or learning new ones. Combinations thus defined behavioral strategies associating digits, responses, and expected feedbacks.

We fit the model free parameters (buffer-capacity, prior reliability, recollection entropy of probe actors, and reinforcement learning parameters) to each participant's series of responses (table S1) (9). In both recurrent and open episodes, the fitted model predicted participants' responses and their statistical dependencies across successive trials (Fig. 2, right). The model fit significantly better than alternative models, independently of model complexity and fitting criteria (fig. S2) (9). Moreover, fitted parameters were independent of which session was fitted ($T < 1$) and, consequently, unrelated to the number of

combinations used in recurrent sessions. The best-fitting capacity—whether fixed or averaged across subjects—included two counterfactual strategies (mean = 2.6, SEM = 0.24, median = 2) (table S1).

The model critically reveals that the gradual variations of responses reported above are actually artifacts from aligning performances from episode onsets and averaging across episodes (Fig. 3). After most episode changes (93.9% and 94.2% of recurrent and open episodes, respectively), indeed, the algorithm switched from exploitation to exploration and created probe actors from long-term memory at variable time points across episodes [on average 3.3 (SD = 0.9) and 4.2 (SD = 1.3) trials after recurrent and open episode onsets, respectively]. We refer to these algorithmic transitions as switch-in events. Realigning model and participants' performances on these switch-in events rather than episode onsets (Fig. 3, left) shows that, in exploitation trials preceding switch-in events, both model and participants' responses were virtually unaffected by episode changes and remained mostly perseverative (~85 to 90%), whereas residual responses remained randomly distributed across exploratory and correct responses (~8% and ~4% of residual responses, respectively). In switch-in trials, by contrast,

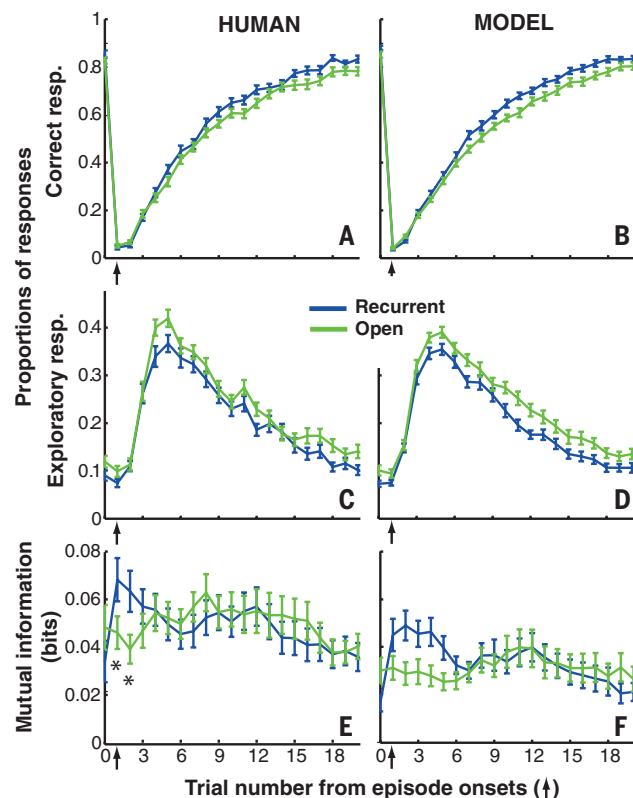


Fig. 2. Behavioral performances following episode changes. Proportion of correct (top) and exploratory (middle) responses from episode onsets (arrows). (A, C, and E) participants' performances. (B, D, and F) Fitted model predictions in every trial given participants' responses in previous trials. (E and F) Statistical dependences between two successive correct responses produced by participants (E) and fitted model simulations (F) (mutual information computed over five-trials sliding windows). Green: open episodes; blue: recurrent episodes. Error bars are SEM across participants. See table S1 for model parameters. * $P < 0.05$.

perseverative responses abruptly dropped off (~40%), and exploratory responses abruptly increased to a plateau (~35 to 40%). In exploration trials following switch-in events, both model and participants' exploratory responses remain on the plateau, whereas perseverative responses slowly decreased (correct responses consequently increased slowly).

In 43% of recurrent episodes, the algorithm terminated these exploration periods by retrieving counterfactual strategies and rejecting probe actors [on average 10.1 (SD = 3.2) after episode onsets]. In the remaining recurrent episodes (57%) and most open episodes (84%), the algorithm

terminated exploration by confirming probe actors in long-term memory [on average 6.7 (SD = 3.3) and 8.1 (SD = 3.9) trials after recurrent and open episode onsets, respectively]. We refer to these algorithmic transitions as rejection and confirmation events, respectively. Realigning again model and participants' performances on these rejection and confirmation events reveals that (Fig. 3, right), when rejection events occurred, both model and participants' correct responses abruptly increased and exploratory responses abruptly dropped off; when confirmation events occurred, by contrast, correct and exploratory responses exhibited no abrupt changes and, as

expected, gradually increased and decreased, respectively (more results in supplementary text).

Brain activations associated with reasoning computations

We then investigated whether fMRI activations confirm the implementation of the proposed algorithm in the PFC. To identify activations associated with inferring strategies' absolute reliability, we considered three reliability variables derived from the best-fitting model: actor and first- and second-alternative reliability. We entered these variables orthogonalized in that order in a unique regression analysis, which also included algorithmic events switch-in, rejection, and confirmation as regressors, along with those modeling exploration and exploitation trials (9). The regression factored out possible confounding variables including reward expectations, outcome predictions, and feedback values. We identified activations using significance thresholds set to $P = 0.05$ (familywise error corrected for multiple comparisons over the frontal lobes), and post hoc analyzes removed selection biases (22).

Strategies' reliability correlated with anterior PFC activations. Actor reliability correlated with ventromedial PFC (vmPFC) and perigenual anterior cingulate (pgACC) cortex activations, whereas right frontopolar cortex (FPC) activations correlated concurrently with both first- and second-alternative reliability (Fig. 4). No other regions exhibited such correlations ($P > 0.01$, uncorrected). vmPFC and pgACC activations that increased with actor reliability further decreased with first- and, more strongly, with second-alternative reliability, whereas right FPC activations decreased with actor reliability while increasing with first- and, more strongly, with second-alternative reliability (Fig. 4). The symmetrical, left FPC region marginally exhibited the same activation pattern as the right FPC (actor and first- and second-alternative reliability: all $T > 1.99$, $P < 0.053$). Accordingly, the less a strategy was eligible as actor, the more its reliability elicited FPC detrimentally to vmPFC and/or pgACC activations. vmPFC-pgACC and left FPC activations were also associated with feedback values ($T > 2.43$, $P < 0.0195$), from which strategies' reliability is inferred.

Using the same regression analysis, we next examined activations in switch-in, rejection, and confirmation events associated with hypothesis testing. These algorithmic events elicited more posterior PFC activations. Medially, the dorsal ACC (dACC) responded selectively to switch-in events (Fig. 5A). Switch-in events elicited larger dACC responses than exploitation and exploration trials ($T < 2.02$, $P > 0.05$). Confirmation events elicited only marginal dACC responses ($T = 2.32$, $P = 0.03$). Laterally, the left PFC [Brodmann's area (BA) 45, middle lateral prefrontal cortex (mid-LPC)] responded selectively to rejection events (Fig. 5B). Rejection events

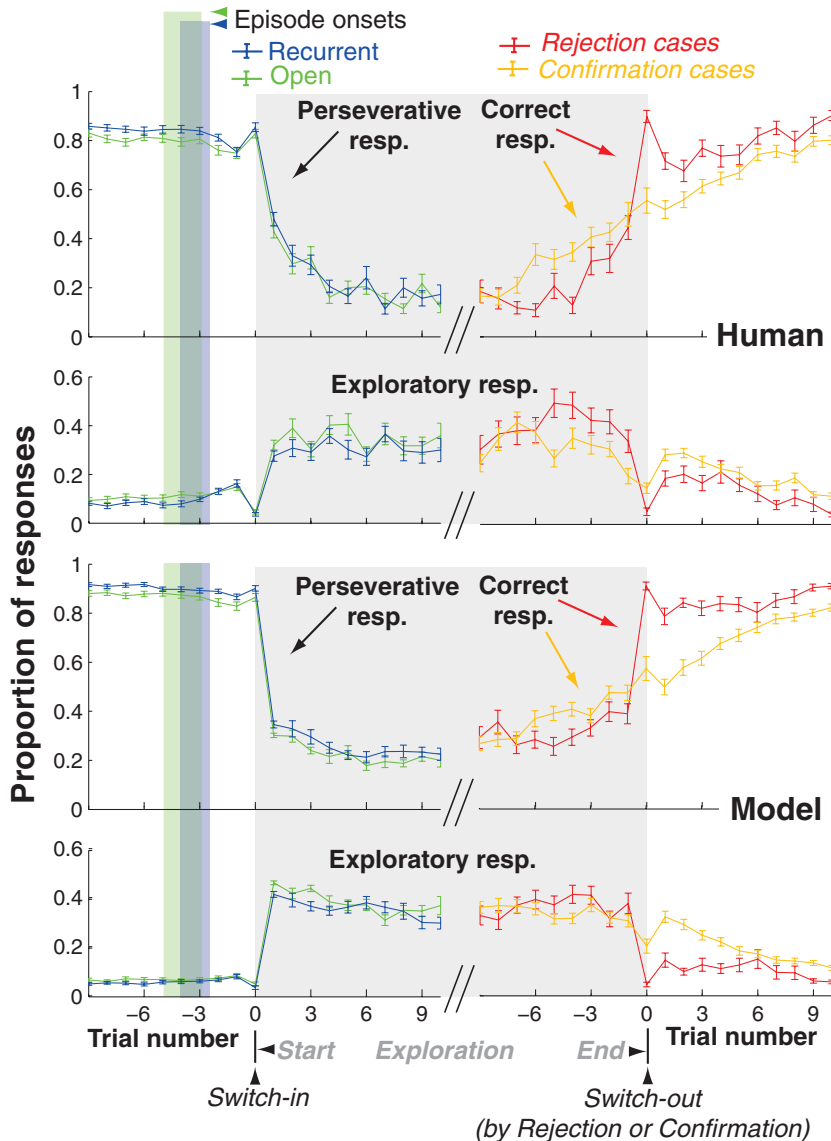


Fig. 3. Behavioral performances according to predicted algorithmic transitions. Model predictions (bottom) and participants' performances (top) realigned on switch-in (left), rejection (right, red) and confirmation (right, orange) events occurring in the algorithm. Data points following switch-in events and preceding rejection and confirmation events included only exploration trials. Model predictions are computed in every trial given participants' responses in previous trials. Green, open episodes; blue, recurrent episodes. Green and blue shaded areas are centered on the average of episode onsets preceding switch-in events (width: standard deviations). Perseverative responses correspond to correct responses in preceding episodes. Error bars are SEM across participants.

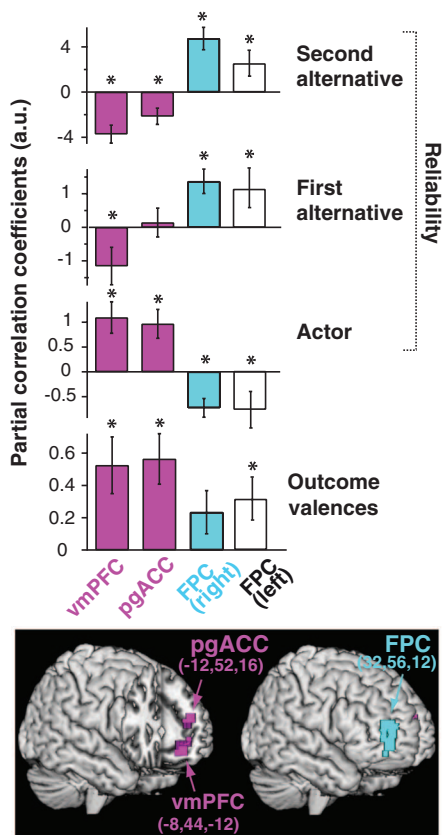


Fig. 4. Brain activations associated with reliability inferences. (Bottom) 3D rendering of all brain activations correlating with actor reliability (magenta) and with first- and second-alternative reliability (cyan) (thresholded at $P < 0.005$ (voxel-wise, uncorrected) and $P < 0.05$ (cluster-wise) for display purpose). Montreal Neurological Institute (MNI) coordinates of activation peaks are shown in brackets. (Top) Partial correlation coefficients for feedback valence, actor, and first- and second-alternative reliability averaged over activation clusters; a.u., arbitrary units. White bars are for the region symmetrical to right FPC activations (left FPC). Error bars are SEM across participants. $*P < 0.05$.

elicited larger mid-LPC activations than exploitation and exploration trials (both $T > 4.53$, $P < 0.00006$) and than switch-in and confirmation events (joint effect: $T = 2.38$, $P = 0.022$). The latter events elicited no significant mid-LPC responses (both $T < 1.69$, $P > 0.10$). Both the dACC and mid-LPC exhibited no differential responses between exploitation and exploration trials ($T_s < 1$) (Fig. 5, A and B) and no responses in the trials immediately preceding and following switch-in and rejection events (Fig. 6). Thus, dACC and mid-LPC responses to switch-in and rejection events, respectively, reflected the algorithmic transitions rather than the differential production of perseverative, exploratory versus correct responses and associated cognitive states around these events. Furthermore, as both switch-in and rejection events involve actor switching based on the same reliability threshold ($= \frac{1}{2}$), these differential activations could not simply reflect choice uncertainty

and general inhibition or selection mechanisms across monitored strategies. Instead, these results indicate that the dACC detects when actors monitored in the pgACC-vmPFC become unreliable for triggering the creation of probe actors, whereas the mid-LPC detects when counterfactual strategies monitored in the FPC become reliable for retrieving them as actor.

Only the ventral striatum responded selectively to confirmation events (Fig. 5C). Confirmation events elicited larger ventral-striatal activations than exploitation and exploration trials (both $T > 3.59$, $P < 0.0009$) and than switch-in and rejection events (both $T > 2.99$, $P < 0.005$). There were no significant ventral-striatal responses to switch-in and rejection events compared with exploitation and exploration trials (all $T < 1.99$, $P > 0.06$) nor differential ventral-striatal responses between exploitation and exploration trials ($T = 1.11$, $P = 0.27$), nor significant ventral-striatal responses in the trials immediately preceding and following confirmation events (Fig. 6). The region concurrently responded to reward prediction errors: Ventral striatal activations correlated both positively with feedback rewarding values ($T = 5.04$, $P = 0.00002$) and negatively with reward expectations ($T = 4.25$, $P = 0.00013$). Thus, beyond its involvement in actor reinforcement learning over trials (16), the ventral striatum exhibited additional responses in confirmation events. Because the vmPFC-pgACC projects to the ventral striatum (23) and encoded actor reliability, the evidence is that the ventral striatum detects when newly created strategies driving behavior become reliable, presumably for confirming their storage in long-term memory.

The dorsal striatum responded selectively to switch-in events (fig. S3), whereas bilateral posterior PFC (BA 44, post-LPC) and left premotor regions responded to both switch-in and confirmation events (fig. S4). These activations accord with the involvement of posterior frontal-striatal circuits in forming and storing action sets (18): Dorsal- and ventral-striatal responses correlated with premotor and post-LPC responses in switch-in and confirmation events, respectively, when the algorithm created and confirmed probe actors in long-term memory (fig. S5). We found no other frontal and basal responses ($P > 0.05$, uncorrected) except bilateral responses to switch-in events in FPC regions reported above (fig. S3), which likely reflected that, concomitant to probe actor creation, the former actor registers as an additional counterfactual strategy in the inferential buffer (more results in supplementary text).

Prefrontal foundations of human reasoning

The predicted algorithmic transitions associated with hypothesis testing and accounting for participants' behavior occurred within the frontal lobes in the expected PFC and striatal regions. Moreover, the anterior PFC encoded the predicted absolute reliability signals associated with the concurrent behavioral strategies the algorithm creates, learns, tests, and retrieves for driving action. These results support the hypothesis that the proposed algorithm describes reasoning PFC processes guiding adaptive behavior

(supplementary text). Accordingly, the frontal lobes implement two concurrent inferential tracks. First, a medial track comprising the vmPFC-pgACC, dACC, and ventral striatum makes inferences about the actor strategy that, through reinforcement learning, selects and learns the actions maximizing reward. Whereas the vmPFC-pgACC infers the actor's absolute reliability, the dACC detects when it becomes unreliable for triggering exploration—i.e., the formation of a new strategy from long-term memory to serve as actor. The ventral striatum then detects when this new actor strategy becomes reliable, which terminates exploration and confirms it in long-term memory. Second, a lateral track comprising the FPC and mid-LPC makes inferences about two or three alternative strategies stored in long-term memory. Whereas the FPC concurrently infers the absolute reliability of these counterfactual strategies from action outcomes, the mid-LPC detects when one becomes reliable for retrieving it as actor.

This medial-lateral segregation stems from the model core notion of absolute reliability, which yields to distinguishing between switching away from ongoing behavior (the actor becomes unreliable) versus switching to another behavioral strategy stored in long-term memory (one counterfactual strategy becomes reliable). In this protocol, the two events never coincided, which would have required alternating between only two recurrent situations associated with two distinct strategies (the actor unreliability then implies the reliability of the alternative strategy) (24). The dACC thus triggers switching away from ongoing behavior with the formation of new behavioral strategies, whereas the mid-LPC enables the switch to counterfactual strategies. The model may thus explain dACC activations observed in detecting unexpected action outcomes (25), switching to exploratory behaviors (26) and starting new behavioral tasks (27), and LPC activations in retrieving task sets (15, 28). Consistent with the model prediction, moreover, the dACC and mid-LPC coactivate when participants switch back and forth between only two alternative behaviors (17).

The model further indicates that the coupling between the medial and lateral track realizes hypothesis testing bearing upon new behavioral strategies created from long-term memory. Serving as a probe actor initially set as being unreliable, newly created strategies are disbanded when the mid-LPC detects that one counterfactual strategy has become reliable for retrieving it as actor. However, the ventral striatum adjusts probe actors to external contingencies through reinforcement learning (16, 19, 20) and detects when probe actors eventually become reliable. In that event, the ventral striatum confirms probe actors in long-term memory as additional, subsequently recoverable, strategies. The interplay between the dACC, mid-LPC, and ventral striatum thus controls switches in and out of exploration periods corresponding to hypothesis testing of newly created strategies. Accordingly, every decision to create new strategies may be subsequently revised according to new

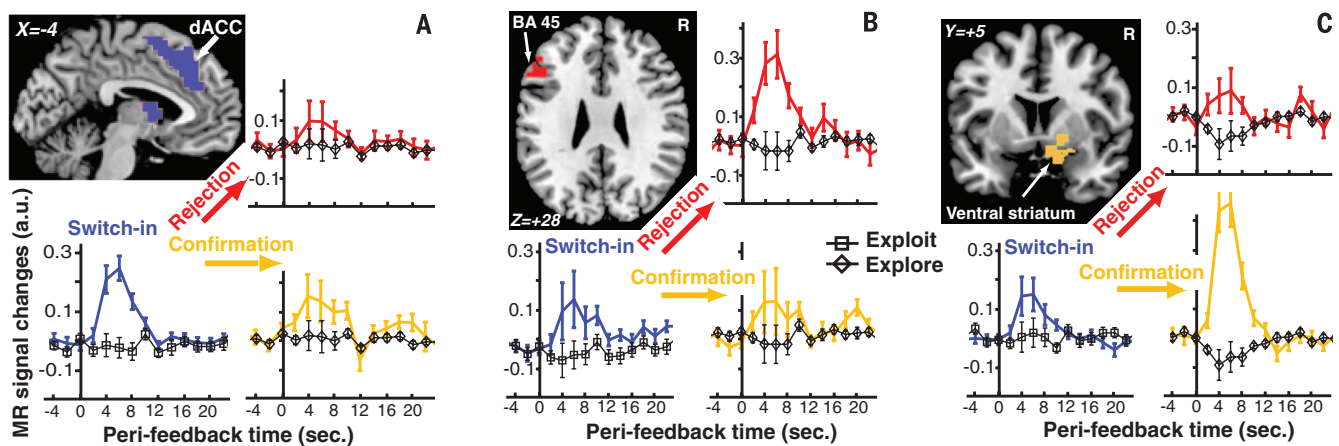
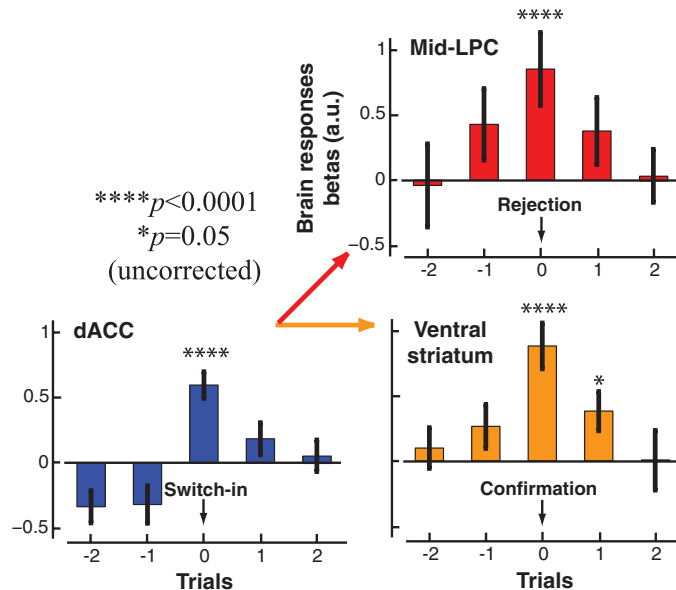


Fig. 5. Prefrontal and basal responses to predicted algorithmic transitions. On the brain slices, activations in switch-in (blue), rejection (red), and confirmation (orange) events superimposed on anatomical templates [thresholded at $P < 0.005$ (voxel-wise, uncorrected) and $P < 0.05$ (cluster-wise) for display purposes]. (A to C) X, Y, and Z are slice MNI coordinates corresponding to activation peaks (table S2). Graphs show perifeedback magnetic resonance responses to switch-in, rejection, and confirmation events averaged over activation clusters and factoring out all other effects. Black lines are perifeedback MR responses in exploitation (square) and exploration (diamond) trials. Error bars are SEM across participants.

Fig. 6. Prefrontal and striatal responses around algorithmic transitions.

Magnetic resonance responses to feedbacks in dACC, mid-LPC, and ventral striatum on trials preceding and following switch-in, rejection, and confirmation events. Bars are partial correlation coefficients (betas) from the regression analysis (a.u., arbitrary units) described in the text and corresponding to event-related regressors modeling switch-in, rejection, and confirmation events shifted 0, 1, or 2 trials preceding and following actual occurrences of these events. Error bars are SEM across subjects. Maximal and significant responses (when corrected for multiple comparisons around algorithmic events) were elicited only when the events occurred in the algorithm.



information, which is critical in optimal adaptive processes operating in open-ended environments for dealing with the intrinsic nonparametric nature of strategy creation (4).

Hypothesis testing derives from inferences about the absolute reliability of actor and two or three counterfactual strategies, which involved the vmPFC-pgACC and FPC, respectively. The dissociation supports the distinction between the notion of actor and a counterfactual strategy and accords with the vmPFC-pgACC and FPC involvement in monitoring ongoing and unchosen courses of action, respectively (3, 11, 12, 29, 30). Strategy absolute reliability measures to which extent the strategy is applicable to the current situation—i.e., current external contingencies and those learned

by the strategy result from the same latent cause. The vmPFC-pgACC thus infers to which extent the latent cause determining current action outcomes remains unchanged. The FPC infers to which extent the latter result from two or three previously identified latent causes. Latent causes are abstract constructs resulting from hypothesis testing implemented through the interplay between the dACC, mid-LPC, and ventral striatum. Latent causes organize long-term memory as a repertoire of behavioral strategies treated as separable entities. By detecting the reliability or unreliability of monitored strategies, the dACC, mid-LPC, and ventral striatum then appear to implement true or false exclusive judgments about possible causes of observed contingencies for selecting

appropriate behavioral strategies. The model thus describes how the PFC forms a unified inferential system subserving reasoning in the service of adaptive behavior. Among the prefrontal regions, the FPC is likely specific to humans (31, 32), which suggests that the ability to jointly infer multiple possible causes of observed contingencies and, consequently, to test new causal hypotheses emerging from long-term memory is unique to humans.

REFERENCES AND NOTES

1. M. Oaksford, N. Chater, *Behav. Brain Sci.* **32**, 69–84; discussion 85–120 (2009).
2. T. E. Behrens, M. W. Woolrich, M. E. Walton, M. F. Rushworth, *Nat. Neurosci.* **10**, 1214–1221 (2007).
3. E. D. Boorman, T. E. Behrens, M. W. Woolrich, M. F. Rushworth, *Neuron* **62**, 733–743 (2009).
4. Y. W. Teh, M. I. Jordan, M. J. Beal, D. M. Blei, *J. Am. Stat. Assoc.* **101**, 1566–1581 (2006).
5. F. Doshi-Velez, *Adv. Neural Inf. Process. Syst.* **21**, 477–485 (2009).
6. N. D. Daw, A. Courville, *Adv. Neural Inf. Process. Syst.* **20**, 369–376 (2007).
7. S. J. Gershman, D. M. Blei, Y. Niv, *Psychol. Rev.* **117**, 197–209 (2010).
8. A. Collins, E. Koehlin, *PLoS Biol.* **10**, e1001293 (2012).
9. Materials and methods are available online as supplementary material on Science Online.
10. N. Cowan, in *Human Learning and Memory*, C. Izawa, N. Ohta, Eds. (Erlbaum, Mahwah, NJ, 2005), pp. 155–175.
11. A. N. Hampton, P. Bossaerts, J. P. O'Doherty, *J. Neurosci.* **26**, 8360–8367 (2006).
12. E. Koehlin, A. Hyafil, *Science* **318**, 594–598 (2007).
13. E. D. Boorman, T. E. Behrens, M. F. Rushworth, *PLoS Biol.* **9**, e1001093 (2011).
14. E. Koehlin, C. Ody, F. Kounieher, *Science* **302**, 1181–1185 (2003).
15. K. Sakai, R. E. Passingham, *Nat. Neurosci.* **6**, 75–81 (2003).
16. J. O'Doherty *et al.*, *Science* **304**, 452–454 (2004).
17. E. Koehlin, C. Summerfield, *Trends Cogn. Sci.* **11**, 229–235 (2007).
18. D. Badre, A. S. Kayser, M. D'Esposito, *Neuron* **66**, 315–326 (2010).
19. W. Schultz, P. Dayan, P. R. Montague, *Science* **275**, 1593–1599 (1997).
20. K. Doya, *HFSP J.* **1**, 30–40 (2007).
21. J. J. Ribas-Fernandes *et al.*, *Neuron* **71**, 370–379 (2011).
22. M. Esterman, B. J. Tamber-Rosenau, Y. C. Chiu, S. Yantis, *Neuroimage* **50**, 572–576 (2010).

23. G. E. Alexander, M. R. DeLong, P. L. Strick, *Annu. Rev. Neurosci.* **9**, 357–381 (1986).
24. S. Charron, E. Koechlin, *Science* **328**, 360–363 (2010).
25. W. H. Alexander, J. W. Brown, *Nat. Neurosci.* **14**, 1338–1344 (2011).
26. N. Kolling, T. E. Behrens, R. B. Mars, M. F. Rushworth, *Science* **336**, 95–98 (2012).
27. N. U. Dosenbach *et al.*, *Neuron* **50**, 799–812 (2006).
28. K. Sakai, *Annu. Rev. Neurosci.* **31**, 219–245 (2008).
29. B. De Martino, S. M. Fleming, N. Garrett, R. J. Dolan, *Nat. Neurosci.* **16**, 105–110 (2013).
30. E. Koechlin, G. Basso, P. Pietrini, S. Panzer, J. Grafman, *Nature* **399**, 148–151 (1999).
31. K. Tefter, K. Semendeferi, *Prog. Brain Res.* **195**, 191–218 (2012).
32. F. X. Neubert, R. B. Mars, A. G. Thomas, J. Sallet, M. F. Rushworth, *Neuron* **81**, 700–713 (2014).

ACKNOWLEDGMENTS

We thank C. Summerfield and S. Palminteri for their helpful comments. Funded by a European Research Council Grant (ERC-2009-AdG#250106) to E.K. MRI data are available at <http://central.xnat.org>, project ID: PROBE.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/344/6191/1481/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S5
Tables S1 and S2
References (33–36)

14 February 2014; accepted 15 May 2014
Published online 29 May 2014;
10.1126/science.1252254

REPORTS

QUANTUM METROLOGY

Optically measuring force near the standard quantum limit

Sydney Schreppler,^{1*} Nicolas Spethmann,¹ Nathan Brahms,^{1,†} Thierry Botter,^{1,‡} Maryrose Barrios,^{1,§} Dan M. Stamper-Kurn^{1,2}

The Heisenberg uncertainty principle sets a lower bound on the noise in a force measurement based on continuously detecting a mechanical oscillator's position. This bound, the standard quantum limit, can be reached when the oscillator subjected to the force is unperturbed by its environment and when measurement imprecision from photon shot noise is balanced against disturbance from measurement back-action. We applied an external force to the center-of-mass motion of an ultracold atom cloud in a high-finesse optical cavity and measured the resulting motion optically. When the driving force is resonant with the cloud's oscillation frequency, we achieve a sensitivity that is a factor of 4 above the standard quantum limit and consistent with theoretical predictions given the atoms' residual thermal disturbance and the photodetection quantum efficiency.

Several decades ago, mounting efforts to detect directly the gravitational radiation produced by distant astrophysical sources prompted investigation of the detection limits imposed by the quantum-mechanical properties of the sensors (1–3). Early gravitational wave detectors (4) and subsequent table-top force-measurement systems (5, 6) were dominated by thermal noise that masked the contributions of detection uncertainty from the measurements themselves. Additionally, technical sources of optical noise emerging at high measurement strength obscured the effect of measurement back-action in macroscopic systems subject to strong optical probing (6), delaying until recently the observation of such back-action noise (7, 8). A purely quantum limit to force-measurement sensitivity for a continuously probed mechan-

ical oscillator occurs when thermal noise has been reduced to the level of zero-point fluctuations. This standard quantum limit (SQL) is reached when noise introduced by the mea-

surement, which uses a coherent state of probe light, is carefully balanced with the statistical fluctuations of the measurement outcome. A wide array of platforms, having sizes spanning many orders of magnitude, have pushed ever closer to the SQL (9–12). Sensitivity to forces as small as $(390 \text{ yN})^2/\text{Hz}$ has been reported (13); however, measured noise levels are still six to eight orders of magnitude larger (that is, much less sensitive) than the SQL (14, 15).

One way to measure a small force is to apply it to a mechanical harmonic oscillator and then to detect the resulting motion by illuminating the oscillator with light and measuring the phase shift of the reflected beam. Measurement strength is enhanced by placing the oscillator within a resonant optical cavity (16). In such an optomechanical system, measurement strength can be characterized by the optomechanical cooperativity $C_{\text{om}} = 4ng_{\text{om}}^2/(\kappa\Gamma)$, where n is the number of probe photons within the cavity, g_{om} is the single-photon optomechanical coupling rate, Γ is the mechanical oscillator full-linewidth, and κ is the optical cavity half-linewidth. C_{om} can be understood as the rate at which one gathers information about the oscillator's motion relative to optical and mechanical decay rates. Quantization of the optical cavity's electromagnetic field (shot noise) sets a lower bound on the imprecision of an optical phase measurement, which varies inversely with C_{om} . In the regime

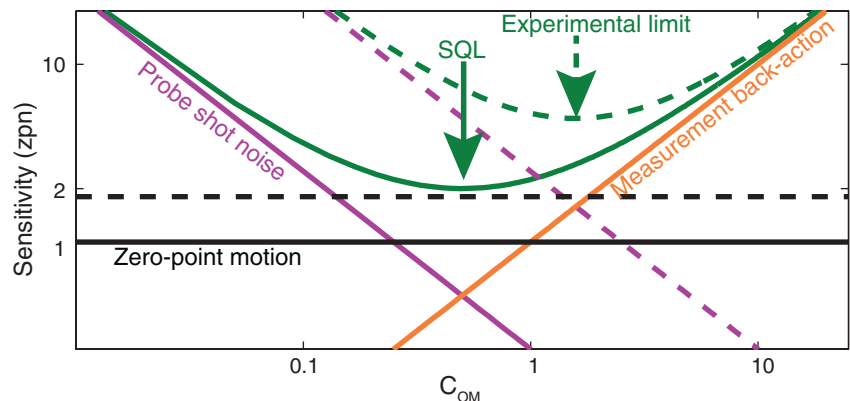


Fig. 1. Contributions to the SQL of force detection. Theoretical shot noise imprecision from probe light (purple lines) decreases and measurement back-action noise (orange line) increases as C_{om} rises. Together with the zero-point fluctuations (black lines), these sum to the total measurement sensitivity (green curves). Dashed lines show noise bounds corrected for imperfect detection efficiency and small nonzero phonon occupation. Sensitivity plotted in units of zero-point force noise ($\text{zpn} = 2\Gamma p_{\text{F}0}$, as in Eq. 1).

¹Department of Physics, University of California, Berkeley, CA 94720, USA. ²Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

*Corresponding author. E-mail: sschreppler@berkeley.edu

†Present address: Cisco Systems, Incorporated, San Francisco, CA 94158, USA.

‡Present address: Jet Propulsion Laboratory, Pasadena, CA 91109, USA.

§Present address: Department of Physics, Harvard University, Cambridge, MA 02138, USA.