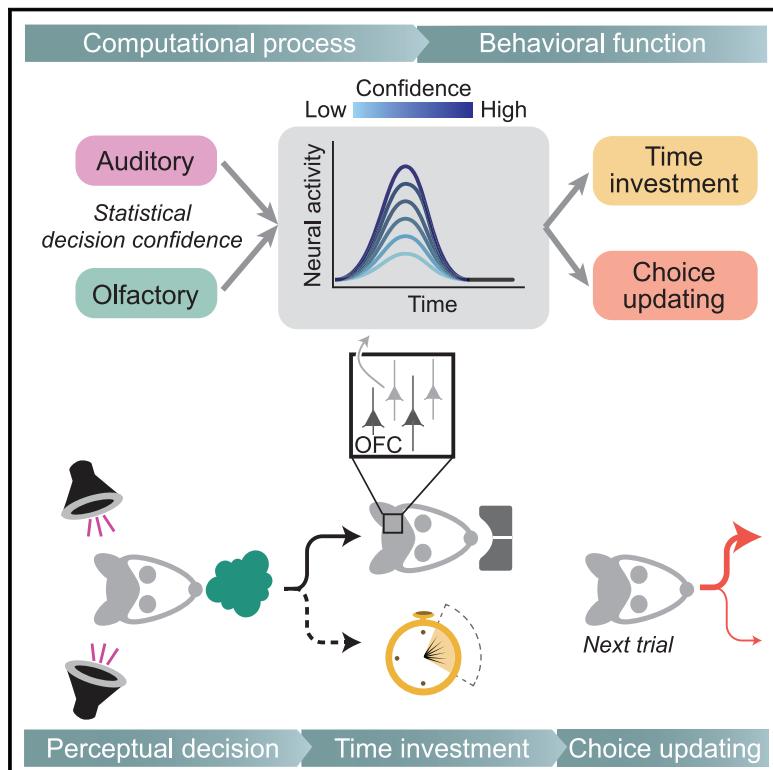


Behavior- and Modality-General Representation of Confidence in Orbitofrontal Cortex

Graphical Abstract



Authors

Paul Masset, Torben Ott, Armin Lak,
Junya Hirokawa, Adam Kepecs

Correspondence

akepecs@wustl.edu

In Brief

Neurons in the orbitofrontal cortex signal decision confidence across sensory modalities to support multiple confidence-guided behaviors, revealing the neural architecture for metacognitive abilities in rats.

Highlights

- Time investment by rats reflects confidence in olfactory and auditory decisions
- Orbitofrontal cortex neurons encode confidence irrespective of sensory modality
- Neurons predict confidence-guided time investment and cross-trial choice updating
- Orbitofrontal cortex contains an abstract representation of decision confidence



Article

Behavior- and Modality-General Representation of Confidence in Orbitofrontal Cortex

Paul Masset,^{1,2,5,7} Torben Ott,^{1,3,7} Armin Lak,⁴ Junya Hirokawa,^{1,6} and Adam Kepecs^{1,2,3,8,*}

¹Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

²Watson School of Biological Sciences, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

³Department of Neuroscience and Department of Psychiatry, Washington University in St. Louis, St. Louis, MO 63110, USA

⁴Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford OX1 3PT, UK

⁵Present address: Department of Molecular and Cellular Biology & Center for Brain Science, Harvard University, Cambridge, MA 02138, USA

⁶Present address: Graduate School of Brain Science, Doshisha University, Kyotanabe, Kyoto 610-0394, Japan

⁷These authors contributed equally

⁸Lead Contact

*Correspondence: akepecs@wustl.edu

<https://doi.org/10.1016/j.cell.2020.05.022>

SUMMARY

Every decision we make is accompanied by a sense of confidence about its likely outcome. This sense informs subsequent behavior, such as investing more—whether time, effort, or money—when reward is more certain. A neural representation of confidence should originate from a statistical computation and predict confidence-guided behavior. An additional requirement for confidence representations to support metacognition is abstraction: they should emerge irrespective of the source of information and inform multiple confidence-guided behaviors. It is unknown whether neural confidence signals meet these criteria. Here, we show that single orbitofrontal cortex neurons in rats encode statistical decision confidence irrespective of the sensory modality, olfactory or auditory, used to make a choice. The activity of these neurons also predicts two confidence-guided behaviors: trial-by-trial time investment and cross-trial choice strategy updating. Orbitofrontal cortex thus represents decision confidence consistent with a metacognitive process that is useful for mediating confidence-guided economic decisions.

INTRODUCTION

In the face of uncertainty, any organism that can perform statistical computations gains an adaptive advantage. An elementary statistical computation is confidence, estimating the probability that a chosen option is correct. An accurate estimate of confidence can confer benefits for a broad range of behaviors, from the essential to the mundane to the sophisticated: deciding between fight and flight, choosing which drink to order, or managing a stock portfolio. As humans, we assess our confidence in our percepts, memories, and decisions all the time, as part of a self-monitoring, reflective process termed metacognition, which is deeply personal and subjective (Dunlosky and Metcalfe, 2009). Consequently, confidence judgments have been studied as a central aspect of metacognition, and thought to require neural architectures unique to higher primates (Dehaene et al., 2017; Metcalfe and Son, 2013). Even though many organisms leverage probabilistic computations to guide behavior (Fetsch et al., 2013; Knill and Pouget, 2004; Kording and Wolpert, 2004; Pouget et al., 2003; Wystrach et al., 2015), it remains unknown what neural architectures support these (Meyniel et al., 2015a; Ott et al., 2018; Pouget et al., 2016).

What are the required properties for a neural representation of confidence? First, neural activity should reflect a confidence

computation; i.e., reflect the statistical likelihood that a proposition is correct, $P(\text{choice} = \text{correct} | \text{subjective evidence}, \text{choice})$ (“statistical decision confidence,” the type of confidence studied here, henceforth “confidence”) (Kepecs et al., 2008; Pouget et al., 2016). Second, neural activity should be correlated with decisions based on confidence; i.e., predict confidence-guided behavior. These two criteria follow from computational considerations, but psychological theories of metacognition provide additional desiderata. A prerequisite for metacognition is an abstract and centralized confidence representation: Abstract confidence representations should generalize both across the source of information used for probability computations and across different types of confidence-guided behaviors.

Previous studies in human and non-human animals have identified representations of confidence that fulfill a subset of these criteria (Fleming and Dolan, 2012; Hanks and Summerfield, 2017; Ott et al., 2018; Shadlen and Kiani, 2013). In animals asked to report their degree of confidence by opting out of a difficult task or placing a bet on their decision, neurons in frontal (Kepecs et al., 2008; Middlebrooks and Sommer, 2012) and parietal cortex (Kiani and Shadlen, 2009), as well as the pulvinar (Komura et al., 2013) and superior colliculus (Basso and Wurtz, 1997; Grimaldi et al., 2018), signal the uncertainty in perceptual decisions.



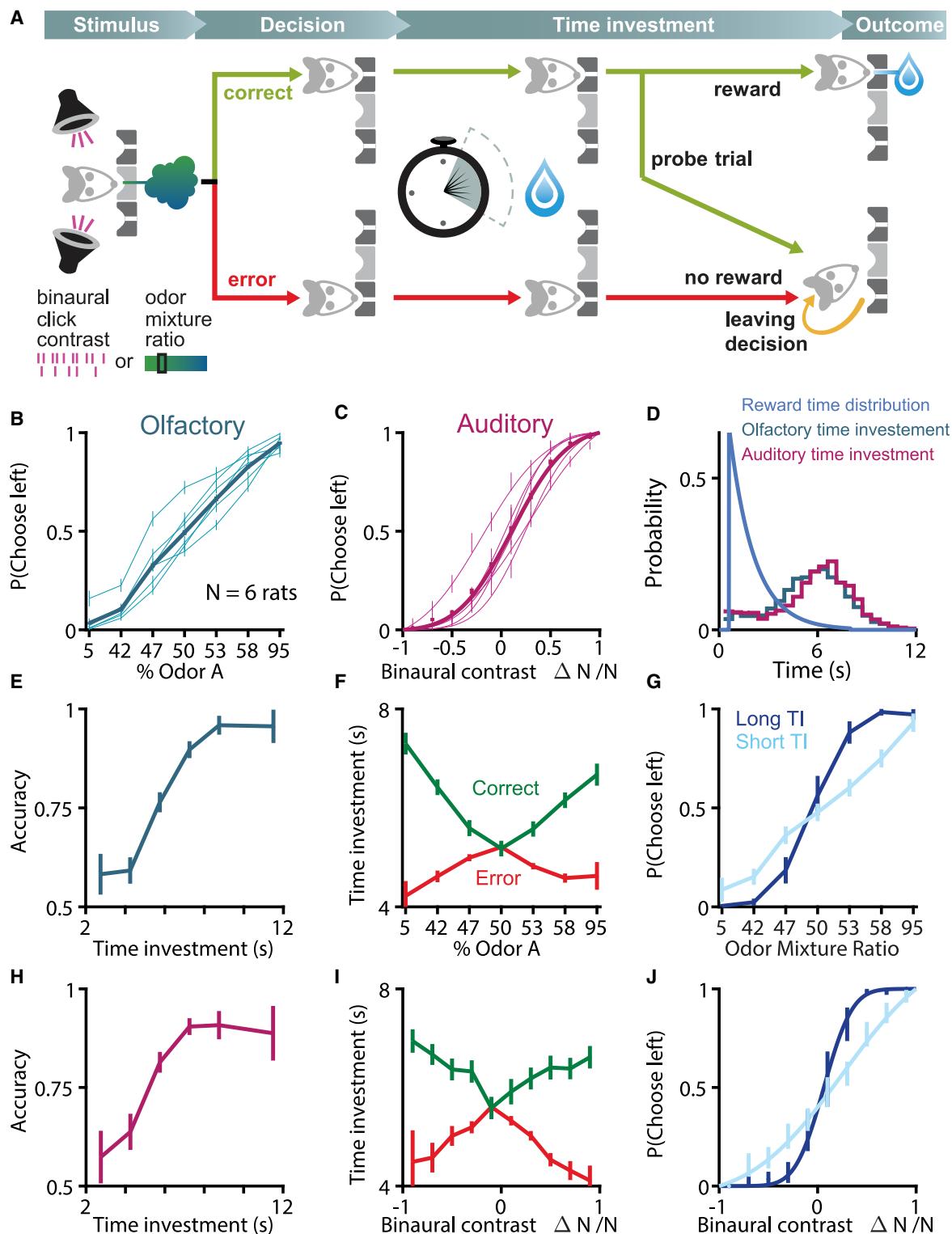


Figure 1. A Post-decision Time Investment Task to Measure Decision Confidence about Olfactory and Auditory Decisions

(A) Behavioral confidence reporting task design. Rats discriminated either an olfactory or an auditory stimulus in each trial (difficulty and sensory modality randomly selected) and indicated their choice by entering one of two side choice ports (left or right). Reward delivery was randomly delayed, and no feedback was

(legend continued on next page)

In humans, single neurons in the parietal and temporal lobes represent the accuracy of memory recalls (Rutishauser et al., 2015, 2018). Although these reports demonstrate that single neurons encode probability estimates about choice correctness, they probed confidence estimates with a single source of uncertain evidence (i.e., one sensory modality or one type of memory recall) and a single confidence-guided behavior (e.g., saccade-based opt-out). Therefore, these studies do not establish whether these neural representations have the properties of abstract confidence.

Here, we set out to test whether single neurons in orbitofrontal cortex (OFC) represent abstract decision confidence by addressing the four key properties for abstract confidence representations: (1) statistical computation of probability of being correct, (2) prediction of confidence-guided choices, (3) independence of the source of information, and (4) generalization across multiple confidence-guided behaviors. We developed a behavioral task in which we could measure choice and confidence reports for decisions based in two sensory modalities and evaluate the neural encoding of single neurons in OFC, a key area for value-based decisions (Hirokawa et al., 2019; Kepcs et al., 2008; Padoa-Schioppa and Conen, 2017; Stalnaker et al., 2015; Wallis, 2007). We trained rats to make perceptual choices based on either an ambiguous smell or sound and invited them to invest variable time for a potential reward after each decision. Post-decision time investment predicted choice accuracy and served as a behavioral report of confidence. We found that time investments reflected a statistically appropriate use of decision confidence. We show that single neurons in OFC encode confidence, predict rats' time investment across both modalities, and predict choice strategy updating across trials. Taken together, our results show that OFC neurons represent decision confidence generalized across sensory modality and behavior, consistent with the notion of metacognitive confidence as described in psychology (Dehaene et al., 2017; Dunlosky and Metcalfe, 2009).

RESULTS

A Behavioral Task to Study Decision Confidence across Sensory Modalities

We trained rats to perform auditory and olfactory perceptual decisions and incentivized them to invest their time in proportion to their decision confidence. Rats initiated a trial by entering the center port where they received a stimulus (Figure 1A). An olfactory or an auditory stimulus was presented in each trial, and trial modality and trial difficulty were randomly interleaved. In olfac-

tory trials, rats had to determine which of two stereoisomers was most prevalent in an odor mixture delivered in one of seven possible ratios (Uchida and Mainen, 2003). In auditory trials, rats had to determine which of two auditory click trains delivered binaurally had the greater number of clicks (Brunton et al., 2013; Sanders and Kepcs, 2012; see STAR Methods). For both modalities, rats performed close to perfect performance on the easiest trials (Figure 1B and 1C) and their accuracy varied with the strength of sensory evidence (average accuracy per session: $\text{Acc}_{\text{olf}} = 0.75 \pm 0.05$ and $\text{Acc}_{\text{aud}} = 0.81 \pm 0.04$).

Rats provided post-decision confidence reports by investing their time to earn an uncertain, delayed reward after making perceptual decisions (Lak et al., 2014). Water reward after correct choices was randomly delayed so that rats had to make an additional decision about how long to wait for potential reward at a choice port. This waiting time in anticipation of a reward is a time investment into their perceptual decision, which can reflect their subjective confidence. On each trial, the timing of the reward delivery was sampled from an exponential distribution with a time constant $\tau = 1.5$ s, offset by 600 ms and truncated at 8 s (Figure 1D, light blue line, mean delay = 2.1 s). On a subset of trials (10% probe trials) we withheld the reward to assess how much time rats would be willing to invest for correct choices. On these probe trials, rats waited on average 5.3 s after olfactory choices and 5.5 s for auditory choices (Figure 1D). This task design allowed us to simultaneously obtain a binary decision and a continuous measure of post-decision time investment in single trials.

Rats Report Decision Confidence about Two Sensory Modalities

We first show that rats' time investment in decisions based on both sensory modalities qualitatively reflects an appropriate use of statistical decision confidence. To operationalize the subjective notion of confidence, we used a statistical definition of decision confidence ("statistical decision confidence"), the probability of choice correctness, $P(\text{choice} = \text{correct} | \text{subjective evidence, choice})$ (Hangya et al., 2016; Meyniel et al., 2015a; Pouget et al., 2016). We tested whether time investment behavior followed three key signatures predicted by the statistical decision confidence model that describe the relationship between confidence, sensory evidence, and choice (Hangya et al., 2016).

The first signature, the calibration curve, describes how reported confidence predicts accuracy (Figures 1E and 1H): the measure of confidence, time investment in catch trials, should predict accuracy. The second signature is the vevaiometric

given in error trials and in 10% of correct trials (probe trials). We thus obtained the rat's decision and amount of time invested in their decision for each probe and error trial.

(B and C) Psychometric curves. Performance for individual rats (light curves) and pooled across animals ($n = 6$, thick curves) as a function of evidence strength for olfactory (B) and auditory (C) trials.

(D) Reward delay distribution (light blue curve) and time investment distribution for both sensory modalities pooled across all animals (probe trials only).

(E and F) Calibration curves for time investment. (E: olfactory trials; F: auditory trials; probe trials only; see STAR Methods).

(G and H) Vevaiometric curves for time investment. (G: olfactory trials; H: auditory trials; probe and error trials).

(I and J) Conditioned psychometric curves for time investment. (I: olfactory trials; J: auditory trials; probe trials only; see STAR Methods)

Error bars represent \pm SEM across trials (D–J, trials pooled across rats).

See also Figure S1.

curve (from the Greek βεβαίος, vevaios, meaning certain), which describes the average confidence report as a function of the strength of evidence and choice (Figures 1F and 1I): the measure of confidence should increase with evidence strength for correct trials (measured in correct probe trials) and decrease with evidence strength for error trials (measured in all error trials). The third signature, the conditioned psychometric curve, describes choice accuracy as a function of the strength of evidence for different confidence levels (Figures 1G and 1J, measured in catch trials): for a given stimulus discriminability, higher confidence trials are more accurate on average. Rats' time investment behavior followed all three signatures in both sensory modalities.

Isolating the Contribution of Statistical Decision Confidence to Time Investment Behavior

We show that rats' time investment quantitatively matches statistically appropriate use of confidence. How would an ideal observer make confidence-guided time investments in perceptual decisions? In sensory decisions, the strength of evidence is controlled by the discriminability of the stimulus (x_{olf} or x_{aud} , Figure 2A, upper left). However, a decision maker only perceives a “percept” \hat{x} , a noisy, subjective version of the provided evidence x (Figure 2A, lower left). We assumed that the decision maker decides based on its internal percept without additional noise ($f(\hat{x})$, Figure 2A, lower center left). Errors therefore arise when, due to noise, percepts fall into the incorrect category with respect to the provided sensory evidence (i.e., when x and \hat{x} are on different sides of the decision boundary). Although choices do not reveal perceptual noise to the observer on a trial-to-trial basis, average choice accuracy provides an estimate of the total perceptual noise (σ , inverse slopes of the psychometric functions, Figure 2A, upper middle). On the other hand, a decision maker has access to its perceived strength of evidence on each trial, which can be used to predict the average accuracy (i.e., probability of being correct) for this given percept. This requires learning a calibration function that maps percepts to average accuracy (Figure 2A, lower center right). Until this point, the computation of the estimated probability of correct, i.e., decision confidence, is specific to the type of sensory evidence. Once computed, the sensory-specific probability estimates might be combined into a single variable and represented in a modality-general manner. The decision-maker would then use its degree of confidence and map it to a time investment amount through a mapping function ($m(c)$, Figure 2A, lower right) to invest time appropriately (Figure 2A, upper right).

To estimate the perceptual noise σ , we fitted the psychometric function with a cumulative Gaussian distribution (Figure 2B, left; see STAR Methods). The standard deviation of this distribution provides an estimate of the perceptual noise (Figure 2B, center). Under the assumption of Gaussian perceptual noise, we inferred the distribution of confidence for a given strength of evidence and choice using the normative statistical model of confidence (Hangya et al., 2016). In other words, we generated percepts based on the observed perceptual noise (σ , Figure 2B, center) and calculated the degree of confidence, $P(\text{correct}|\hat{x}, \text{choice})$, associated with each percept. Next, we converted each confidence estimate to a time investment prediction by remapping

confidence c monotonically into a time investment Tl using a mapping function $Tl = m(c)$ by using the empirical cumulative time investment distribution (Figure 2B, right; see STAR Methods). This approach allowed us to isolate the contribution of a confidence computation that informs single-trial time investment behavior.

For all rats and both modalities, model-predicted optimal time investment matched the rats' behavior well (Figures 2C–2E and S1, solid lines for model predictions mostly fall within error bars of observed time investment behavior). Shuffling single-trial time investment data while retaining the same model parameters (perceptual noise σ and mapping function $m(c)$, i.e., single trials are not informed by decision confidence) produced uninformative relationships between time investments and task variables that were qualitatively distinct from observed time investment data (Figure S1, dashed lines). These results show that the rats' time investment reflected an appropriate use of decision confidence in two sensory modalities.

OFC Single Neurons Are Tuned to Decision Confidence for Both Sensory Modalities

How can we determine if a neuron's activity is informed by a statistical decision confidence computation? We searched for neural signatures of statistical decision confidence, shown in Figures 1E–1J and 2C–2E, which can be viewed as “confidence tuning curves” that constrain the possible underlying computational process (Fleming and Daw, 2017; Hangya et al., 2016). We recorded 1,593 well-isolated single neurons from six rats in OFC using custom-built 32- or 64-channel tetrode micro-drives (Figure S2; see STAR Methods) and analyzed a subset of 1,211 neurons with an average firing rate above 1 s^{-1} . We focused our initial analysis on the beginning 2 s of the waiting time period, after rats had entered the choice port and started to invest time into their decision (“anticipation period”). Figures 3A, S3A, and S3B show three example neurons that are tuned to statistical decision confidence across both sensory modalities.

Figure 3Ai shows a neuron whose activity increased after the entry into the choice port and was graded according to olfactory evidence supporting the choice. Activity was highest for easy correct choices, the trials with the highest average confidence, whereas activity was lowest for easy error choices, the trials with the lowest average confidence (cf. Figure 2C). We computed confidence tuning curves using the average firing rate in the anticipation period (see STAR Methods). Choice accuracy systematically increased with firing rate in olfactory trials (calibration curve, Figure 3Aii, left panel, cf. Figure 2C). Firing rate increased with evidence for correct trials and decreased with evidence for error trials (vevaiometric curve, Figure 3Aii, right panel, cf. Figure 2D). The same neuron also encoded decision confidence for auditory trials (Figures 3Aiii and 3Av). In two other neurons, we observed confidence encoding across sensory modalities, albeit with inverted tuning and different time courses; i.e., they were negatively tuned to confidence and therefore increased their firing rates with decreasing confidence (Figures S3A and S3Bi–iv).

To identify OFC neurons that encode decision confidence during the anticipation period, we computed an outcome prediction index (OPI, a normalized area under the receiver operating characteristic [auROC], see STAR Methods) as a proxy metric for decision

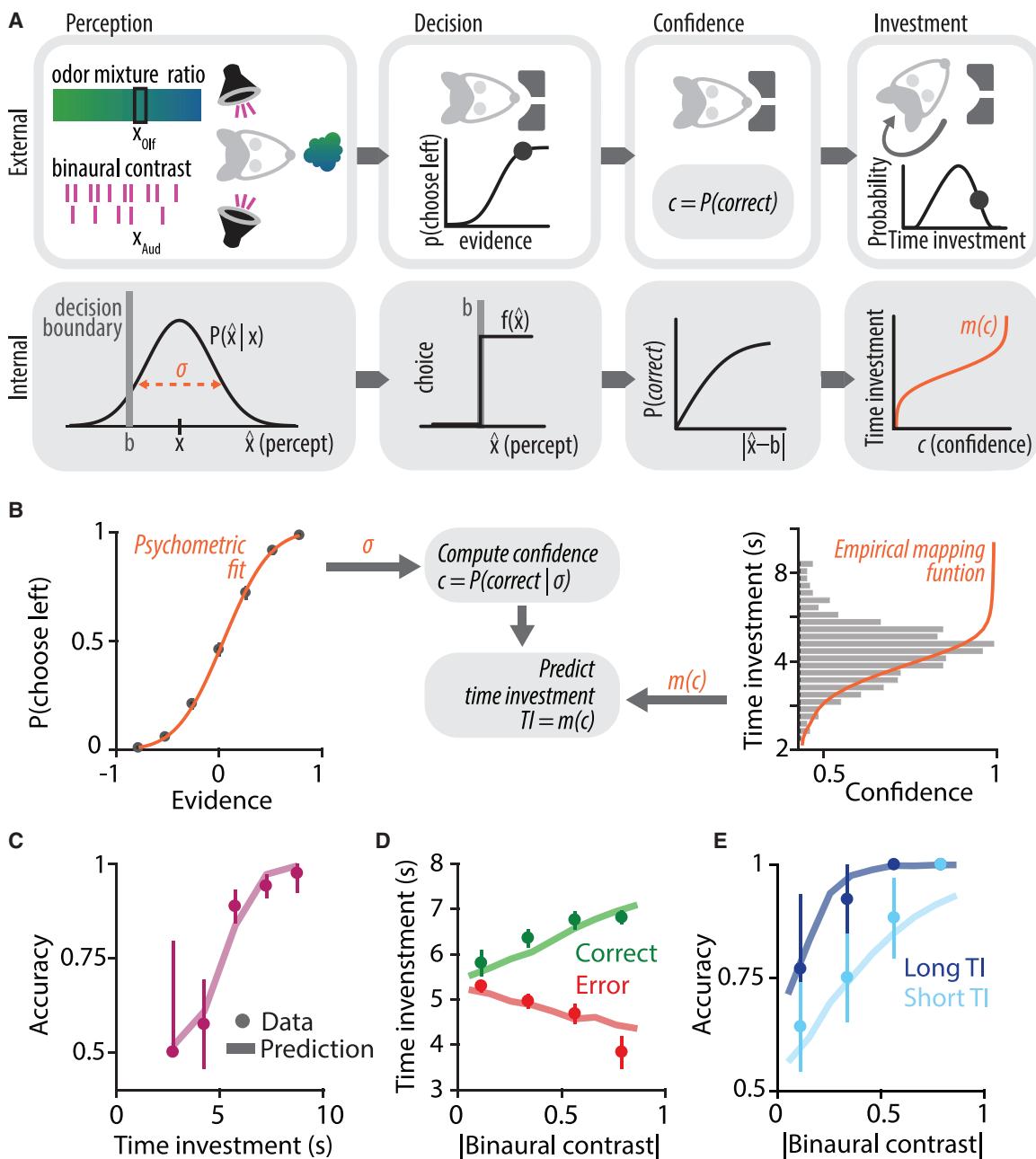


Figure 2. A Statistical Model Predicts the Optimal Time Investment Strategy

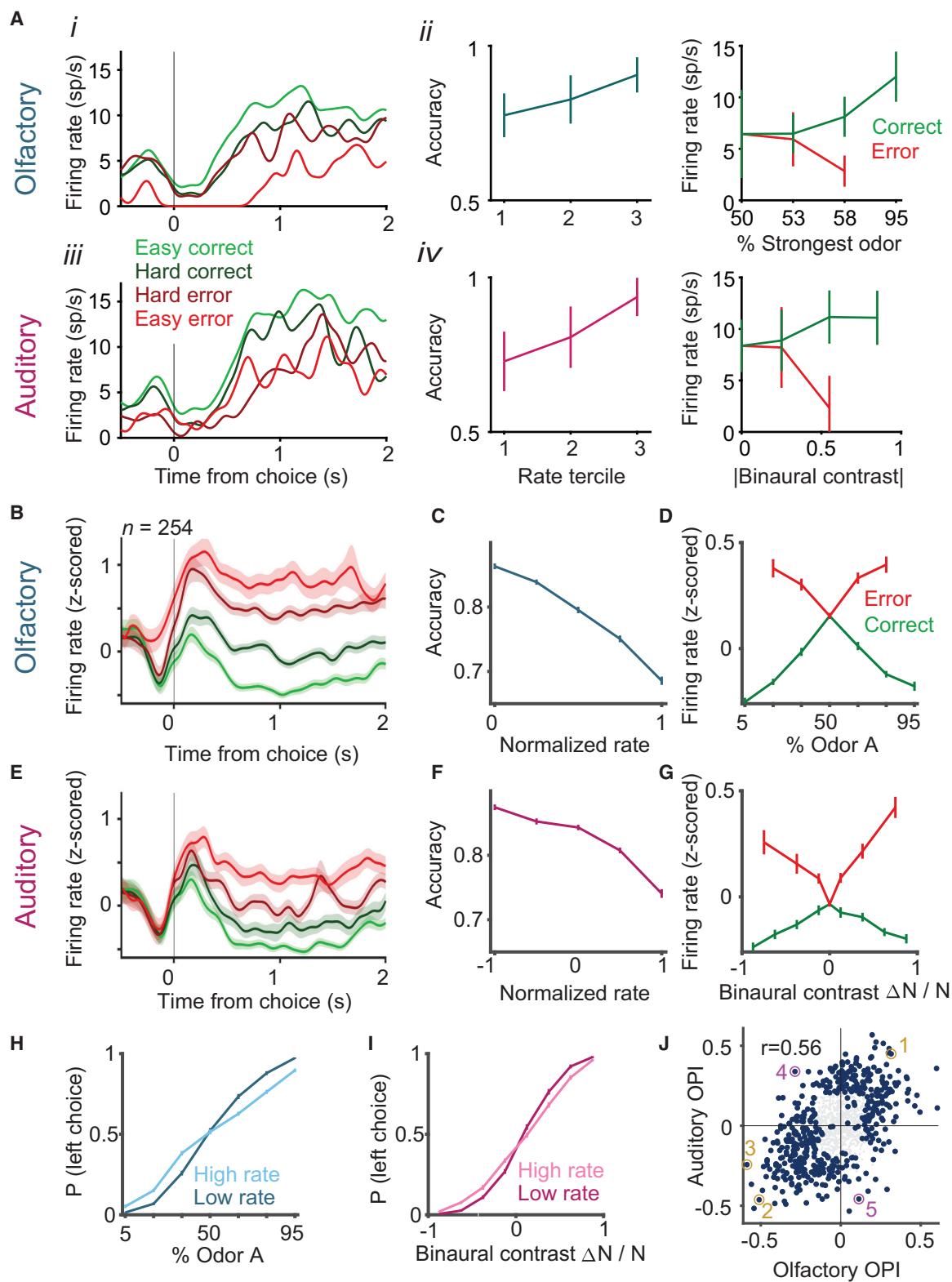
(A) A model for reporting choice and confidence (see text and STAR Methods). x , evidence (olf, olfactory; aud, auditory); b , decision boundary; \hat{x} , percept; σ , perceptual noise; $f(\hat{x})$, decision rule; c , confidence; $m(c)$, confidence mapping function.

(B) Estimating perceptual noise σ and confidence mapping function $m(c)$ (left, points, data; line, model fit) (right, histogram, time investment distributions; line, estimated mapping function $m(c)$).

(C–E) Rat's time investment behavior reflects appropriate use of decision confidence (single rat example, auditory trials) as shown by the calibration curve (C), vevaiometric curve (D), and conditioned psychometric curve (E) (cf. Figures 2E–2J). Dots, data; lines, model. Note that optimal model predictions fell within error bars of the rat's time investment behavior.

Error bars represent \pm SEM across trials.

See also Figure S1.



(legend on next page)

confidence. We identified outcome-predictive neurons with significant positive or negative OPI ($n_{\text{selective}} = 459$, $p < 0.025$ in either sensory modality using shuffling test). We split the population of outcome-predictive neurons into neurons with positive OPI (confidence(+), i.e., higher firing rate for correct trials, $n = 205$, sign of OPI measured using trials of both modalities) and negative OPI (confidence(-), i.e., higher firing rate for error trials, $n = 254$). For both sensory modalities, confidence(-) and confidence(+) neuronal populations showed a graded increase of activity at the beginning of the anticipation period according to choice and difficulty (Figures 3B, 3E, S4A, and S4B) and the confidence tuning curves in the anticipation period followed the signatures of decision confidence (Figures 3C, 3D, 3F–3I, and S4C–S4H).

To assess whether single neurons generalize outcome prediction across sensory modalities, we show that OPIs were strongly correlated across sensory modalities; i.e., outcome-predictive neurons for one modality tended to be also outcome-predictive in the other while preserving tuning direction (Figure 3J, $r = 0.56$, $p < 0.0001$, t test). Visual examination (Figures S3C and S3D) and quantitative tests of self-consistency suggest that the remaining variability of OPIs across sensory modalities could be constrained by our ability to estimate it (see STAR Methods). Together, these results reveal that single neurons in OFC encode decision confidence generalized across both sensory modalities tested.

Neural Activity in OFC Predicts Confidence-Guided Time Investments across Modalities

We next show that OFC neurons predict time investment, our behavioral report of confidence, in single trials. Figures 4Ai and 4Aii show that, for the same neuron shown in Figure 3A, the increase in activity after choice port entry in both modalities was proportional to the rat's eventual time investment. Its firing rate in the anticipation period predicted trial-by-trial time investment in both modalities (Figures 4Aii and 4Av, Pearson correlation $r_{\text{olf}}(59) = 0.42$, $p = 0.0008$, $r_{\text{aud}}(38) = 0.39$, $p = 0.012$). The firing rate in the anticipation period of the two neurons negatively tuned to decision confidence was negatively correlated with the time investment (Figure S3A and S3Bv–S3Bvii). We observed similar signatures for both the confidence(–) and confidence(+) populations (Figures 4B–4E and S5A–S5D, respectively).

We tested whether time investment predictions of single neurons were similar across sensory modalities. We computed the correlation between firing rates in the anticipation period and time investments independently for each neuron and modality. The prediction metrics were strongly correlated across both sensory modalities (Figure 4F, $r = 0.76$, $p < 0.0001$, t test). Thus, a neuron predicting time investment in one modality tended to also predict time investment in the other modality the same way.

We developed a measure for firing rate behavior co-variations to determine the contribution of single neurons to time investment beyond what can be predicted based on their tuning, conceptually similar to choice probability analysis (Britten et al., 1996). We first calculated each neuron's firing rate not explained by the strength of evidence for a choice by taking the residuals of a linear model expressing firing rate as a function of evidence strength and choice. We then calculated the correlation between residual firing rates and residual time investments across trials separately for each sensory modality (see STAR Methods). For olfactory trials, there was a strong correlation between the time investment metric computed using the residuals and the time investment metric computed using neutral evidence trials only (Figure 4G, $r = 0.64$, $p < 0.0001$ and $r = -0.04$, $p = 0.47$ for shuffled trials, t test). The time investment metrics based on residuals were strongly correlated across both modalities (Figure 4H, $r = 0.75$, $p < 0.0001$ and $r = -0.05$, $p = 0.34$ for shuffled trials, t test). Thus, firing rates of single neurons predict time investment on a trial-by-trial basis beyond the information provided by stimulus difficulty and choice.

We further investigated the similarity in neuronal time investment representations across sensory modalities using a cross-modal decoding approach. Using simultaneously recorded neurons, decoders trained to predict time investment in one sensory modality predicted time investments equally well for trials in the other, untrained, sensory modality. To assess decoding quality, we computed the correlation between predicted and observed time investments (10-fold cross-validation, see STAR Methods). Irrespective of the modality used to train the decoder, single-session decoding quality was similar for both sensory modalities ($p > 0.05$, Kolmogorov-Smirnov test, Bonferroni corrected), while they were statistically better than decoders

Figure 3. Orbitofrontal Neuronal Responses during Outcome Anticipation Reflect Statistical Confidence in both Olfactory and Auditory Decisions

- (A) Positively tuned confidence-encoding single neuron. (i) Average firing rate aligned to the time of choice port entry for olfactory trials grouped by evidence supporting the choice. Green, correct trials; red, error trials; brightness, evidence strength. (ii) Calibration curve (left) and vevaiometric curve (right) for average firing rate in the anticipation period (first 2 s after choice port entry) in olfactory trials. (cf. Figures 1E and 1F). (iii–iv) Same conventions as in (i–ii) for auditory trials. Error bars represent \pm SEM across trials.
- (B) Average population activity of confidence(–) neurons aligned to choice port entries for olfactory trials. Same conventions as in (A); shaded area, \pm SEM across neurons.
- (C) Population calibration curve for olfactory trials. Firing rates predict accuracy (cf. Figure 1E).
- (D) Population vevaiometric curve for olfactory trials (cf. Figure 1F). Green trace, correct trials; red trace, error trials.
- (E–G) Same conventions as in (B–D) for auditory trials.
- (H) Population conditioned psychometric curve for olfactory trials (cf. Figure 1G). Difference between low firing rate trials and high firing rate trials for %Odor A = 42, 47, 53, and 58, $p < 1e-7$, two-tailed t test.
- (I) Population conditioned psychometric curve for auditory trials. Difference between low firing rate trials and high firing rate for absolute value of binaural contrast in the bins 0.75–0.5, 0.5–0.25 $p < 0.001$, two-tailed t test.
- (J) Correlation between OPIs in olfactory trials and auditory trials (each point corresponds to one neuron). Neurons labeled 1–5 correspond to the neurons shown in (A) and Figures S4A–S4D. Blue dots: outcome-predictive neurons; gray labels: non-significant neurons; $p < 0.025$ using bootstrap. Error bars represent \pm SEM across neurons (or across trials in A).

See also Figures S2, S3 and S4.

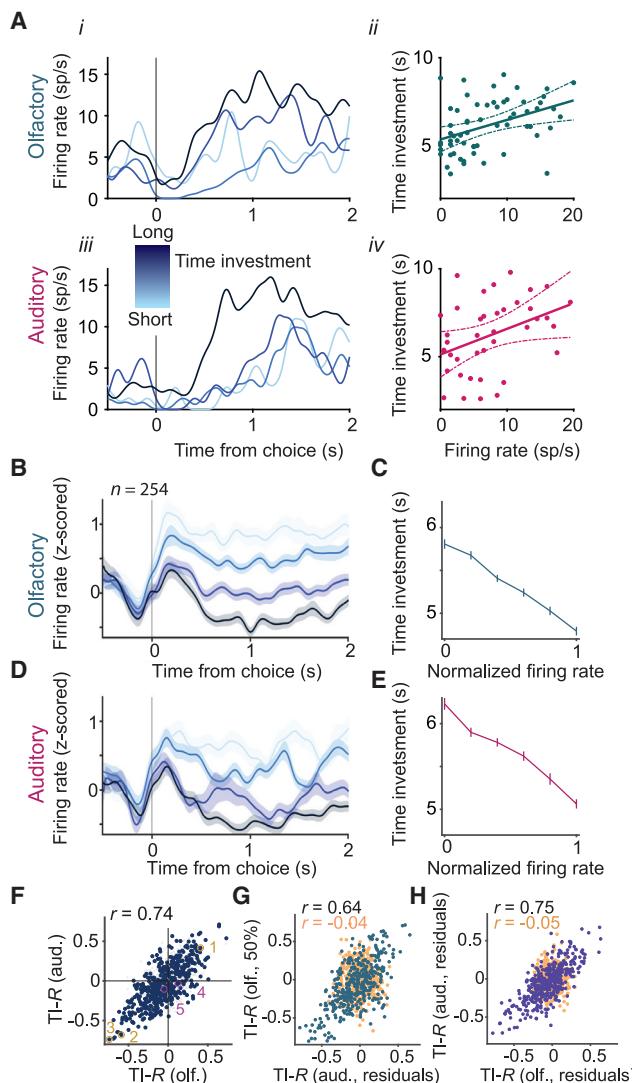


Figure 4. Orbitofrontal Neuronal Activity during Outcome Anticipation Predicts Time Investment across Both Sensory Modalities

(A) Same positively tuned confidence-encoding neuron as in Figure 3A. (i) Average firing rate aligned to choice port entry for olfactory trials grouped by the rat's eventual amount of time invested in the decision. Blue gradient, amount of time investment (see STAR Methods). (ii) The neuron's average activity in the anticipation period predicts the rat's waiting time in olfactory trials. Pearson correlation $r_{\text{olf}}(59) = 0.42$, $p = 0.0008$. (iii and iv) Same conventions as in (i) and (ii) for auditory trials. Pearson correlation $r_{\text{aud}}(38) = 0.39$, $p = 0.012$ (t test). Error lines represent 95% confidence intervals for linear model across trials.

(B) Average activity of confidence(+) population sorted by the rats' time investment in olfactory trials aligned to choice port entries. Blue gradient, time investment. Shaded area, SEM across neurons.

(C) Population activity predicts time investment in olfactory trials. Error bars, SEM across neurons.

(D) Average activity for negatively tuned outcome-predictive neurons in auditory trials (same convention as in B).

(E) Population activity predicts time investment in auditory trials (same layout as in C).

(F) Correlation between firing rate and time investment correlations in olfactory and auditory trials for outcome-predictive neurons (each point corresponds to

predicting shuffled trials (Figure 5A). Furthermore, single-session decoding quality was strongly correlated across sensory modalities (Figures 5B and 5C) and improved with the number of cells used (Figures S5E and S5F). Together, these analyses reveal that the firing pattern of OFC neurons is predictive of trial-by-trial time investment across both sensory modalities.

Neural Activity in OFC Predicts Confidence-Guided Choice Updating

So far, we have considered a specific behavioral report of confidence: time investment, which was incentivized by our task design. We also noticed that rats showed small systematic changes in perceptual decision strategy after rewarded trials. Reinforcement learning models that incorporate perceptual ambiguity predict similar choice strategy updating driven by confidence-scaled reward prediction errors (Drugowitsch et al., 2019; Lak et al., 2017, Lak et al., 2020b). These models predict that when decision confidence is low, but the choice is rewarded—a surprising outcome—there is a prediction error that produces a bias in the subsequent trial's choice (Figures 6A–6C and S6A–S6C; see STAR Methods).

We found that rats' choices were systematically modulated depending on the previous outcome and trial difficulty. If the previous trial was rewarded, rats tended to repeat the same choice again; i.e., they tended to choose the same side (left or right) again, which produced a shift in the psychometric function depending on the previous choice (Figures 6D, 6E, 6G, and 6H). Choice updating scaled with the evidence strength of the previous trial both within and across sensory modalities—albeit this effect was less pronounced across sensory modalities—and was therefore consistent with confidence-scaled updating of choice values (Figures 6F, 6I, S6D, and S6G for within modality choices, and Figures S6E, S6F, S6H, and S6I for choices across modalities; $p < 10^{-6}$ for evidence for choice effect on predicting choice updating; logistic regression, see STAR Methods). We ruled out that choice updating was due to slow drifts in choice bias: there was no systematic relationship between choice updating and the choice in the subsequent trial in either modality (Figure S6J).

This behavioral phenomenon enabled us to examine whether OFC neurons encoding decision confidence and predicting time investments also predict cross-trial choice strategy updating. We split trials according to the firing rate of the confidence(+) neurons by their median firing rate in each session and compared choice updating, i.e., the tendency to repeat the same choice, in the subsequent trial. When the activity of

one neuron). Neurons labeled 1–5 correspond to the neurons shown in A and Figures S3A–S3D. Blue dots: outcome-predictive neurons; gray labels: non-significant neurons; $p < 0.025$ using bootstrap.

(G) Correlation between neutral evidence time investment correlates and firing rate and time investment correlates in olfactory trials. Each point corresponds to one neuron. Dark green, data $r = 0.64$, $p < 0.0001$; pink, shuffled control $r = -0.04$, $p = 0.47$, Pearson correlation, t test.

(H) Correlation between firing rate and time investment correlates between olfactory and auditory trials. Each point corresponds to one neuron. Blue, data $r = 0.75$, $p < 0.0001$ and; pink, shuffled control $r = -0.05$, $p = 0.34$, Pearson correlation, t test.

See also Figures S2, S3 and S5.

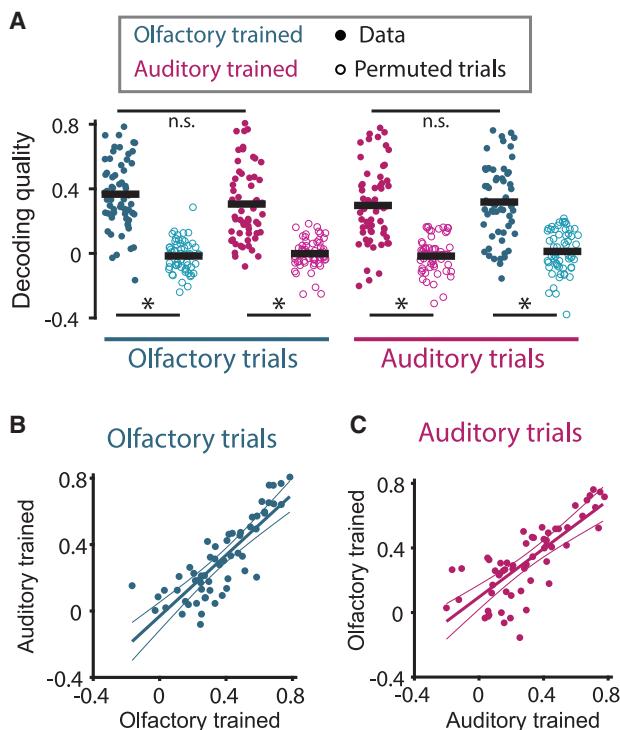


Figure 5. Sensory-Modality-General Decoding of Time Investment by OFC Neural Populations

(A) Decoding quality (measured as the correlation between predicted and observed time investment from ensemble of simultaneous recorded neurons). Each dot corresponds to one recording session. Empty circles, decoder trained on permuted data. Ten-fold cross validations, n.s., $p > 0.1$, * $p < 1e-10$ (two-sample Kolmogorov-Smirnov test, Bonferroni corrected p values). (B) For olfactory trials, decoding quality between the two decoders is strongly correlated across sessions (Spearman's $r(57) = 0.83$, $p < 1e-10$, t test). Error lines represent 95% confidence intervals for linear model fit. (C) Same as (B) for auditory trials (Spearman's $r(57) = 0.78$, $p < 1e-10$, t test). See also Figures S2 and S5.

confidence(+) neurons was high, we observed almost no choice updating in the subsequent trials, whereas when the activity of confidence(+) neurons was low, we observed strong choice updating in the subsequent trial (Figures 6J–6M). We observed this effect both within (Figures 6J and 6K, $p < 0.003$ and $p < 0.03$ for olfactory and auditory trials respectively, logistic regression) and across (Figures 6L and 6M, $p < 0.008$ auditory to olfactory and olfactory to auditory respectively, logistic regression) sensory modalities. Since the activity of confidence(+) neurons is modulated by the sensory evidence, we repeated this analysis by using a model that considered sensory evidence in addition to confidence-neuron firing rates in the previous trial, which revealed a significant additional contribution of firing rates to the prediction of choice updating ($\Delta AIC = -813 [-805]$ when omitting firing rate for predicting choice updating in olfactory [auditory] trials, $p < 0.001$, bootstrapping, see STAR Methods). Curiously, the confidence(−) neuronal population did not predict choice updating on average, suggesting that a specialized OFC subpopulation might be specifically involved in choice updating (Figure S6K).

OFC Neurons Share Activity Dynamics and Encoding across Sensory Modalities

Lastly, we assessed the time course and tuning similarities across sensory modalities throughout each trial. Firing rate dynamics during a trial were similar across sensory modalities both during the anticipation and the outcome periods, as assessed by peak averaged firing rate (Figures 7A–7D; see STAR Methods) and by the distribution of spike density correlations across modalities (Figure S7A, $p < 2e-16$ for choice period, and Figure S7B, $p < 1.41e-17$ for outcome period, Kolmogorov-Smirnov test). Similarly, the temporal dynamics of confidence tuning (Figures 7E and 7F) and time investment prediction (Figures 7G and 7H) were correlated across modalities (Figures S7C–S7E; see STAR Methods).

To further separate the dynamics of confidence encoding from the overall activity dynamics in OFC, we trained a generalized linear model (GLM) to predict the firing rate of single neurons across time based on several task variables such as *chosenside* and *evidenceforchoice* (see STAR Methods). For both olfactory and auditory trials, confidence encoding (captured by the coefficient $\beta_{EvidenceForChoice}$) arose at the beginning of the choice period and remained sustained during the anticipation period, unlike the coefficient for chosen side or trial outcome (Figures 7I and 7J). Coding similarity across modalities emerged late in the movement period and plateaued within the first 2 s of the time investment period (Figure 7K). Activity at the time of the outcome (reward or leaving) was best captured by the binary outcome variable rather than the graded evidence for choice (Figure 7L).

Taken together, these results show that both the overall activity dynamics, as well as confidence encoding and time investment predictions that emerge after the rats' choice, are shared across sensory modalities.

DISCUSSION

Here, we showed that neuronal representations of decision confidence in rat OFC satisfy characteristics expected for an abstract confidence signal that could mediate metacognition. First, neural activity reflected a statistical confidence computation and predicted behavioral reports of confidence. Further, these confidence signals generalized across the sensory modalities and predicted distinct confidence-guided behaviors: time investment and cross-trial choice updating.

Limitations and Challenges in Linking Confidence to Neural Activity

Our claim about confidence representations relies on several lines of evidence that allow us to distinguish abstract confidence representations from other signals. First, we showed that OFC firing rates during the time investment period reflect critical signatures of a statistical confidence computation (Hangya et al., 2016; Sanders et al., 2016). While statistical confidence computations normatively predict the observed signatures (Hangya et al., 2016), there may be other confidence-unrelated processes (e.g., attention, motivation) that could lead to similar patterns, although modeling studies suggest that they are unlikely to produce all three signatures of decision confidence (Sanders, 2014).

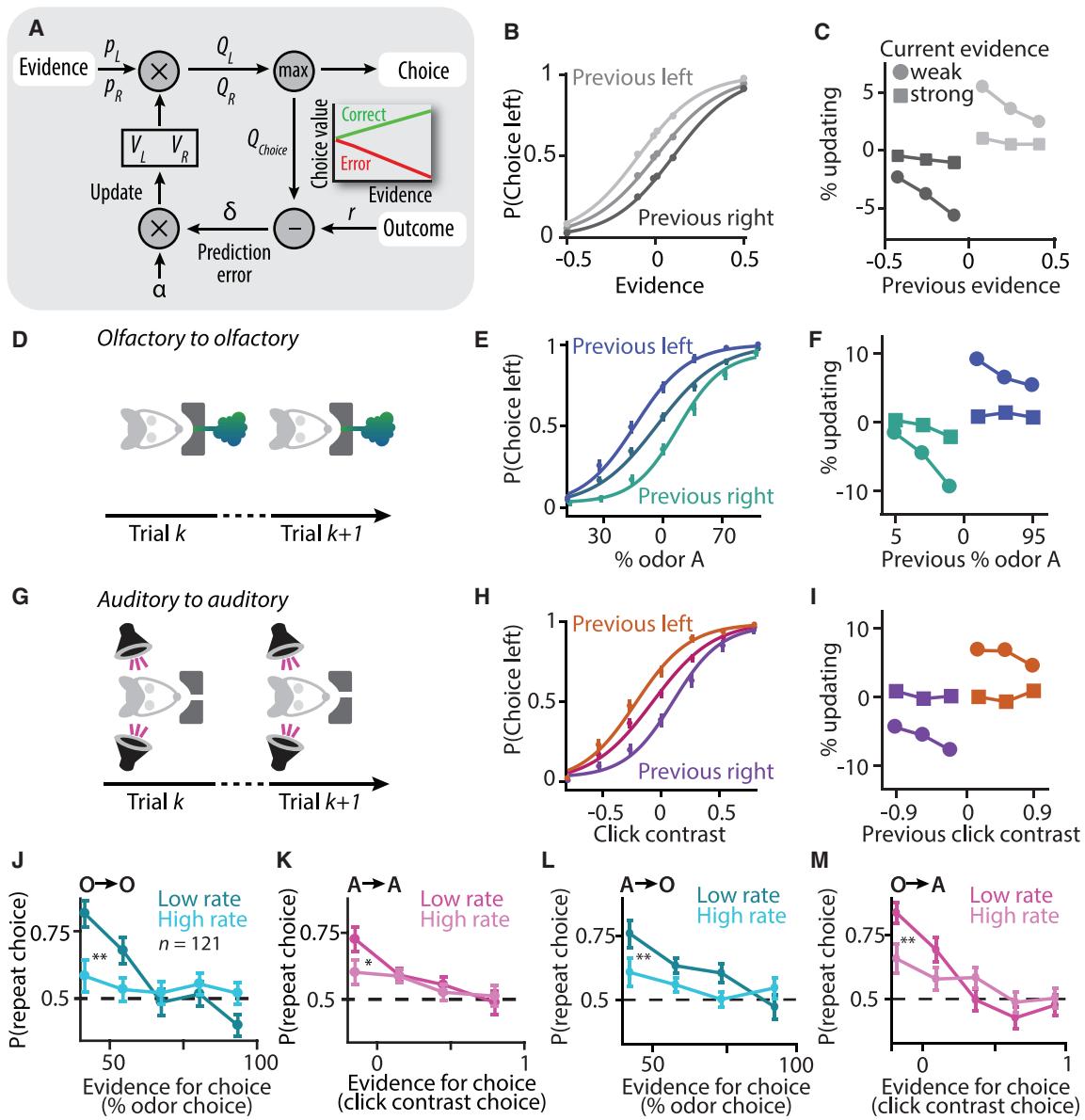


Figure 6. Orbitofrontal Neurons Predict Confidence-Guided Choice Updating

- (A) Belief-state reinforcement learning model (for details, see STAR Methods). Subjective sensory evidence (stimulus contrast with perceptual noise) results in a belief about the probability that the left or right choice is rewarded (p_L or p_R). Action values Q are computed by multiplying this probability with cached values V for right and left choices. A greedy policy selects the action (left or right) with higher action value Q . Choice values Q_{choice} follow signatures of confidence (inset). Upon reward r , a prediction error δ is used to update cached value of the chosen option.
- (B) The model produces stochastic choice behavior resulting in a psychometric function. Conditioning the psychometric on the previous trial's choice (left or right) reveals trial-by-trial choice updating, a bias to repeat the same choice.
- (C) The model agent's bias in choice behavior, i.e., trial-by-trial updating, as a function of current evidence (circles, high; squares, low) and previous evidence.
- (D) Analysis for olfactory trials preceded by a rewarded olfactory trial (E and F).
- (E) Psychometric functions in olfactory trials conditioned on the previous trial's correct choice (left or right).
- (F) Trial-by-trial updating as a function of current evidence (circles, high; squares, low) and previous evidence in consecutive olfactory trials. Updating was highest when previous and current evidence was weak; i.e., uncertainty high.
- (G) Analysis for auditory trials preceded by a rewarded auditory trial (H and I).
- (H and I) Same conventions as in (E and F) for auditory choices.
- (J) The probability of repeating an olfactory choice split by the firing rate of confidence(+) neurons in the anticipation period of the previous olfactory trial. Low firing rate predicted strong choice updating.
- (K) Same convention as in (J) for auditory choices preceded by an auditory choice.
- (L) Same convention as in (J) for olfactory choices preceded by an auditory choice (cross-modality updating).

(legend continued on next page)

Second, OFC firing rates predict time investment on a trial-to-trial basis, beyond what is expected from confidence tuning alone. This suggests that OFC activity drives confidence-guided time investments, consistent with previous observations that OFC inactivation specifically impairs rats' ability to invest time in their decisions according to their confidence level (Lak et al., 2014), and reminiscent to findings that frontal cortical regions determine action timing in rats (Murakami et al., 2014, 2017).

Third, we used two sensory modalities to test the generality of confidence representations, a first step toward showing meta-cognitive representations of confidence. Nevertheless, confidence about other decisions, such as a value-guided (De Martino et al., 2013) or memory-based choices (Miyamoto et al., 2017, 2018) and novel experiences may require different brain regions. OFC has a well-established role in value-guided decisions (Padoa-Schioppa and Conen, 2017; Stalnaker et al., 2015; Wallis, 2007; Xie and Padoa-Schioppa, 2016; van Duuren et al., 2009), and therefore, we consider it likely that our results will generalize beyond sensory-based decisions.

Finally, we showed that OFC neurons predict confidence-guided choice strategy updating. Appropriately estimating confidence is essential for learning in probabilistic environments (Drugowitsch et al., 2019; Lak et al., 2020a; Lak et al., 2020b; Meyniel et al., 2015b), and dopamine neurons have been shown to represent confidence-scaled prediction errors that are used for choice updating (Lak et al., 2017; Lak et al., 2020b). These results raise the intriguing possibility that confidence computations in OFC relay this information to neuromodulatory systems to enable confidence-guided learning, consistent with recent evidence that frontal cortex relays information about expected reward to ventral tegmental area (Starkweather et al., 2018; Takahashi et al., 2011). How the dynamic activity patterns across sensory modalities enable linking abstract confidence representations to subsequent behaviors for both learning and time investment remains an open question.

Relevance to Orbitofrontal and Prefrontal Function

Our observation that rat OFC neurons signal decision confidence across modalities supports OFC's conceptualization in outcome prediction and valuation. OFC is known to represent reward expectancies and value of choice options (Hirokawa et al., 2019; Padoa-Schioppa and Assad, 2006; Rudebeck and Murray, 2014; Stalnaker et al., 2015; Wallis, 2007; van Duuren et al., 2009; Constantinople et al., 2019; Roitman and Roitman, 2010; Sul et al., 2010). OFC lesions have been shown to affect a number of goal-directed behaviors in which values have to be inferred (Jones et al., 2012; Rudebeck and Murray, 2014; Stalnaker et al., 2015; Gremel and Costa, 2013). OFC has also been shown to support economic decisions where the value of each option is estimated either based on past experience or subjective preference between options (Padoa-Schioppa and Conen, 2017; Steiner and Redish, 2014). OFC activity has been studied less under perceptual ambiguity, when the cues that predict reward are uncertain. In these situations, the value of a

choice should reflect the probability of making a correct perceptual decision; i.e., decision confidence. Thus, a confidence computation transforms a proposition about ambiguous evidence into a proposition about outcome likelihood, indispensable for any valuation process and hence consistent with a hidden state representation of task space in OFC (Gershman, 2018; Wilson et al., 2014). Abstract confidence signals could thus serve as a basis to make economic decisions and guide actions based on values inferred from perceptual uncertainty.

A sensory-modality-general representation of decision confidence might have been promoted by our task space, which is indifferent to the sensory modality after making a choice. It remains an open question whether situations in which sensory modality information is required for learning would promote a sensory-modality-specific representation of decision confidence. On the other hand, we observed stronger within-modality than cross-modality choice updating, suggesting that rats considered stimulus identity and statistics across trials for within-modality learning ("model based") in addition to learning about choice values irrespective of sensory stimuli ("model free"). Our task design included trials that were never rewarded (probe trials), which could incentivize model-based learning. These observations suggest that state space representations reflect a mixture of model-based and model-free learning processes, as has been reported in complex behavioral tasks (Kim et al., 2019; Kool et al., 2016). However, we did not observe a difference in neuronal predictions of choice updating within or across sensory modalities (Figures 6J–6M), supporting the notion that confidence-encoding neurons in OFC generalize across sensory modalities, consistent with a model-free state-space representation.

In humans, representations of confidence have been reported in agranular posterior medial cingulate cortex (Bang and Fleming, 2018; Fleming et al., 2018; Rushworth and Behrens, 2008). Indeed, a precise homology between rodent and primate OFC remains to be established due to differences in criteria as well as methodological discrepancies across studies (Uylings et al., 2003; Wallis, 2011; Ongür and Price, 2000). Moreover, increased parcellation of frontal cortex in primates could lead to a separation of functions that are intermingled within rodent OFC (Carlén, 2017). Nevertheless, a number of studies have found similar functional properties across species in OFC, such as reward value and risk (Kuwabara et al., 2020; Li et al., 2016; O'Neill and Schultz, 2010; Padoa-Schioppa and Assad, 2006; Wallis, 2011; Constantinople et al., 2019; van Duuren et al., 2009; Rudebeck and Murray, 2014).

A Metacognitive Representation in Orbitofrontal Cortex?

Our core observations that OFC neurons represent decision confidence and predict confidence reports fit into a statistical framework of decision confidence (Hangya et al., 2016; Ott et al., 2018). Beyond the study of neural mechanisms, this framework has been used to argue that pre-verbal infants use

(M) Same convention as in (J) for auditory choices preceded by an olfactory choice (cross-modality updating). Error bars represent SEMs across pooled trials. * $p < 0.05$ ** $p < 0.01$ (logistic regression with post hoc t test). See also Figure S6.

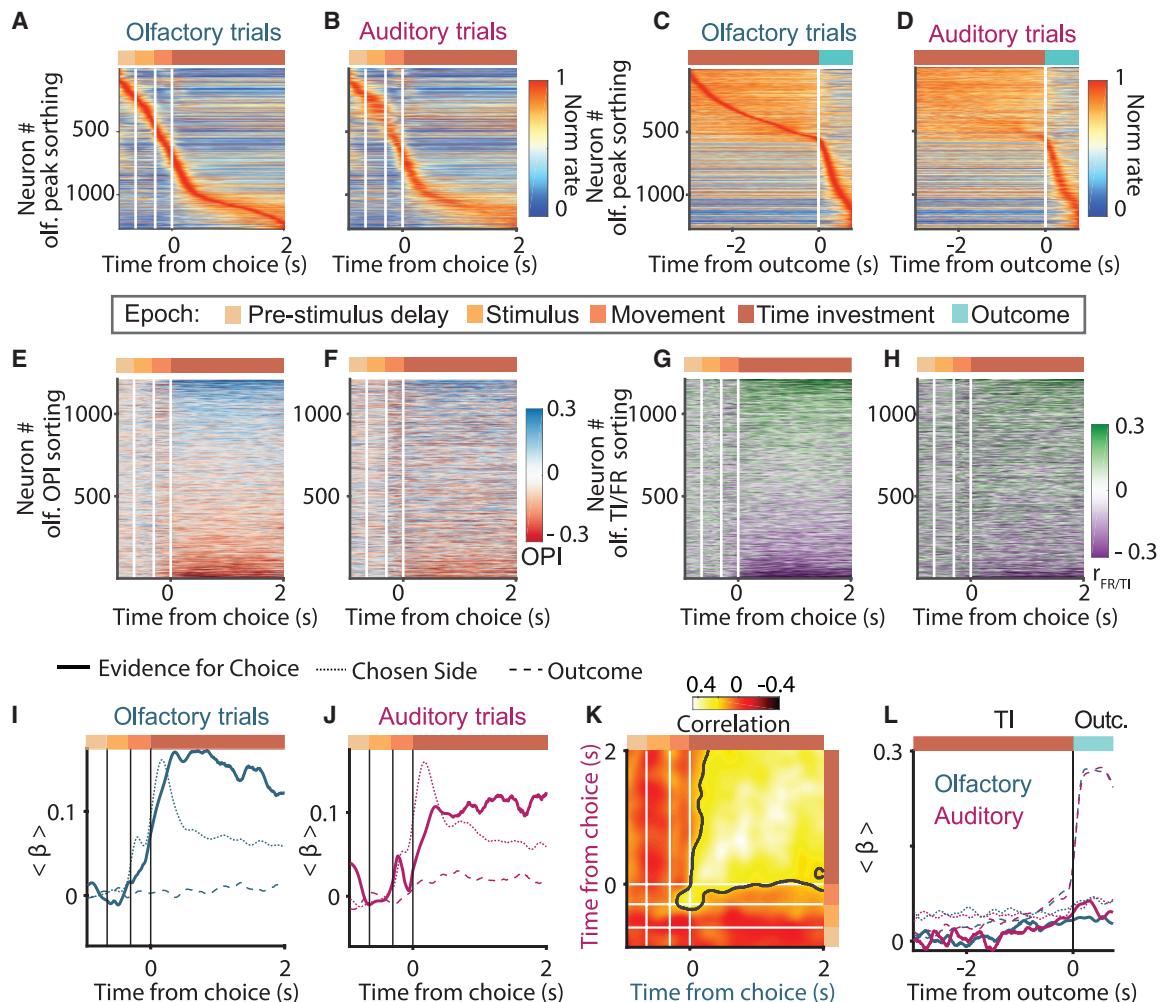


Figure 7. Orbitofrontal Neurons Show Similar Activity Dynamics and Encoding across Sensory Modalities

- (A) Trial-averaged activity of each neuron in the anticipation period sorted by time point of peak firing rate for olfactory trials.
- (B) Same as (A) for auditory trials and using the same alignment as for olfactory trial.
- (C) Trial-averaged activity in the outcome period sorted by temporal location of peak firing rate for olfactory trials.
- (D) Same as (C) for auditory trials and using the same alignment as for olfactory trial.
- (E) Time course of outcome prediction for olfactory trials. Neurons are sorted using the average outcome prediction in the anticipation period in olfactory trials only.
- (F) Time course of outcome prediction for auditory trials (same conventions as in A). Note that neurons are sorted based on outcome prediction in olfactory trials only.
- (G) Time course of correlations between firing rate and time investment for olfactory trials. Neurons are sorted based on their average correlation between firing rate and time investment in the anticipation period in olfactory trials only.
- (H) Time course of correlations between firing rate and time investment for auditory trials. Note that neurons are sorted based on their average correlation between firing rate and time investment in olfactory trials only.
- (I) Average GLM coefficients (normalized for sign, see STAR Methods) of outcome-predictive neurons ($n = 495$) across time aligned to choice port entry for olfactory trials.
- (J) Similar time course as in (I) but for auditory trials.
- (K) Cross-temporal correlation across sensory modalities (horizontal axis, olfactory; vertical axis, auditory) of normalized coefficient $\beta_{EvidenceForChoice}$ aligned to choice port entry. Black line, area denoting correlation coefficients with $p < 0.05$, t test, Bonferroni-corrected for multiple comparisons.
- (L) Average normalized GLM coefficients (normalized for sign, see STAR Methods) of outcome-predictive neurons ($n = 495$) across time aligned to outcome for both modalities.

See also Figure S7.

metacognitive evaluation to adapt their behavior (Goupil and Kouider, 2016; Goupil et al., 2016). However, such statistical framework does not constrain the neural architecture that realizes confidence computations and therefore is not sufficient to account for our observation that confidence representations are centralized (i.e., within one brain area) and abstract (i.e., generalized across sensory modalities and behavioral outputs). These observations are consistent with key predictions from metacognitive notions of confidence in psychology that conjecture a “metacognitive bottleneck” for computing abstract confidence (Ais et al., 2016; Dehaene et al., 2017; Dunlosky and Metcalfe, 2009). Our results provide evidence that, in rodents, OFC supports metacognitive abilities with a centralized, abstract representation of decision confidence.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead Contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Rats
- METHOD DETAILS
 - Behavioral task
 - Training
 - Analysis of Behavioral Data
 - Statistical Confidence Model
 - Surgery and electrophysiology
 - Histology
 - Analysis of Neuronal Data
 - Confidence tuning curves
 - Time investment decoding analysis
 - Time course analyses
 - GLM analysis
 - Choice updating analysis
 - Additional interpretation of results
- QUANTIFICATION AND STATISTICAL ANALYSIS

ACKNOWLEDGMENTS

We thank members of the Kepecs lab, as well as Michael Lagler, Thomas Klausberger, and Anqi Zhang for comments on an earlier version of this manuscript. We thank Barry Burbach and Rob Eifert for technical help. This research was supported by NIH R01DA038209 and R01MH097061 to A.K., by the Deutsche Forschungsgemeinschaft (DFG) OT 562/1-1 to T.O., and by Wellcome Trust grant 213465 to A.L.

AUTHOR CONTRIBUTIONS

P.M. and A.K. designed the experiments. P.M. and J.H. conducted pilot experiments. P.M. and T.O. conducted the experiments. P.M., T.O., A.L., and A.K. performed analyses. P.M., T.O., and A.K. wrote the manuscript with input from A.L. and J.H.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 27, 2019

Revised: February 27, 2020

Accepted: May 11, 2020

Published: June 5, 2020

REFERENCES

- Ais, J., Zylberberg, A., Bartfeld, P., and Sigman, M. (2016). Individual consistency in the accuracy and distribution of confidence judgments. *Cognition* **146**, 377–386.
- Bach, D.R., and Dolan, R.J. (2012). Knowing how much you don't know: a neural organization of uncertainty estimates. *Nat. Rev. Neurosci.* **13**, 572–586.
- Bang, D., and Fleming, S.M. (2018). Distinct encoding of decision confidence in human medial prefrontal cortex. *Proc. Natl. Acad. Sci. USA* **115**, 6082–6087.
- Basso, M.A., and Wurtz, R.H. (1997). Modulation of neuronal activity by target uncertainty. *Nature* **389**, 66–69.
- Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., and Movshon, J.A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* **13**, 87–100.
- Brunton, B.W., Botvinick, M.M., and Brody, C.D. (2013). Rats and humans can optimally accumulate evidence for decision-making. *Science* **340**, 95–98.
- Carlén, M. (2017). What constitutes the prefrontal cortex? *Science* **358**, 478–482.
- Chamberlain, S.R., Menzies, L., Hampshire, A., Suckling, J., Fineberg, N.A., del Campo, N., Aitken, M., Craig, K., Owen, A.M., Bullmore, E.T., et al. (2008). Orbitofrontal dysfunction in patients with obsessive-compulsive disorder and their unaffected relatives. *Science* **321**, 421–422.
- Chung, J.E., Magland, J.F., Barnett, A.H., Tolosa, V.M., Tooker, A.C., Lee, K.Y., Shah, K.G., Felix, S.H., Frank, L.M., and Greengard, L.F. (2017). A Fully Automated Approach to Spike Sorting. *Neuron* **95**, 1381–1394.e6.
- Cohen, J.R., Asarnow, R.F., Sabb, F.W., Bilder, R.M., Bookheimer, S.Y., Knowlton, B.J., and Poldrack, R.A. (2011). Decoding continuous variables from neuroimaging data: basic and clinical applications. *Front. Neurosci.* **5**, 75.
- Constantinople, C.M., Piet, A.T., and Brody, C.D. (2019). An Analysis of Decision under Risk in Rats. *Curr. Biol.* **29**, 2066–2074.e5.
- Dasgupta, P., and Maskin, E. (2005). Uncertainty and Hyperbolic Discounting. *Am. Econ. Rev.* **95**, 1290–1299.
- De Martino, B., Fleming, S.M., Garrett, N., and Dolan, R.J. (2013). Confidence in value-based choice. *Nat. Neurosci.* **16**, 105–110.
- Dehaene, S., Lau, H., and Kouider, S. (2017). What is consciousness, and could machines have it? *Science* **358**, 486–492.
- Drugowitsch, J., Mendonça, A.G., Mainen, Z.F., and Pouget, A. (2019). Learning optimal decisions with confidence. *Proc. Natl. Acad. Sci. USA* **116**, 24872–24880.
- Dunlosky, J., and Metcalfe, J. (2009). Metacognition (Sage Publications).
- Feil, J., Sheppard, D., Fitzgerald, P.B., Yücel, M., Lubman, D.I., and Bradshaw, J.L. (2010). Addiction, compulsive drug seeking, and the role of frontostriatal mechanisms in regulating inhibitory control. *Neurosci. Biobehav. Rev.* **35**, 248–275.
- Fetsch, C.R., DeAngelis, G.C., and Angelaki, D.E. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nat. Rev. Neurosci.* **14**, 429–442.
- Fineberg, N.A., Potenza, M.N., Chamberlain, S.R., Berlin, H.A., Menzies, L., Bechara, A., Sahakian, B.J., Robbins, T.W., Bullmore, E.T., and Hollander, E. (2010). Probing compulsive and impulsive behaviors, from animal models to endophenotypes: a narrative review. *Neuropsychopharmacology* **35**, 591–604.

- Fleming, S.M., and Daw, N.D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychol. Rev.* 124, 91–114.
- Fleming, S.M., and Dolan, R.J. (2012). The neural basis of metacognitive ability. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1338–1349. <https://doi.org/10.1098/rstb.2011.0417>.
- Fleming, S.M., van der Putten, E.J., and Daw, N.D. (2018). Neural mediators of changes of mind about perceptual decisions. *Nat. Neurosci.* 21, 617–624.
- Gershman, S.J. (2018). The successor representation: its computational logic and neural substrates. *J. Neurosci.* 38, 7193–7200.
- Gillan, C.M., and Robbins, T.W. (2014). Goal-directed learning and obsessive-compulsive disorder. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369 <https://doi.org/10.1098/rstb.2013.0475>.
- Goupil, L., and Kouider, S. (2016). Behavioral and Neural Indices of Metacognitive Sensitivity in Preverbal Infants. *Curr. Biol.* 26, 3038–3045.
- Goupil, L., Romand-Monnier, M., and Kouider, S. (2016). Infants ask for help when they know they don't know. *Proc. Natl. Acad. Sci. USA* 113, 3492–3496.
- Gremel, C.M., and Costa, R.M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat. Commun.* 4, 2264. <https://doi.org/10.1038/ncomms326>.
- Grimaldi, P., Cho, S.H., Lau, H., and Basso, M.A. (2018). Superior colliculus signals decisions rather than confidence: analysis of single neurons. *J. Neurophysiol.* 120, 2614–2629.
- Hangya, B., Sanders, J.I., and Kepcs, A. (2016). A Mathematical Framework for Statistical Decision Confidence. *Neural Comput.* 28, 1840–1858.
- Hanks, T.D., and Summerfield, C. (2017). Perceptual Decision Making in Rodents, Monkeys, and Humans. *Neuron* 93, 15–31.
- Hirokawa, J., Vaughan, A., Masset, P., Ott, T., and Kepcs, A. (2019). Frontal cortex neuron types categorically encode single decision variables. *Nature* 576, 446–451.
- Jones, J.L., Esber, G.R., McDannald, M.A., Gruber, A.J., Hernandez, A., Mirenzi, A., and Schoenbaum, G. (2012). Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science* 338, 953–956.
- Kable, J.W., and Glimcher, P.W. (2009). The neurobiology of decision: consensus and controversy. *Neuron* 63, 733–745.
- Kepcs, A., Uchida, N., Zariwala, H.A., and Mainen, Z.F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature* 455, 227–231.
- Kiani, R., and Shadlen, M.N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324, 759–764.
- Kim, D., Park, G.Y., O'Doherty, J.P., and Lee, S.W. (2019). Task complexity interacts with state-space uncertainty in the arbitration between model-based and model-free learning. *Nat. Commun.* 10, 5738.
- Knill, D.C., and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–719.
- Kobak, D., Brendel, W., Constantinidis, C., Feierstein, C.E., Kepcs, A., Mainen, Z.F., Qi, X.L., Romo, R., Uchida, N., and Machens, C.K. (2016). Demixed principal component analysis of neural population data. *eLife* 5, 1–37.
- Komura, Y., Nikkuni, A., Hirashima, N., Uetake, T., and Miyamoto, A. (2013). Responses of pulvinar neurons reflect a subject's confidence in visual categorization. *Nat. Neurosci.* 16, 749–755.
- Kool, W., Cushman, F.A., and Gershman, S.J. (2016). When Does Model-Based Control Pay Off? *PLoS Comput. Biol.* 12, e1005090.
- Kording, K.P., and Wolpert, D.M. (2004). Bayesian integration in sensorimotor learning. *Nature* 427, 244–247.
- Kuwabara, M., Kang, N., Holy, T.E., and Padoa-Schioppa, C. (2020). Neural mechanisms of economic choices in mice. *eLife* 9, e49669. <https://doi.org/10.7554/eLife.49669>.
- Lak, A., Costa, G.M., Romberg, E., Koulakov, A.A., Mainen, Z.F., and Kepcs, A. (2014). Orbitofrontal cortex is required for optimal waiting based on decision confidence. *Neuron* 84, 190–201.
- Lak, A., Nomoto, K., Keramati, M., Sakagami, M., and Kepcs, A. (2017). Midbrain Dopamine Neurons Signal Belief in Choice Accuracy during a Perceptual Decision. *Curr. Biol.* 27, 821–832.
- Lak, A., Hueske, E., Hirokawa, J., Masset, P., Ott, T., Urai, A.E., Donner, T.H., Carandini, M., Tonegawa, S., Uchida, N., and Kepcs, A. (2020a). Reinforcement biases subsequent perceptual decisions when confidence is low, a wide-spread behavioral phenomenon. *eLife* 9, 1–33.
- Lak, A., Okun, M., Moss, M.M., Gurnani, H., Farrell, K., Wells, M.J., Reddy, C.B., Kepcs, A., Harris, K.D., and Carandini, M. (2020b). Dopaminergic and Prefrontal Basis of Learning from Sensory Confidence and Reward Value. *Neuron* 105, 700–711.E6. <https://doi.org/10.1016/j.neuron.2019.11.018>.
- Li, Y., Vanni-Mercier, G., Isnard, J., Mauguire, F., and Dreher, J.C. (2016). The neural dynamics of reward value and risk coding in the human orbitofrontal cortex. *Brain* 139, 1295–1309.
- Magland, J.F., and Barnett, A.H. (2015). Unimodal clustering using isotonic regression: ISO-SPLIT, pp. 1–29.
- Mar, A.C., Walker, A.L.J., Theobald, D.E., Eagle, D.M., and Robbins, T.W. (2011). Dissociable effects of lesions to orbitofrontal cortex subregions on impulsive choice in the rat. *J. Neurosci.* 31, 6398–6404.
- Metcalfe, J., and Son, L.K. (2013). Anoetic, noetic, and auto-noetic metacognition. In *Foundations of Metacognition*, M.J. Beran, J.L. Brandl, J. Perner, and J. Proust, eds. (Oxford University Press).
- Meyniel, F., Sigman, M., and Mainen, Z.F. (2015a). Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron* 88, 78–92.
- Meyniel, F., Schluenzer, D., and Dehaene, S. (2015b). The Sense of Confidence during Probabilistic Learning: A Normative Account. *PLoS Comput. Biol.* 11, e1004305.
- Middlebrooks, P.G., and Sommer, M.A. (2012). Neuronal correlates of metacognition in primate frontal cortex. *Neuron* 75, 517–530.
- Miyamoto, K., Osada, T., Setsuie, R., Takeda, M., Tamura, K., Adachi, Y., and Miyashita, Y. (2017). Causal neural network of metamemory for retrospection in primates. *Science* 355, 188–193.
- Miyamoto, K., Setsuie, R., Osada, T., and Miyashita, Y. (2018). Reversible Silencing of the Frontopolar Cortex Selectively Impairs Metacognitive Judgment on Non-experience in Primates. *Neuron* 97, 980–989.e6.
- Murakami, M., Vicente, M.I., Costa, G.M., and Mainen, Z.F. (2014). Neural antecedents of self-initiated actions in secondary motor cortex. *Nat. Neurosci.* 17, 1574–1582.
- Murakami, M., Shteingart, H., Loewenstein, Y., and Mainen, Z.F. (2017). Distinct Sources of Deterministic and Stochastic Components of Action Timing Decisions in Rodent Frontal Cortex. *Neuron* 94, 908–919.e7.
- O'Neill, M., and Schultz, W. (2010). Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron* 68, 789–800.
- Ongür, D., and Price, J.L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex* 10, 206–219. <https://doi.org/10.1093/cercor/10.3.206>.
- Ott, T., Masset, P., and Kepcs, A. (2018). The Neurobiology of Confidence: From Beliefs to Neurons. *Cold Spring Harb. Symp. Quant. Biol.* 83, 9–16. <https://doi.org/10.1101/sqb.2018.83.038794>.
- Padoa-Schioppa, C., and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.
- Padoa-Schioppa, C., and Conen, K.E. (2017). Orbitofrontal Cortex: A Neural Circuit for Economic Decisions. *Neuron* 96, 736–754.
- Pascoli, V., Hiver, A., Van Zessen, R., Loureiro, M., Achargui, R., Harada, M., Flakowski, J., and Lüscher, C. (2018). Stochastic synaptic plasticity underlying compulsion in a model of addiction. *Nature* 564, 366–371.
- Paxinos, G., and Watson, C. (2007). *The rat brain in stereotaxic coordinates*, Sixth Edition (Acad. Press).
- Pouget, A., Dayan, P., and Zemel, R.S. (2003). Inference and computation with population codes. *Annu. Rev. Neurosci.* 26, 381–410.
- Pouget, A., Drugowitsch, J., and Kepcs, A. (2016). Confidence and certainty: distinct probabilistic quantities for different goals. *Nat. Neurosci.* 19, 366–374.

- Richards, J.B., Mitchell, S.H., de Wit, H., and Seiden, L.S. (1997). Determination of discount functions in rats with an adjusting-amount procedure. *J. Exp. Anal. Behav.* 67, 353–366.
- Roesch, M.R., and Olson, C.R. (2005). Neuronal activity in primate orbitofrontal cortex reflects the value of time. *J. Neurophysiol.* 94, 2457–2471.
- Roesch, M.R., Taylor, A.R., and Schoenbaum, G. (2006). Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* 51, 509–520.
- Roitman, J.D., and Roitman, M.F. (2010). Risk-preference differentiates orbitofrontal cortex responses to freely chosen reward outcomes. *Eur. J. Neurosci.* 31, 1492–1500. <https://doi.org/10.1111/j.1460-9568.2010.07169>.
- Rudebeck, P.H., and Murray, E.A. (2014). The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. *Neuron* 84, 1143–1156.
- Rushworth, M.F.S., and Behrens, T.E.J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.* 11, 389–397.
- Rutishauser, U., Ye, S., Koroma, M., Tudusciuc, O., Ross, I.B., Chung, J.M., and Mamelak, A.N. (2015). Representation of retrieval confidence by single neurons in the human medial temporal lobe. *Nat. Neurosci.* 18, 1041–1050.
- Rutishauser, U., Aflalo, T., Rosario, E.R., Pouratian, N., and Andersen, R.A. (2018). Single-Neuron Representation of Memory Strength and Recognition Confidence in Left Human Posterior Parietal Cortex. *Neuron* 97, 209–220.e3.
- Sanders, J.I., 2014. A computational framework for understanding decision confidence. PhD thesis. Watson School of Biological Sciences, Cold Spring Harbor Laboratory.
- Sanders, J.I., and Kepcs, A. (2012). Choice ball: a response interface for two-choice psychometric discrimination in head-fixed mice. *J. Neurophysiol.* 108, 3416–3423.
- Sanders, J.I., and Kepcs, A. (2014). A low-cost programmable pulse generator for physiology and behavior. *Front. Neuroeng.* 7, 43.
- Sanders, J.I., Hangya, B., and Kepcs, A. (2016). Signatures of a Statistical Computation in the Human Sense of Confidence. *Neuron* 90, 499–506.
- Schoenbaum, G., Roesch, M.R., and Stalnaker, T.A. (2006). Orbitofrontal cortex, decision-making and drug addiction. *Trends Neurosci.* 29, 116–124.
- Shadlen, M.N., and Kiani, R. (2013). Decision making as a window on cognition. *Neuron* 80, 791–806.
- Stalnaker, T.A., Cooch, N.K., and Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. *Nat. Neurosci.* 18, 620–627.
- Starkweather, C.K., Gershman, S.J., and Uchida, N. (2018). The Medial Prefrontal Cortex Shapes Dopamine Reward Prediction Errors under State Uncertainty. *Neuron* 98, 616–629.e6.
- Steiner, A.P., and Redish, A.D. (2014). Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nat. Neurosci.* 17, 995–1002.
- Sul, J.H., Kim, H., Huh, N., Lee, D., and Jung, M.W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron* 66, 449–460. <https://doi.org/10.1016/j.neuron.2010.03.033>.
- Takahashi, Y.K., Roesch, M.R., Wilson, R.C., Toreson, K., O'Donnell, P., Niv, Y., and Schoenbaum, G. (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat. Neurosci.* 14, 1590–1597.
- Teichert, T., Yu, D., and Ferrera, V.P. (2014). Performance monitoring in monkey frontal eye field. *J. Neurosci.* 34, 1657–1671.
- Uchida, N., and Mainen, Z.F. (2003). Speed and accuracy of olfactory discrimination in the rat. *Nat. Neurosci.* 6, 1224–1229.
- Uylings, H.B.M., Groenewegen, H.J., and Kolb, B. (2003). Do rats have a prefrontal cortex? *Behav. Brain Res.* 146, 3–17.
- van Duuren, E., van der Plasse, G., Lankelma, J., Joosten, R.N.J.M.A., Feenstra, M.G.P., and Pennartz, C.M.A. (2009). Single-cell and Population Coding of Expected Reward Probability in the Orbitofrontal Cortex of the Rat. *J. Neurosci.* 29, 8965–8976. <https://doi.org/10.1523/JNEUROSCI.0005-09.2009>.
- Wallis, J.D. (2007). Orbitofrontal cortex and its contribution to decision-making. *Annu. Rev. Neurosci.* 30, 31–56.
- Wallis, J.D. (2011). Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nat. Neurosci.* 15, 13–19.
- Wilson, R.C., Takahashi, Y.K., Schoenbaum, G., and Niv, Y. (2014). Orbitofrontal cortex as a cognitive map of task space. *Neuron* 81, 267–279.
- Wystrach, A., Mangan, M., and Webb, B. (2015). Optimal cue integration in ants. *Proc. Biol. Sci.* 282, 20151484.
- Xie, J., and Padoa-Schioppa, C. (2016). Neuronal remapping and circuit persistence in economic decisions. *Nat. Neurosci.* 19, 855–861.
- Yeung, N., and Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1310–1321.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Fluriso (Isoflurane)	VetOne	Fluriso
Lidocaine	Henry Schien	LidoJect
Ketoprofen	Zoetis	Ketofen
Dil fluorescent labeling	Invitrogen	D3911
+2-octanol	Sigma-Aldrich	SKU 147982
-2-octanol	Sigma-Aldrich	SKU 147990
Experimental Models: Organisms/Strains		
Long Evans Rat	Taconic	LONGEV-M (discontinued)
Software and Algorithms		
MATLAB	Mathworks	https://www.mathworks.com/products/matlab.html
MountainSort	Chung et al., 2017	https://github.com/flatironinstitute/mountainsort_examples
Cheetah	Neuralynx	N/A
Other		
Bpod State Machine r1	Sanworks	https://www.sanworks.io/
PulsePal	Sanworks	https://www.sanworks.io/
64 port olfactometer	Island Motion	P/N 109-024
Tetrode wire (RO800, 1/4 hard PAC)	Sandvik	N/A
Digital Lynx system	Neuralynx	N/A

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Adam Kepcs (akepeecs@wustl.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The data and custom code that support the findings from this study are available from the Lead Contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Rats

We used 6 adults male Long Evans rats (Taconic). Rats were housed in pairs during training and then single-housed after implantation of the microdrive and maintained on a reverse 12 h light/dark cycle. The rats had *ad libitum* access to food and were under a liquid restriction schedule with daily monitoring of water intake to maintain a body weight of at least 85% of free-drinking weight. Rats obtained water during daily training sessions and received *ad libitum* water on weekends and as needed. All procedures involving animals were carried out in accordance with National Institute of Health standards and were approved by the Cold Spring Harbor Laboratory Institutional Animal Care and Use Committee.

METHOD DETAILS

Behavioral task

Behavioral procedures were conducted similarly as previously described (Kepecs et al., 2008; Lak et al., 2014). A rectangular behavioral box contained three ports aligned on one wall equipped with LEDs, infrared photodiodes, and phototransistors. Interruption of the infrared photo beam was used to determine port entries and port exits. The two side ports (choice ports) were additionally equipped with valve-controlled water spouts for reward delivery. The center port (stimulus port) was used to deliver olfactory stimuli through a custom-built olfactometer. During olfactory stimulus presentation, odor solutions (+/– 2-octanol, Sigma Aldrich) were placed daily on fresh filters to mix 1:10 with pure air and delivered at the stimulus port at an overall flow rate of 1 l/min. Air delivery without odors was maintained during the trial. For auditory stimuli, two speakers were placed symmetrically outside of the behavioral box at the left and right panel, aligned to the animal's head. The olfactometer, speakers, water valves, LEDs and phototransistors were controlled by the behavioral measurement system Bpod and Pulse Pal (Sanworks, NY) (Sanders and Kepecs, 2014) and custom software (MATLAB).

Six naive rats were trained on a post-decision temporal wager task to assess the rats' level of confidence in their decision. The first part of the trial was constructed as a two-alternative forced choice task. Rats self-initiated each trial by entering the central stimulus port. After a random delay of 0.2–0.4 s, either an olfactory or auditory stimulus was presented (randomly interleaved). For olfactory stimuli, rats had to determine the dominant odor of a mix of pure odorants +2-octanol and –2-octanol. Odor stimuli were delivered for at least 0.35 s or until the rat left the center port (max. 3 s). Odor mixtures were fixed at 7 concentration ratios, which we adjusted to match the performance levels for each mixture ratio across animals. We always included a ratio of 5/95 (easiest discrimination, left side rewarded), 50/50 (randomly rewarded) and 5/95 (easy discrimination, right side rewarded). Intermediate discrimination difficulty were typically around 42/58 (intermediate discrimination, left rewarded, presented ratio were 44/56, 42/58, 40/60, 35/65 or 30/70 depending on the animal) and 58/42 (intermediate discrimination, right rewarded, presented ratio was 56/44, 58/42, 60/40, 65/35 or 70/30 depending on the animal) as well as 47/53 (difficult discrimination, left rewarded, presented ratio was 48/52, 47/53, 46/54, 45/55 or 44/56 depending on the animal) and 53/47 (difficult discrimination, right rewarded, presented ratio was 52/48, 53/47, 54/46, 55/45 or 56/44 depending on the animal). After a variable odor sampling time, rats exited the stimulus port, which terminated odor delivery, and indicated their choice by entering one of two choice ports (left or right) with a maximum response time of 3 s after leaving the stimulus port. Choices were rewarded according to the dominant odor component in the mixture: Left port entries were rewarded for mixture ratio < 50; right port entries were rewarded for mixture ratios > 50. 50/50 ratios were randomly rewarded. For auditory stimuli, rats had to determine the side with the higher number of clicks in binaural streams of clicks. Auditory stimuli were random Poisson-distributed click trains played binaurally at the two speakers placed outside of the behavioral box for a fixed time of 0.25 s. For each rat, we chose a maximum click rate $\text{click}_{\text{max}}$ according to the performance of the animal, typically 50 clicks/s (full range 40–100 clicks/s). This maximum click rate was fixed for each animal. For each trial, we randomly chose a delta click rate between left and right from a uniform distribution between 0 and $\text{click}_{\text{max}}$. The sum of the left and right click rate was kept constant at $\text{click}_{\text{max}}$. Rats indicated their choice by exiting the stimulus port and entering one of two choice ports (left or right) with a maximum response time of 3 s after leaving the stimulus port. Choices were rewarded according the higher number of clicks presented between the left and right click train (equal number of licks were randomly rewarded). Exiting the stimulus port during the pre-stimulus delay or during the stimulus time (first 0.25 s) were followed by a white noise and a time out of 3–7 s.

To assess the rats' degree of confidence, we asked the rats to place a temporal wager on their decision by investing a self-determined amount of time after committing to a choice. For correct choices, we withheld reward delivery for a random time between 0.6–8 s drawn from an exponential distribution with time constant 1.5 s, shifted by 0.6 s and truncated at 8 s. We chose the exponential distribution to maintain a relatively constant reward expectation, i.e., a flat hazard rate. Rats had to keep poking the choice ports during the entire time period. On a subset of trials (10% probe trials), we withheld reward entirely by setting the reward timing to 20 s. In addition, there was no feedback on error trials. Consequently, rats decided to leave the choice ports after a variable amount of time to initiate the next trial. To avoid false detections of leaving decisions rats had a grace period of 0.2–0.4 s in which reentry into the choice ports was not considered as a leaving decision. After a leaving decision was complete, a short auditory tone indicated to the rat that the trial was over. This provided us with a continuous time investment as well as a binary choice on each non-rewarded trial (probe trials and error trials). Time investment served as an index of the rats' confidence estimate in their decision (see *Statistical Confidence Model* and Figures 1 and 2).

Training

We initially trained naive rats to enter the central stimulus port and subsequently collect a water reward at either of the two choice ports (left or right). Rats learned to enter the stimulus port for at least 0.5 s. Subsequently we introduced one sensory modality only, starting with stimuli of high discriminability. After performance levels reached above ~90% we introduced stimuli with decreasing discriminability, i.e., increasing difficulty, to sample the full range of the rats' choice behavior. During this early training stage, all correct choices were rewarded and for incorrect choices we provided a feedback of white noise and a time out (3–7 s). After stable performance levels (lapse rates < 5%–10%) we removed feedback for error trials and introduced randomly delayed rewards gradually, starting from a maximum of 2 s to a maximum of 8 s over the course of several sessions. We never reinforced leaving behavior at the choice port. Note that during training with one sensory modality we sampled ~5% of trials from the other sensory

modality. Subsequently, we repeated training procedure for the second modality. Finally, we randomly interleaved trials of both modalities. Shortly before recording sessions started, we introduced probe trials. We switched the order of modalities in the training procedure between animals. Overall training took typically between 40–60 training sessions over the course of 8–12 weeks. Each session lasted about 2–3 h, in which the rats typically performed 700–1000 trials. The first 50 trials of each session were sampled from easier stimuli, no probe trials were provided and the reward delay distribution was shortened. This ensured stable choice and time investments across both modalities.

Analysis of Behavioral Data

We collected 89,406 trials from six rats across 119 behavioral sessions (rats performed between 15 and 28 sessions each, number of trials within a session varied from 375 to 1172, mean = 751). We did not analyze the first 75 trials of each session (50 trials with easy contingencies – see *Training* – and an additional 25 trials to ensure the rats had reach plateau behavior) and trials with a time investment shorter than 0.6 s (< 3% of trials). For olfactory trials, we ranked stimuli according to their mixture ratio yielding a common evidence strength measure across rats. For auditory trials, we computed the binaural contrast $x_A = (N_{Left} - N_{Right}) / (N_{Left} + N_{Right})$ using the number of clicks presented in each trial. Psychometric curves for auditory trials ([Figures 1C, 2B, and S1](#)) were fitted using the model presented in [Figure 2A](#), left lower panel. We fitted the center (bias b) and standard deviation of a cumulative truncated Gaussian distribution as binaural contrast was constrained by definition between –1 and 1:

$$P(\text{choice} = 1 | x, b, \sigma) = \frac{\Phi(x, b, \sigma) - \Phi(-1, b, \sigma)}{\Phi(1, b, \sigma) - \Phi(-1, b, \sigma)}$$

Where $\Phi(x, \mu, \sigma)$ is the cumulative density function of the normal distribution of mean μ and standard deviation σ evaluated at point x :

$$\Phi(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

We used the *fit* function in MATLAB to find the parameters μ and σ that minimized the mean squared error. Error bars represent SEM across trials unless otherwise noted.

For the calibration and conditioned psychometric curves ([Figures 1E, 1G, 1H, and 1J](#)), we restricted our analysis to probe trials to keep the average accuracy equivalent to the rest of the trials. For the vevaiometric curve ([Figures 1F and 1I](#)), we included correct probe trials or all error trials with time investment longer than two seconds. Data analysis was performed using custom software (MATLAB).

Statistical Confidence Model

We developed a model-based approach to isolate the contribution of statistical decision confidence to time investment behavior, without requiring the fitting of several free parameters related to subjective valuations that were not explicitly measured, such as opportunity cost ([Lak et al., 2014](#)). To isolate the confidence computation, we first estimate the perceptual noise and the mapping function between time investment and confidence and then predict the statistically appropriate time investment in each trial.

We used a statistical model of confidence to evaluate if the rats' time investment behavior reflects an appropriate use of decision confidence ([Hangya et al., 2016; Sanders et al., 2016](#)). In our model approach (for details see below), sensory evidence is first transformed into an internal percept, i.e., a noisy representation of the evidence. The percept determines a choice (left or right) and a probability of being correct using a statistical definition of confidence. Confidence is then mapped onto a time investment in a calibration step. To isolate the confidence computation, we estimate the amount of perceptual noise using the animal's choices and we estimate the calibration function that maps confidence to time investment using the animal's empirical time investment distribution. This allows us to predict an optimal time investment for each trial that we can compare against the data and models with noise-corrupted confidence values to assess the optimality of time investment behavior.

We defined a generative model for our perceptual decision task by assuming that the percept \hat{x} is a noisy representation of the evidence strength x (Gaussian noise, i.e., $\hat{x} \sim \mathcal{N}(x, \sigma)$), and that the choice d is based on the percept alone, by comparing the percept to a decision boundary b (i.e., $d = 1$ (left choice) for $\hat{x} > b$, $d = -1$ (right choice) otherwise) (cf. [Figure 2](#)). Confidence c was defined as the probability of being correct given the subjective level of evidence (percept \hat{x}) and choice d , i.e., $c = P(\text{Correct} | \hat{x}, d)$. We can then generate a percept \hat{x} , choice d , and confidence c for each evidence strength x , for which the confidence c is given by $c = 0.5 * (1 + \text{erf}(\hat{x} - b) / \sigma\sqrt{2})$ ([Lak et al., 2014; Hangya et al., 2016](#)). To arrive at a non-parametric mapping function between confidence c and time investment T , we matched the rats' empirical time investment distribution h_T and the model confidence distribution h_c . Specifically, we define a mapping function $T = m(c)$ given by $m(c) = H_T^{-1}(H_c(c))$ with H_T and H_c the cumulative distribution functions of empirical time investment and model confidence distributions, respectively. This calibration of confidence to time investment values captures the rats' subjective cost of time and utility function, which we did not independently measure, without making assumptions about the shape of this mapping function. Importantly, this mapping function recovers any monotonic function between confidence and time investment. This calibration step is entirely agnostic to the appropriate use of confidence in a given trial (i.e., independent of a given x, d). Thus, model time investment predictions reflect an optimal use of statistical confidence in our generative

decision model given a specific amount of perceptual noise σ and given an empirical time investment distribution h_{Tl} , i.e., optimal time investment is given by $Tl_{opt} = m(c)$. Note that optimality in this context refers to an optimal use of statistical confidence to maximize reward rate. Since we did not independently determine the cost of time and reward utility, we cannot make any claims about maximization of subjective reward value. Therefore, our claim about optimality refers to the statistically correct use of evidence, which is required for a reward maximization strategy. Model predictions can then be evaluated for three key signatures of confidence (calibration curve, vevaiometric curve, and conditioned psychometric curve, cf. [Figure 2](#)) against the rats' time investment behavior.

For each rat, we defined the difficulty level x for auditory trials as the binaural contrast $x_A = (N_{Left} - N_{Right}) / (N_{Left} + N_{Right})$. For olfactory trials, we transformed odor mixture ratios (ranging from 5% to 95%) onto a common, continuous axis as the inverse cumulative density function of the standard normal distribution at the percentage of left choices ($d = 1$) for each odor mixture ratio, i.e., $x_O = \Phi^{-1}(P(\text{left}|\text{odor ratio}))(\mu = 0, \sigma = 1)$. We then fitted each rat's psychometric function using a Gaussian cumulative distribution with parameters σ (perceptual noise) and b (bias) using the same procedure described above (without truncation). Next, we generated 100,000 model samples using the same evidence strength x as used in behavioral sessions. We then generated model percepts \hat{x} ($\hat{x} \sim \mathcal{N}(x, \sigma)$), choices d (choice function f : $f(\hat{x}) = 1$ if $\hat{x} > 0$ and $f(\hat{x}) = -1$ if $\hat{x} < 0$), and confidence values c ($c = 0.5 * 1 + \text{erf}(|\hat{x}|/\sigma\sqrt{2})$) (note that we simplified equations using $b = 0$ since bias in all subjects was close to zero). Model confidence and empirical time investment (in probe trials) distributions were estimated using kernel density estimators with standard normal kernels (`ksdensity` function in MATLAB) evaluated at 30 points. Cumulative distribution functions were estimated by the cumulative sum of densities and inverse functions were estimated using MATLAB's quantile function to arrive at a mapping function $Tl = m(c)$. This procedure yielded model time investment prediction Tl for each trial, together with a percept \hat{x} and choice d for a given difficulty level x .

Note that the 'optimal' duration of time investment will vary with numerous internal factors, such as the cost of time and reward utility, which we treated as stable in our model. Lacking an independent assessment of these factors it is not possible to conclude whether the specific durations of time investment are overall optimal from the organism's perspective ([Bach and Dolan, 2012](#); [Constantinople et al., 2019](#); [Dasgupta and Maskin, 2005](#); [Kable and Glimcher, 2009](#); [Richards et al., 1997](#)). Nevertheless, we kept reward and time contingencies constant, so that we could evaluate whether variations in time investment reflected an optimal use of sensory evidence, i.e., statistical decision confidence.

Surgery and electrophysiology

After animals were fully trained, we implanted micro-drives for tetrode recordings under general anesthesia (2%–3% isoflurane, VettOne) in 6 rats. We stereotactically implanted custom-built micro-drives with 32 or 64 channels in the left OFC (target coordinates: +3.7 mm AP; 3.2 mm ML; -2.8 mm DV). Individual tetrodes were custom-made from four twisted polyimide-coated nichrome wires with 12 μm diameter (Sandvik AB). Wires were gold plated to a final impedance of 200–300 kOhm and organized in 2–4 independently movable bundles of 4 tetrodes each. After post-surgery recovery and analgesics (5 mg/kg ketoprofen) of at least 7 days, we started behavioral sessions and extracellular recordings by attaching a weight-balanced tether cable to the microdrive's electrical interface board while the rat was in the behavioral chamber. Voltage signals were first amplified using a unity-gain op-amp preamplifier fitted to the microdrive and subsequently digitized at 32,556 Hz (DigiLynx, Neuralynx) using Cheetah software (Neuralynx). Behavioral event pulses were generated by the behavioral control system (Bpod, Sanworks, NY) and recorded by Cheetah. After each recording session, we advanced tetrode bundles by at least 40 μm .

Single units were clustered using the automated clustering platform MountainSort (Simons Foundation, NY) ([Chung et al., 2017](#); [Magland and Barnett, 2015](#)). Continuously sampled data for each tetrode was bandpass-filtered (300 Hz–6,000 Hz) and normalized removing correlations between individual leads of one tetrode. Subsequently, events with negative voltage deflections were detected by thresholding normalized voltage signals at three standard deviations. Extracted events (event width of 1.67 ms) were clustered using a non-parametric, density-based clustering algorithm (ISO-SPLIT), which clusters individual events based on their representation in a low-dimensional feature space ([Magland and Barnett, 2015](#)). Subsequently, we manually evaluated each sorted cluster based on several cluster quality metrics. We used MountainSort's isolation metric ([Chung et al., 2017](#)), which quantifies how well each cluster is isolated from neighboring clusters using a non-parametric nearest-neighbor approach. Next, we used MountainSort's noise overlap metric ([Chung et al., 2017](#)), which quantifies the similarity between each cluster and with randomly chosen events (noise events). Both metrics roughly correspond to the fraction of events overlapping with another cluster (isolation) or noise cluster (noise overlap). In general, we only included clusters with an isolation larger than 0.96 and a noise overlap of less than 0.05. In addition, we carefully inspected the auto-correlogram for each cluster, excluding clusters without a clear dip in their respective auto-correlograms (more than 0.1% events within an inter-spike interval of 1.5 ms). In some instances, we manually merged two clusters which showed a clear dip in their respective cross-correlogram as well as a pairwise low isolation metric (indicating a high cluster overlap) after evaluation of waveform shape and spread. Overall, this procedure ensured that we only analyzed well-isolated clusters considered to be single units.

Histology

After finishing recordings, rats were euthanized with pentobarbital and transcardially perfused with saline and 4% paraformaldehyde. Brains were removed and post-fixed in paraformaldehyde for several days. Coronal brain sections with 100–150 μm thickness were

prepared using a vibratome. Recording tracks and sites were localized by coating tetrode bundles with fluorescent dye (Dil, Invitrogen) before implantation and/or by targeted electrolytic lesions before perfusion applying 20 μ A for 5 s at 1–4 leads.

Analysis of Neuronal Data

We recorded from 1593 well-isolated single units in the OFC of six rats. 1211 neurons with an average firing rate above 1 s⁻¹ entered subsequent analyses. Data analysis was performed using custom software (MATLAB).

We only used sessions in which average choice accuracy was above 60% for both modalities. We obtained average activity time courses for single neurons, its spike density function, by computing the mean firing rate across trials within a condition for each time point using time bins of size $dt = 10$ ms. The resulting time courses were smoothed using a $\sigma = 50$ ms Gaussian window. For olfactory trials, difficult trials included neutral trials and the hardest odor ratio. For auditory trials, difficult trials were defined as an absolute value of the binaural contrast smaller than 0.2. The time investment duration bins used to compare time courses in single neuron examples and population averages were 2.5 s < TI < 4 s, 4 s < TI < 5.5 s, 5.5 s < TI < 6.5 s and 6.5 s < TI (We only used trials in which the time investment was above 2 s). To compute population time courses, we z-scored the values of the single neuron time courses and averaged the activity across the population.

Confidence tuning curves

For confidence and time investment tuning, we used the average firing rate in the first two seconds of the reward anticipation period ('anticipation period'). We did not include trials where rats waited less than 2 s, since we observed that very short waiting times were indicative of low task engagement. Note that rats usually waited substantially longer than 2 s (up to 12 s, mean time investments were 5.3 s after olfactory choices and 5.5 s for auditory choices, see [Figures 1D](#) and [2B](#)). For the calibration curve, we used probe trials and error trials and resampled the trials to match the behavioral session's average accuracy. For the firing rate time investment correlation, we fitted a linear model and plotted the model prediction and computed the correlation coefficient ([Figures 4Aii](#), [4Aiv](#), [S3A–S3C](#), [S3Dvi](#) and [S3Dvii](#)).

The outcome prediction indices (OPI) provides a simple scalar measure that captures a critical aspect of confidence, yet outcome prediction could also arise from alternative computational mechanisms such as error detection or recall of past experiences based on the strength of evidence ([Kepes et al., 2008](#); [Teichert et al., 2014](#); [Yeung and Summerfield, 2012](#)). To compute OPI we computed the area under the receiver operating characteristic (auROC) between the average firing rate in the anticipation period for correct and error trials re-scaled from -1 to 1, i.e., OPI = 0 corresponds to zero outcome prediction. Statistical significance was assessed using bootstrap ($n = 200$), resampling trials with replacement trials each category and defining outcome predictive neurons with OPI different from zero with $p < 0.025$ in either category. To obtain time investment predictions, we computed the Spearman correlation between the spike density function and time investments across probe and error trials.

The correlation of the OPIs across modalities was strong, albeit not perfect, with some neurons appearing to predict outcome differently for olfactory and auditory decisions. Therefore, we examined whether the correlation was limited by the noise due to low firing rates and trial numbers (trials with time investments longer than 2 s, olfactory trials: $n = 169 \pm 44$ and auditory trials: $n = 124 \pm 35$). For each neuron, we divided the trials for each modality in half and re-computed outcome prediction indices for each subset of trials. We then computed the correlation within modality across the two subsets of trials. For outcome predictive neurons, the average Pearson correlation for olfactory decisions was $r_{olf} = 0.76 \pm 0.02$ and the average Pearson correlation for auditory decisions was $r_{aud} = 0.67 \pm 0.02$ (average correlation over 200 random partitions, $p < 1e-10$). This reveals that the variability in the outcome prediction index across modalities is in part constrained by our ability to estimate it. Furthermore, visual examination of neurons that were apparently incoherent in their OPI across modalities are consistent with this interpretation ([Figures S3C](#) and [S3D](#)).

For population tuning curves ([Figures 3B–3I](#), [4C](#), and [4E](#)), we first z-scored the firing rates of outcome-predictive neurons, computed their tuning curves, and averaged these across neurons.

Confidence-encoding neurons, by definition, are expected to predict confidence-guided time investments, because confidence is determined by the amount of evidence supporting a choice. Therefore, to determine whether confidence-encoding neurons contribute to confidence-guided time investments, we developed measures of firing rate and behavior co-variation that take into account their correlation with the strength of sensory evidence supporting a choice. This is conceptually similar to choice probability analysis (which estimates the probability of predicting the choice from a single neuron's activity) for determining the contribution of single neurons to binary decisions ([Britten et al., 1996](#)). Using a two-step linear model, we first calculated each neuron's firing rate not explained by the strength of evidence by taking the residuals of a linear model expressing firing rate as a function of evidence strength.

To quantify the contribution of the firing rates to time investment behavior we considered how well firing rates predicted time investment duration after taking the information provided by the strength of evidence into account. Because time investments provide a continuous measure of decision confidence we used a linear model to analyze the contribution of each neuron's firing rate to the rats' time investments after accounting for evidence strength (signed value of the evidence supporting the choice: +1 indicates an easy correct trial while -1 indicates an easy error trial). For olfactory trials we acquired a sufficient number of trials with neutral evidence (50% odor mixture trials) to determine the contribution of neural activity to time investments with uninformative sensory evidence. For these trials the choice is driven by internal noise as there is no sensory evidence to support one choice over the other, and we thus computed the correlation between firing rate and time investment directly. For trials in which there was evidence to support

the choice, we built two linear models to remove the contribution of the strength of evidence to both the firing rates and the time investments. One expressed firing rate as a function of evidence strength, and the other expressed time investment as a function of evidence strength. The residuals of these models provided the portion of neuronal firing rate and time investment not explained by evidence strength.

For each trial we remove the linear contribution of the strength of evidence supporting the choice to the firing rate and the time investment. For olfactory trial, we used the ranked difficulty of the trials and uniformly distributed it between -0.95 to 0.95 . An easy error trial will have a strength of evidence of -0.95 while an easy correct trial will have an evidence for the choice of 0.95 . For auditory trials we used the value of the binaural contrast supporting the choice: $\text{Evidence} = (N_{\text{for}} - N_{\text{against}}) / (N_{\text{for}} + N_{\text{against}})$. We used the evidence value to build linear models to estimate both the firing rate and the time investments independently for each modality using the following models:

$$\text{FiringRate} = \beta_{0,\text{ev}} + \beta_{1,\text{ev}} \cdot \text{Evidence}$$

$$\text{TimeInvestment} = \beta_{0,\text{ti}} + \beta_{1,\text{ti}} \cdot \text{Evidence}$$

We denote FR_{evidence} and TI_{evidence} the estimated firing rate and time investment respectively. We then compute the residual metrics (firing rate or time investment) by subtracting the estimated value of the metric according to the linear model from the true value of the metric:

$$\text{Residual}_{FR} = \text{FiringRate} - FR_{\text{evidence}}$$

$$\text{Residual}_{TI} = \text{TimeInvestment} - TI_{\text{evidence}}$$

We then used the residual metrics to estimate the contribution of the firing rates to the time investment behavior.

$$\text{Residual}_{TI} = \beta_{0,\text{res}} + \beta_{1,\text{res}} \cdot \text{Residual}_{FR}$$

In Figures 7E and 7F, we analyze the spearman correlation between the values of $\beta_{1,\text{res}}$ across modalities.

Time investment decoding analysis

For the time investment decoding analysis (Figure 5) we constructed a linear predictor of the time investment for each session by using the neurons simultaneously recorded in a given session. For stability reasons, we only used sessions ($n_{\text{sessions}} = 94$ out of 119) with more than 50 auditory and 50 olfactory trials with a time investment superior to 2 s leading to a decision to leave (excluding rewarded trials). For each neuron, we used the firing rate in the anticipation period (2 s after entry in choice port) and we excluded neurons with an average firing rate below 1 s^{-1} over the whole session. We used the MATLAB function `stepwiselm` to construct an estimate of the time investment on each trial given simultaneously recorded neurons ($n_{\text{neurons}} = 7.7 \pm 6.1$). To avoid overfitting and lack of generalization, the linear model only used a neuron as predictor if adding the neuron would decrease the Akaike information criterion (AIC). We used a ten-fold cross-validation procedure (Cohen et al., 2011) to compute estimates of time investment for each trial. For each session we computed the Spearman correlation between predicted time investment and observed time investment as a measure of the quality of the decoding. For each session and for each of the ten partitions, we inferred two decoders: One trained on 90% of the olfactory trials and one trained on 90% of the auditory trials. We then tested the decoders on the remaining 10% of the trials. The comparison was performed on sessions in which the decoders used at least one neuron on average ($n_{\text{sessions}} = 59$). To assess the significance of the decoding quality between two conditions across sessions we computed a two-sample Kolmogorov-Smirnov test on the distribution of session-wise correlations and corrected for multiple comparisons (Bonferroni correction).

Time course analyses

For time course analyses (Figure 7), we normalized variable time periods of the trials (pre-stimulus delay, stimulus sampling period and movement period). The number of time bins used for these period in the spike density function was matched to the mean duration of those epochs across all trials (see Figure 10 in (Kobak et al., 2016)). We used a time bin of 10 ms and the lengths of the time normalized epochs were as follows: pre-stimulus delay: 30 time bins; stimulus sampling period: 35 time bins; movement period: 30 time bins. To assess the neurons' overall time course (Figures 7A–7D), we normalized each neuron to its peak activity. We then sorted neurons according to the time point of the peak activity for olfactory trials. We used the same sorting order for auditory trials. For the early stage plot (Figures 7A and 7B) we used the time from choice port entry to the first 2 s of time investment (unnormalized). For the late stage (Figures 7C and 7D), we did not use time normalization and used the last 3 s of the time investment and 750 ms after outcome reveal (reward or decision to leave).

For OPI and time investment prediction time courses, we computed OPI and the correlation between firing rates and time investment in each time bin separately. We sorted neurons according to their average outcome prediction index in the anticipation period based on olfactory trials only (Figure 7E). Although sorted based on their olfactory selectivity, OPIs for auditory trials were similar to

olfactory trials (Figure 7F). Next, we quantified each neuron's ability to predict time investment separately for each sensory modality and time point (Figures 7G and 7H). We sorted neurons according to the strength of the correlation between the anticipation period firing rates and time investments based on olfactory trials only (Figure 7G). Although sorted based on their correlation during olfactory trials, correlations between firing rate and time investments for auditory trials were similar to olfactory trials (Figure 7H).

To assess the difference in the time course of the correlation across modalities of the outcome prediction indices and the time investment correlations we performed a bootstrap analysis (Figure S7E). We perform 200 runs by subsampling the neuronal population by 50% and computed the cross-correlation between outcome prediction indices and time investment correlations across the sub-population (diagonal of Figures S7C and S7D). We smoothed the resulting correlations using a moving average (5 time bins) and to assess how time courses differed, we computed an area under the ROC measure the ability to discriminate at a given time point whether a given time course was from the outcome prediction index or time investment correlation. The green bar indicate in Figure S7E indicates an area under the ROC above 0.975.

GLM analysis

In order to analyze the simultaneous coding of different task variables, we implemented a GLM using the spike density function presented above (See section *Analysis of neuronal data*). For each neuron and at each time point, the model attempts to fit the following equation to the spike density. The fit is done using the MATLAB function *fitglm* and is performed independently for olfactory and auditory trials:

$$\begin{aligned} \text{SpikeDensity}(t) = & \beta_{\text{Baseline}}(t) + \beta_{\text{ChosenSide}}(t) \cdot \text{ChosenSide} + \beta_{\text{EvidenceForChoice}}(t) \cdot \text{EvidenceForChoice} \\ & + \beta_{\text{Outcome}}(t) \cdot \text{Outcome} + \beta_{\text{Evidence}}(t) \cdot \text{Evidence} \end{aligned}$$

The factor $\beta_{\text{Baseline}}(t)$ captures the variation in the firing rate that is not captured by the other task-relevant variables, for example, a neuron could increase its activity when entering a port but there could be no information in this activity about the sensory evidence or the identity of the port. The factor $\beta_{\text{ChosenSide}}(t)$ captures the contribution of the side (left or right) that the animal chooses to the firing rate. The factor $\beta_{\text{EvidenceForChoice}}(t)$ captures the contribution of the strength of the evidence for the chosen side to the firing rate. An easy correct trial would correspond to a strength of evidence for the chosen side of 1 while an easy error trial would correspond to a strength of evidence for the chosen side of -1. The factor $\beta_{\text{Outcome}}(t)$ captures the contribution of the accuracy of the choice to the firing rate. The factor $\beta_{\text{Evidence}}(t)$ captures the contribution of the sensory evidence to the firing rate.

For the choice period analysis (Figures 7I–7K), the normalized coefficients were obtained as follow: for $\beta_{\text{EvidenceForChoice}}$ and β_{Outcome} , we used $+ \beta$ for neurons belonging to the confidence⁽⁺⁾ population and $- \beta$ for neurons belonging to the confidence⁽⁻⁾ population. For $\beta_{\text{ChosenSide}}$, we used $+ \beta_{\text{ChosenSide}}$ for neurons that were right selective in the anticipation period and $- \beta_{\text{ChosenSide}}$ for neurons that were left selective in the anticipation period. For Figure 7K we smoothed the spike density functions with a moving window average with a window of 20 time steps (200ms) before computing cross-temporal correlations.

For the outcome period analysis (Figure 7L), we used the sign of the Reward Index (RI) to normalize the coefficient. A positive reward index indicated higher firing in rewarded trials and negative index indicated higher firing in unrewarded trials. We normalized the parameters for $\beta_{\text{EvidenceForChoice}}$ and β_{Outcome} using $+ \beta$ for positive RI and $- \beta$ for neurons with negative RI. For $\beta_{\text{ChosenSide}}$, we used side selectivity on the outcome period.

Choice updating analysis

Perceptual choices can be biased by past experience through reward-based learning mechanisms. We quantified the degree to which rats biased their choices based by splitting trials according to whether the previous correct choice was 'left' or 'right' separately for olfactory and auditory trials (Figures 6B and 6F) pooling all trials from all recording sessions (same sessions as above). We quantified choice updating by calculating the probability of repeating the same choice (previous correct trial) considering all four possible sensory-modality transitions (olfactory-to-olfactory, olfactory-to-auditory, auditory-to-auditory, auditory-to-olfactory) (Figures 6C and 6G). We assessed different predictors on repeating a choice, namely evidence for choice (β_1) in the current trial k , previous trial's evidence (β_2), and a constant term (β_0) using a logistic regression model (f is the logistic function):

$$\text{Repeat}_k = f(\beta_0 + \beta_1 \cdot \text{Evidence}_k + \beta_2 \cdot \text{Evidence}_{k-1})$$

We assessed whether confidence encoding neurons predicted choice updating using the same logistic regression model adding a firing rate term (β_3) (average firing rate of confidence encoding neurons in each trial) and an interaction term (not shown):

$$\text{Repeat}_k = f(\beta_0 + \beta_1 \cdot \text{Evidence}_k + \beta_2 \cdot \text{Evidence}_{k-1} + \beta_3 \cdot \text{Firing}_{k-1})$$

We assessed whether firing rate added explanatory power to this model by calculating the change in Akaike information criterion (AIC), a measure to compare model quality that considers the different number of free parameters. We compared AIC of this model with a model omitting firing rates and used bootstrap to calculate statistical significance.

Trial-history dependent choice updating can be explained by reinforcement learning models. In perceptual decisions there is no overt learning after subjects are proficient in the task. Nevertheless, choice biases based on past trial's outcome are still prevalent even in perceptual decision tasks. To explain this phenomenon, we used a reinforcement learning model that used a perceptual belief

state to learn about action values ('left' or 'right'). This model produced confidence-scaled prediction errors that are used to update state values, therefore resulting in confidence-dependent choice biases in the next trial.

We used a temporal difference reinforcement learning model as previously described (Lak et al., 2017, 2020). In brief, subjects made a left (L) or right (R) choice by a greedy policy selecting the action (L or R) which maximizes action values Q (i.e., Q_L or Q_R). In each trial, Q was given by stored action values V weighed by the subjective probability of receiving a reward at the left or right port p : $Q_L = p_L \cdot V_L$ and $Q_R = p_R \cdot V_R$. Similar to our perceptual decision model (see above), p_R represents the probability of the stimulus x being on the right side of the decision boundary b based on the subjective percept \hat{x} , the noisy representation of the sensory evidence, i.e., $p_R = P(x > b | \hat{x})$ and analogous for p_L . The outcome of a trial is thus choice d (L or R), decision confidence $c = p_d$ and predicted value Q_d .

After a reward r , the prediction error is given by $\delta = r - Q_d$. Choice values V are updated according to a learning rule: $V_d \leftarrow V_d + \alpha\delta$, where α is the learning rate. For simplicity, this model does not include temporal discounting. We ran this model for 3,000,000 trials with $\alpha = 0.5$, $r = 1$ for correct trials, initial action values $Q_L = Q_R = 0.5$, and producing a percept in each trial according to $\hat{x} \sim \mathcal{N}(x, \sigma)$ with $\sigma = 0.2$ (perceptual noise) and used the last 100,000 trials for plots.

Potential slow drifts in choice bias could lead to similar behavioral patterns in choice updating as observed here (Lak et al., 2020). Slow drifts in choice bias will be observed in a sequence of trials both before and after a rewarded trial and can therefore be isolated by calculating the choice bias conditioned on the subsequent trial's choice. In a control analysis quantifying the influence of a slow drift in choice bias across trials we did not observe strong or systematic slow drifts in choice bias in either modality transition (Figure S6J).

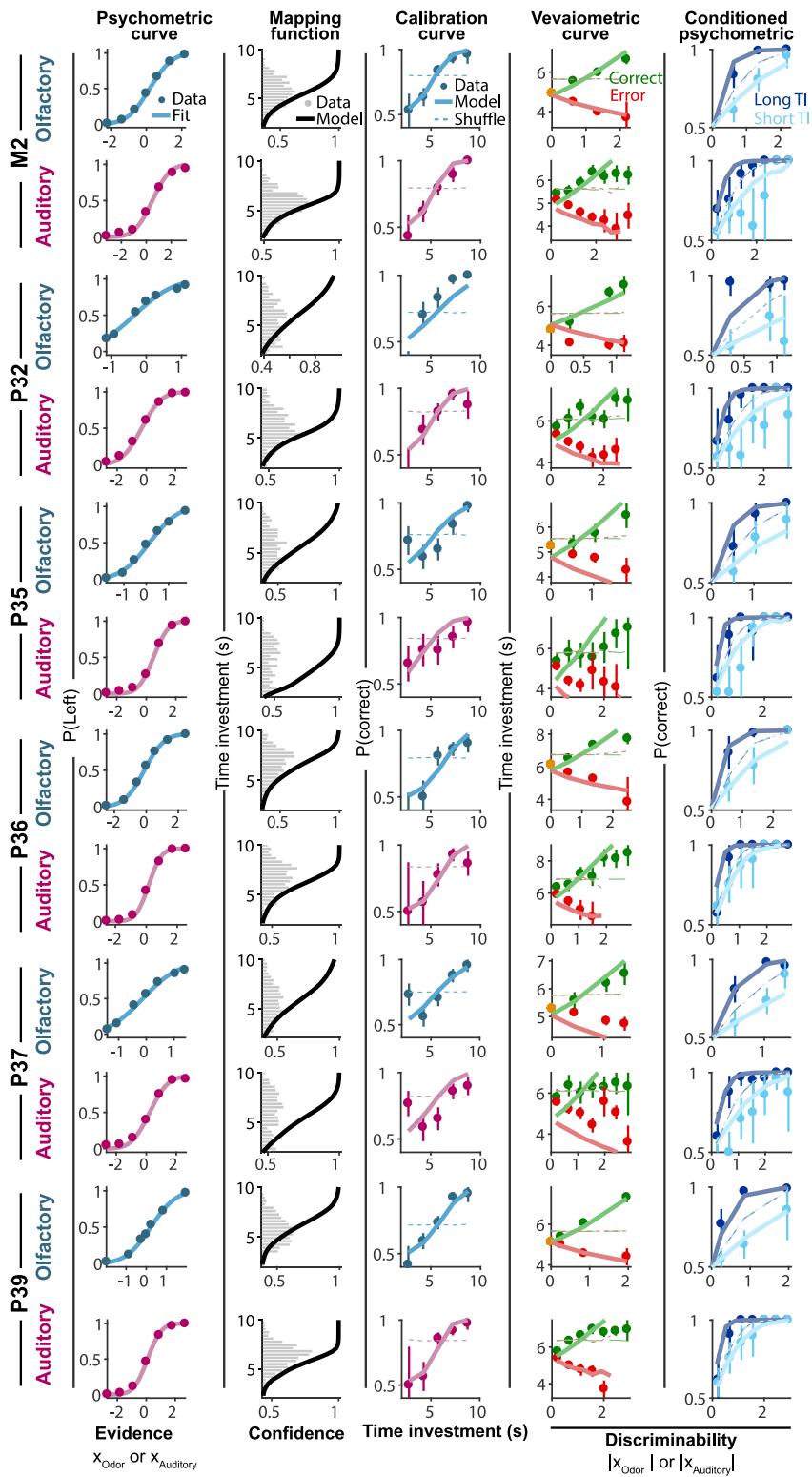
Additional interpretation of results

OFC has also been implicated in the control of impulsivity, compulsivity and addictive behaviors (Feil et al., 2010; Fineberg et al., 2010; Gillan and Robbins, 2014; Pascoli et al., 2018; Schoenbaum et al., 2006). It has been suggested that the role of OFC in reversal learning or delay discounting could mediate these effects (Chamberlain et al., 2008; Mar et al., 2011) and single unit recordings in both rodents and monkeys have shown that valuation due to reward size and time discounting are encoded in OFC (Roesch and Olson, 2005; Roesch et al., 2006). In our work, the activity in the later stages of the investment period (see Figures 7C, 7D, and S7E) supports a contribution of OFC to appropriately sustain confidence or value information from decision to outcome. Perturbation of this process could lead to inability to act on delayed rewards depending on the appropriate value or confidence. The observation that OFC inactivation impairs the ability of rats to invest time according to their confidence, while keeping the mean time investment unchanged (Lak et al., 2014), supports this view.

QUANTIFICATION AND STATISTICAL ANALYSIS

Error bars indicate mean \pm SEM unless otherwise noted. Statistical tests and associated p value are reported in the Results or associated Methods sections for data presented in the main Figures and in Supplemental Figure legends or associated Methods section for data presented in Supplemental Figures. All statistical tests were two-sided.

Supplemental Figures



(legend on next page)

Figure S1. Time Investment Reflects Signatures of Statistical Decision Confidence and Matches Statistical Predictions for All Rats and Modalities, Related to Figures 1 and 2

For each rat and sensory modality (rows), the 1st column shows the psychometric curve (points) and cumulative Gaussian model fit to estimate perceptual noise (line). The second column shows the estimated mapping function $m(c)$ from confidence to time investment (black line; based on cumulative distribution functions of empirical time investment and model-derived confidence; see [STAR Methods](#)). Gray histograms indicate the observed time investment distribution. The mapping function was based on the each rat's time investment distribution (gray bars) (see Methods). The third to fifth columns show key signatures of statistical confidence for data (points) and model prediction (thick lines). Dashed lines show model predictions with shuffled confidence values. The calibration curve (3rd column) shows that time investment predicted accuracy for all rats and modalities. The vevaiometric curve (4th columns) shows that time investment was positively correlated with evidence in correct trials, and negatively correlated with evidence in error trials. The conditioned psychometric curve (5th column) shows stimulus discriminability was high in long time investment trials compared to short time investment trials. Error bars represent SEMs across trials pooled from recording sessions.

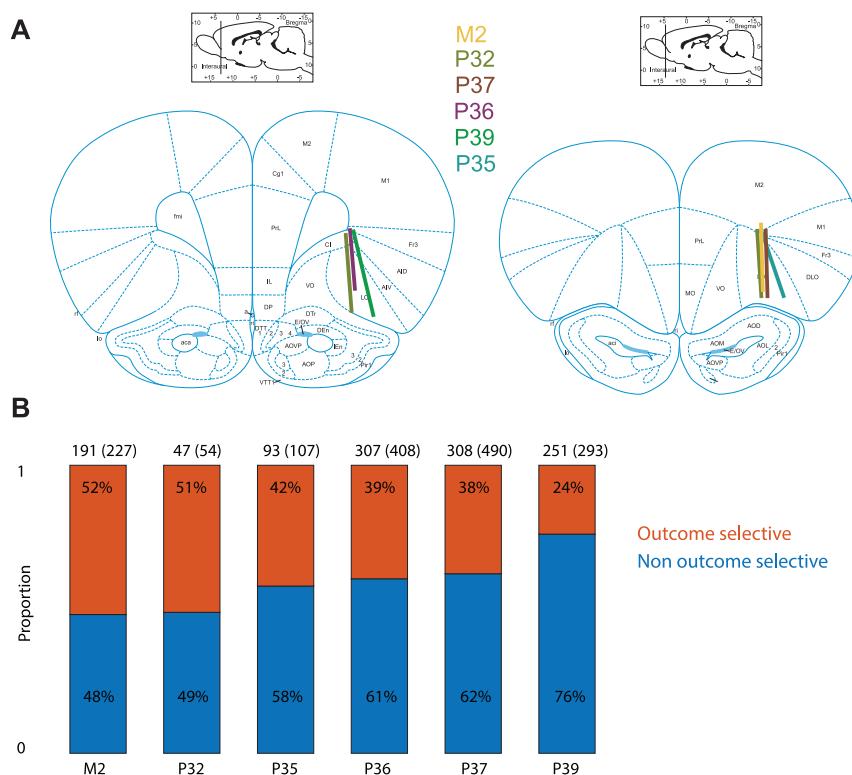
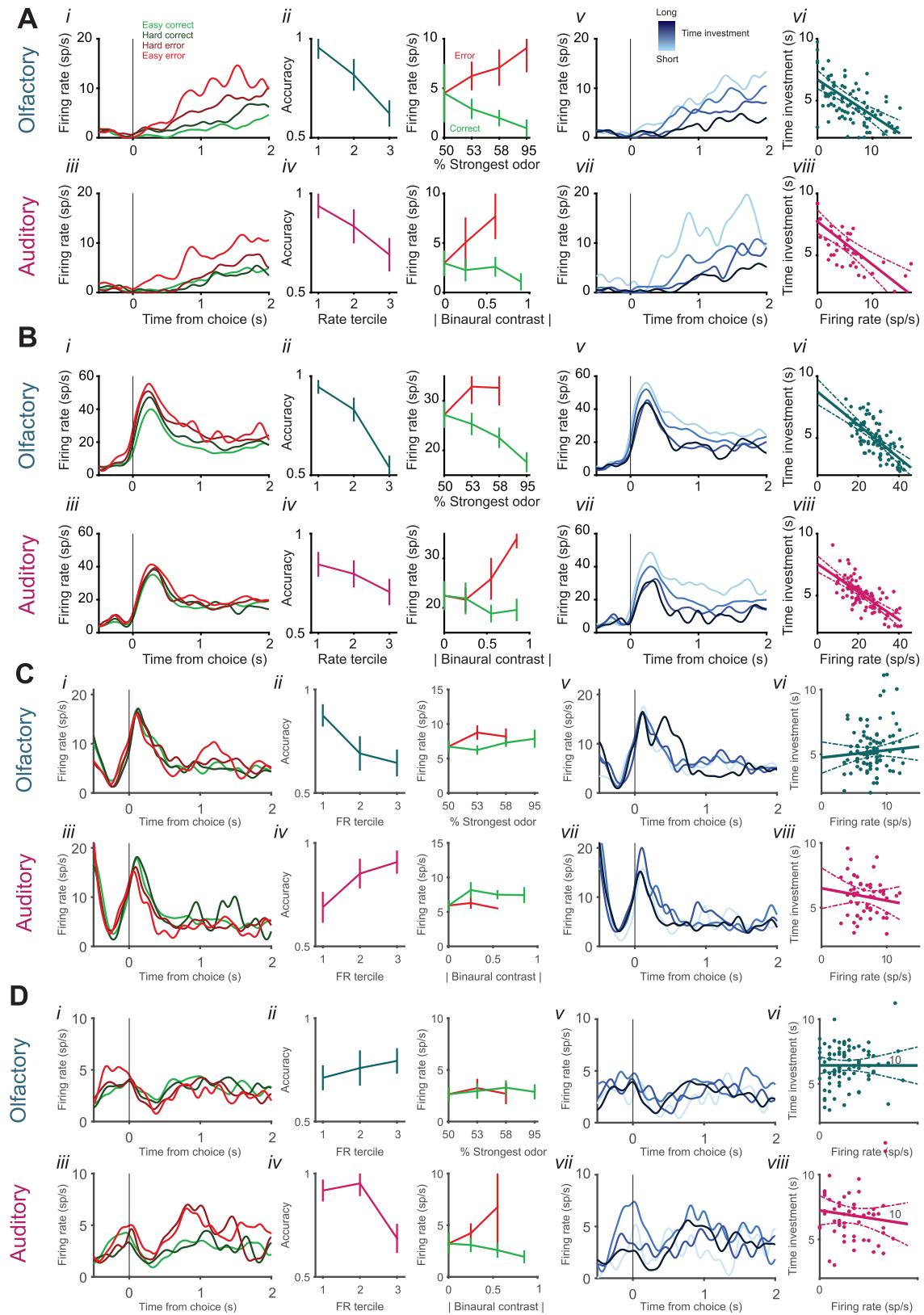


Figure S2. Tetrode Recordings in OFC, Location and Numbers of Neurons across Animals, Related to Figures 3, 4, 5, 6, and 7

(A) Average location of the tetrode bundle for each animal estimated using electrolytic lesions and/or fluorescent marker. Left panel represents the coronal section positioned at 3.7 mm anterior to Bregma. Right panel represents the coronal section at 4.7 mm anterior to Bregma (Paxinos and Watson, 2007).

(B) Proportion of outcome-predictive neurons per animal. Number on top of each bar corresponds to the number of neurons analyzed for the corresponding animal and the number in parentheses correspond to the total number of neurons (including neurons with firing rates below 1 s^{-1}).



(legend on next page)

Figure S3. Example Single OFC Neuron Responses Reflect Statistical Decision Confidence and Predict Time Investments, Related to Figures 3 and 4

(A) Negatively tuned confidence-encoding single neuron (i)-(viii). Same conventions as in (Figure 3Ai-iv and Figure 4Ai-iv) for different single neuron negatively tuned to decision confidence. Note that the neuron's activity follows a different time course and tuning curve patterns are inverted as compared to the neuron in Figure 3A. (vi) $r_{\text{olf}}(85) = -0.60$, $p = 1.2e^{-9}$; (vii) $r_{\text{aud}}(41) = -0.68$, $p = 5e^{-7}$, Pearson correlation, t test.

(B) Negatively tuned confidence-encoding single neuron (i)-(viii). Same conventions as in (Ai-viii) for different single neuron negatively tuned to decision confidence. Note that the neuron's activity follows a different time course and all tuning curve patterns are inverted as compared to the neuron in Figure 3A. (vi) $r_{\text{olf}}(89) = -0.71$, $p = 2.4e^{-15}$; (viii) $r_{\text{aud}}(88) = -0.74$, $p = 6.6e^{-17}$, Pearson correlation, t test.

(C) Single neuron incoherently tuned to decision confidence (negative in olfactory trials; positive in auditory trials). Figure conventions as in Figure 3A. (vi) $r_{\text{olf}}(91) = 0.1$, $p = 0.33$; (viii) $r_{\text{aud}}(54) = -0.15$, $p = 0.29$, Pearson correlation, t test.

(D) Single neuron incoherently tuned to decision confidence (positive in olfactory trials; negative in auditory trials). Figure conventions as in Figure 3A. (vi) $r_{\text{olf}}(87) = -0.002$, $p = 0.99$; (viii) $r_{\text{aud}}(57) = -0.14$, $p = 0.30$, Pearson correlation.

Error bar represent SEM across trials. Error lines represent 95% confidence intervals for fitted linear model.

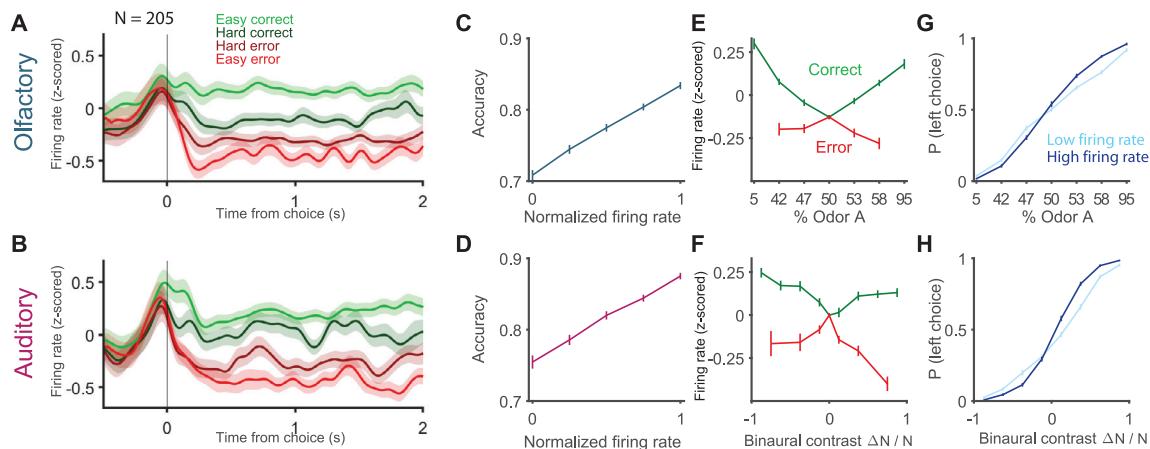


Figure S4. Positively Tuned Neuronal Responses during Outcome Anticipation Reflect Statistical Decision Confidence (Confidence⁽⁺⁾ Neurons), Related to Figure 3

(A) Average activity of positive outcome-predictive – confidence(+) – population ($n = 205$) aligned to choice port entries for olfactory trials. Green traces, correct trials; red traces, error trials; brightness, evidence strength; shaded area, SEM across neurons.

(B) Same conventions as in (A) for auditory trials.

(C) Population calibration curve for olfactory trials. Firing rates predict accuracy (cf. Figure 1E).

(D) Calibration curve for auditory trials (same layout as in (C)).

(E) Population vevaiometric curve for olfactory trials reflecting the pattern expected by statistical decision confidence (cf. Figure 1F). Green trace, correct trials; red trace, error trials.

(F) Population vevaiometric curve for auditory trials (same conventions as in (E)).

(G) Population conditioned psychometric curve for olfactory trials. Psychometric curve is steeper for high firing rate trials (dark blue) as compared to low firing rate trials (light blue). For a fixed difficulty level, average accuracy is higher in high-firing rate trials as predicted by statistical decision confidence (cf. Figure 1G). Difference between low firing rate trials and high firing rate trials for %Odor A = 42,47,53,58, $p < 0.008$ and for absolute value of binaural contrast in the bins 0.75-0.5, 0.5-0.25 $p < 0.01$ (two tailed t test).

(H) Population conditioned psychometric curve for auditory trials (same conventions as in (G)).

Error bars represent SEM across neurons.

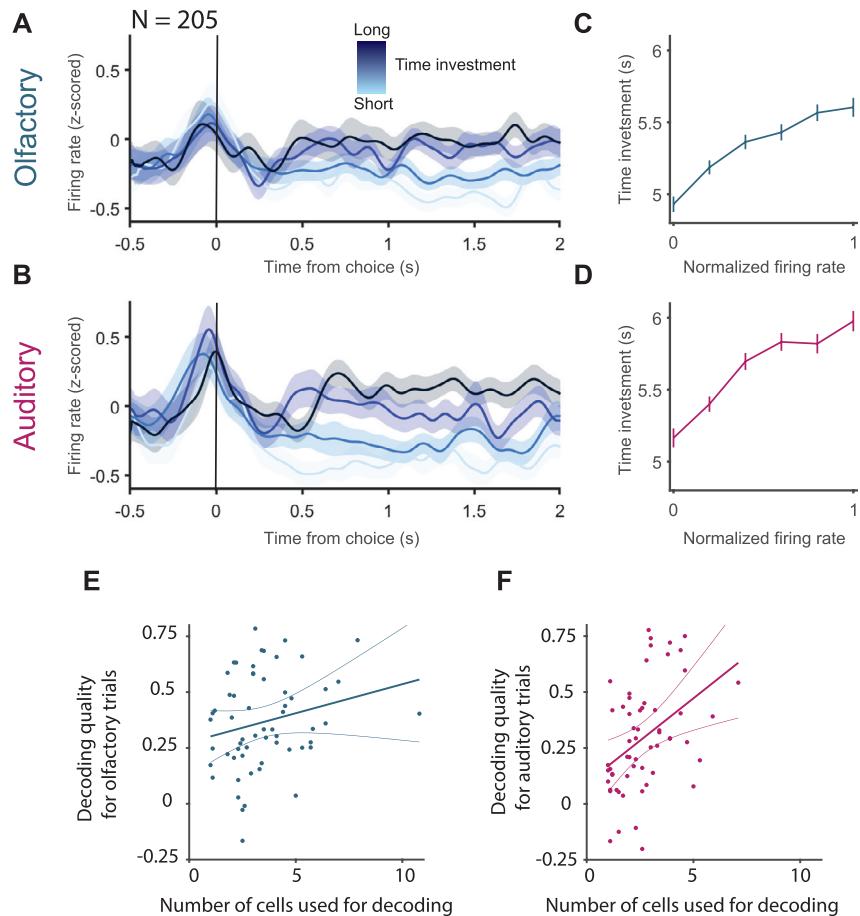


Figure S5. Positively Tuned Neuronal Responses during Outcome Anticipation Predict Time Investment Decisions (Confidence⁽⁺⁾ Neurons), Related to Figures 4 and 5

(A) Average activity of positive outcome-predictive – confidence(+) – population ($n = 205$) sorted by the rats' time investment in olfactory trials aligned to choice port entries. After the rats made a choice, population activity increased proportionally to the amount of time rats invested in a decision. Blue gradient, time investment. Shaded area, SEM across neurons.

(B) Average activity for positively tuned outcome-predictive neurons in auditory trials (same conventions as in (A)).

(C) Population activity predicts time investment in olfactory trials. Error bars represent SEM across neurons.

(D) Population activity predicts time investment in auditory trials (same conventions as in (C)).

(E) Decoding quality tends to increase with the mean number of cells used in the decoding in a given session for olfactory trials (Spearman's $r(57) = 0.25$, $p = 0.06$, t test, error lines represent 95% confidence intervals for fitted linear model).

(F) Same layout as in panel E for auditory trials (Spearman's $r(57) = 0.46$, $p = 2.1e-4$, t test, error lines represent 95% confidence intervals for fitted linear model).

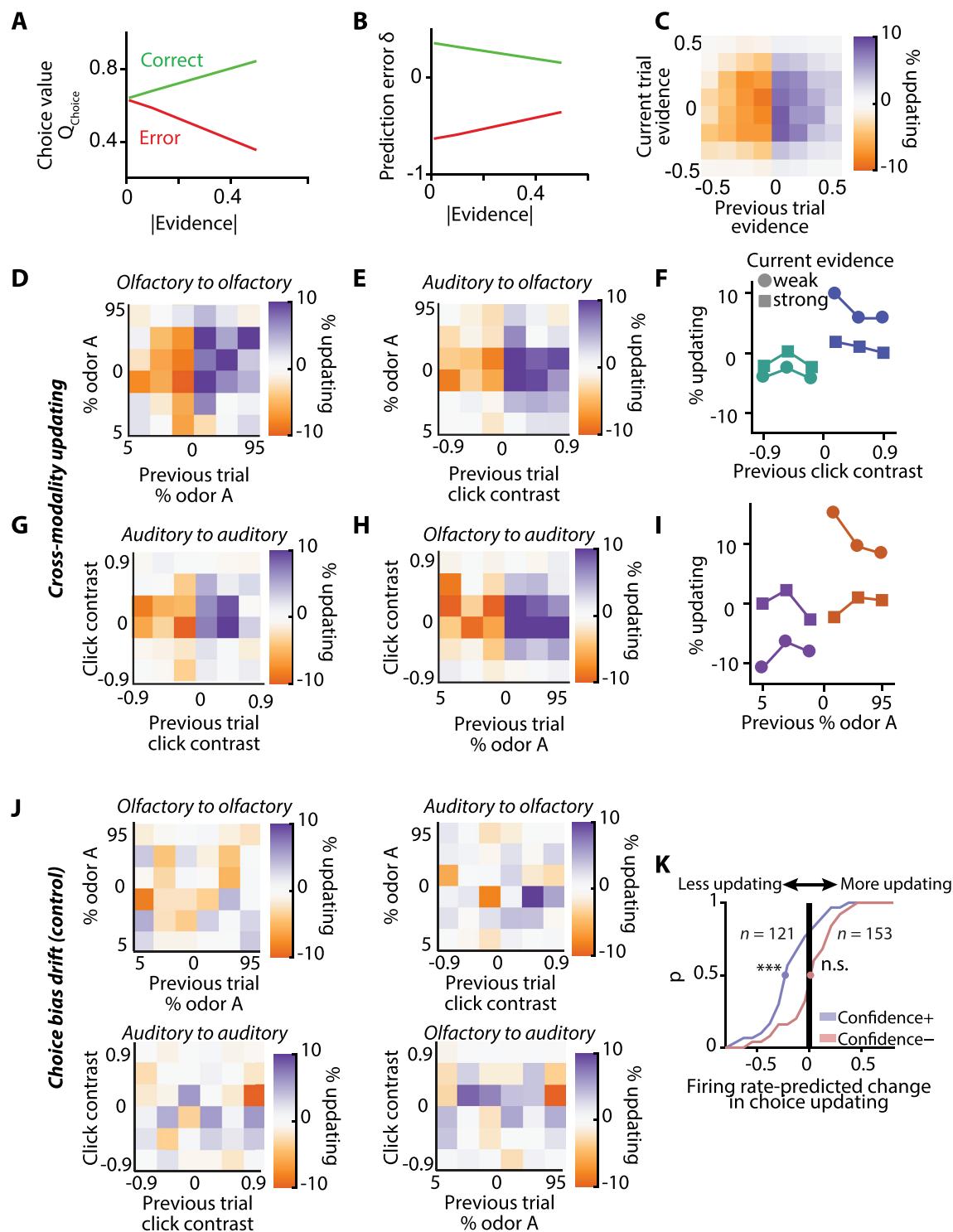


Figure S6. Belief-State Reinforcement Learning Model for Choice Strategy Updating, Related to Figure 6

(A) Belief state reinforcement learning model (cf. Figure 6A) produces confidence-scaled action values Q_{choice} .

(B) Prediction errors, $\delta = r - Q_{\text{choice}}$, show flipped confidence signatures leading to confidence-scaled updating of V_L and V_R .

(C) The model agent's magnitude of choice updating (shift in the psychometric function) after a correct trial for both current evidence (vertical axis) and previous evidence (horizontal axis) shows that choice updating is highest after difficult trials if current evidence is weak (center area).

(D) Choice updating (shift in psychometric function) as a function of current and previous evidence for two consecutive olfactory trials (within-modality updating).

(legend continued on next page)

-
- (E) Same conventions as in D for olfactory decisions followed by auditory decisions (cross-modality updating).
- (F) Choice updating as a function of previous evidence for olfactory decisions followed by auditory decisions (cross-modality updating).
- (G) Same conventions as in D for two consecutive auditory decisions (within-modality updating).
- (H) Same conventions as in D for auditory decisions followed by olfactory decisions (cross-modality updating).
- (I) Choice updating as a function of previous evidence for auditory decisions followed by olfactory decisions (cross-modality updating).
- (J) Control analysis quantifying the influence of a slow drift in choice bias across trials. Slow drift will be observed in a sequence of trials both before and after a rewarded trial and can therefore be isolated by calculating the choice bias conditioned on the subsequent trial's choice. We did not observe strong or systematic slow drifts in choice bias in either modality transition (sub panels).
- (K) Cumulative distribution of the change in choice updating predicted by confidence neurons firing rates in the previous trials (choice bias for high versus low firing rate trials). At the population level, only confidence(+) neurons predict choice updating: Higher firing rate predicts less choice updating ($p < 0.001$, rank sum test). Confidence(-) neurons do not predict choice updating ($p = 0.2$, rank sum test).
- *** $p < 0.001$, n.s. $p > 0.05$ (rank sum test).

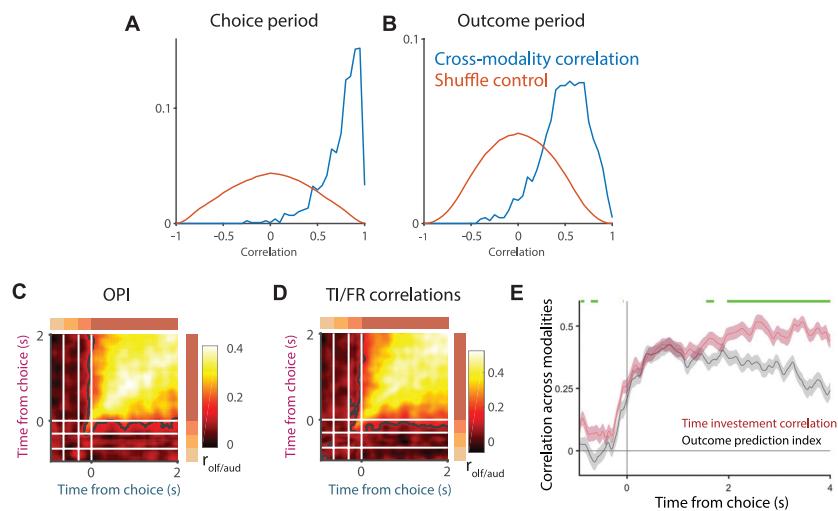


Figure S7. OFC Activity and Encoding Dynamics Are Shared across Sensory Modalities, Related to Figure 7

(A) Distribution of Pearson correlation between the spike density function for olfactory and auditory trials for the choice period (see Figures 7A and 7B). Blue: distribution for correlations within neurons; orange: shuffle control, distribution for correlations across neurons.

(B) Same conventions as in (B) for the correlations in the outcome period (see Figures 7C and 7D).

(C) Time course of OPI correlation across sensory modalities. For each time point in (Figures 7E and 7F), we computed the correlation coefficient of OPIs across modalities. The diagonal line can be understood as the correlation between OPIs across modalities in the respective time bin. Off-diagonal entries quantify cross-temporal correlation of OPIs. Black line, area denoting correlation coefficients with $p < 0.05$, t test Bonferroni corrected for multiple comparisons.

(D) Time course of cross-modality correlation of firing rate/time investment correlations. For each time point in Figures 7G and 7H, we computed the correlation coefficient of firing rate/time investment correlations across sensory modalities. Heatmap conventions as in (C).

(E) Time courses of the correlation across modalities of OPIs (see Figure S7C) and time investment correlations (see Figure S7D). Strength and time course of observed correlations was similar for both OPIs and time investment correlations in the first two seconds of the time investment period and remained higher for the time investment correlations into the later stages of the investment period. Green bar represents times points with normalized auROC > 0.975 between the distributions of the correlations for subsampled neuronal populations. Shaded areas represent \pm SEM across neurons (bootstrap, see Methods).