nature neuroscience

Neuronal remapping and circuit persistence in economic decisions

Jue Xie[1] & Camillo Padoa-Schioppa[1–3]

The orbitofrontal cortex plays a central role in good-based economic decisions. When subjects make choices, neurons in this region represent the identities and values of offered and chosen goods. Notably, choices in different behavioral contexts may involve a potentially infinite variety of goods. Thus a fundamental question concerns the stability versus flexibility of the decision circuit. Here we show in rhesus monkeys that neurons encoding the identity or the subjective value of particular goods in a given context 'remap' and become associated with different goods when the context changes. At the same time, the overall organization of the decision circuit and the function of individual cells remain stable across contexts. In particular, two neurons supporting the same decision in one context also support the same decision in different contexts. These results demonstrate how the same neural circuit can underlie economic decisions involving a large variety of goods.

Clinical[1–3] and lesion[4–6] studies indicate that the orbitofrontal cortex (OFC) is necessary for economic choice behavior. Neurons in this area encode the value of individual goods[7–11], the identity of the choice outcome[8,12–14] and the chosen value[8,15–17]. These signals capture both the input and the output of a good-based decision, suggesting that choices between goods emerge from a decision circuit within the OFC[18,19]. In this perspective, a fundamental question concerns the stability versus flexibility of the circuit. Choices in different behavioral contexts may involve a potentially infinite variety of goods. To handle this enormous variability, the decision circuit must adapt to represent the goods available for choice at any given time. However, to generate effective decisions in different behavioral contexts, the overall organization of the decision circuit should persist across contexts.

To shed light onto how the decision circuit reorganizes across behavioral contexts, we examined the activity of neurons in the OFC while monkeys executed economic decisions between different sets of goods. We found that neurons encoding the identity or the subjective value of a particular good in a given context remapped and became associated with different goods when the context changed. At the same time, the overall organization of the decision circuit and the functional role of individual cells remained stable across contexts. Specifically, two neurons supporting the same (or opposite) decision in one context also supported the same (or opposite) decision, respectively, in different contexts. In other words, neuronal pools persisted across behavioral contexts. Our results challenge the understanding that neuronal responses in OFC are primarily driven by sensory features and demonstrate how the same neural circuit can underlie economic decisions involving a large variety of goods.

RESULTS

In the experiments, rhesus monkeys chose between different juices offered in variable amounts. Offers were represented by symbols on a computer monitor and the animals indicated their choices with an eye movement. Previous work in similar conditions shows that different groups of cells in OFC encode the subjective value of individual goods (Offer Value), the binary choice outcome (Chosen Juice) and the subjective value of the chosen good (Chosen Value)[8,20]. For each of these variables, the slope of the encoding can be positive (higher firing rates for higher values) or negative (higher firing rates for lower values). In the present study we examined whether and how neuronal representations in the OFC adapted when the goods available for choice changed. Here A, B, C and D indicate four different juice types and "X:Y" indicates choices between juices X and Y, with X preferred to Y. Each recording session consisted of two blocks of trials (**Fig. 1**). In experimental sessions, different juice pairs were offered in the two blocks (A:B, C:D design). We recorded the activity of neurons in the central OFC and analyzed it in multiple time windows (Online Methods).

Encoding stability and neuronal remapping

Inspection of individual cells revealed that most neurons encoded the same variable in the two trial blocks. We illustrate five examples (**Fig. 2**). One cell (**Fig. 2a**) encoded the Offer Value of kiwi punch in the first block; the same cell encoded the Offer Value of apple juice in the second block. Similarly, another cell (**Fig. 2b**) encoded the Offer Value of peach juice in the first block and the Offer Value of cherry juice in the second block. A third cell (**Fig. 2c**) encoded the Chosen Juice in both blocks. In the first block, the activity was high when the animal chose lemon Kool-Aid; in the second block, the activity was high when the animal chose grape juice. Finally, two cells (**Fig. 2d**,**e**) encoded the Chosen Value in both trial blocks. The slope of the encoding was positive for one neuron (**Fig. 2d**) and negative for the other (**Fig. 2e**). Note that for Offer Value cells (**Fig. 2a**,**b**) and Chosen Juice cells (**Fig. 2c**), encoding the same variable in both trial blocks

[1]Department of Neuroscience, Washington University in St. Louis, St. Louis, Missouri, USA. [2]Department of Economics, Washington University in St. Louis, St. Louis, Missouri, USA. [3]Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, Missouri, USA. Correspondence should be addressed to C.P.-S. (camillo@wustl.edu).

Received 4 February; accepted 4 April; published online 9 May 2016; doi:10.1038/nn.4300

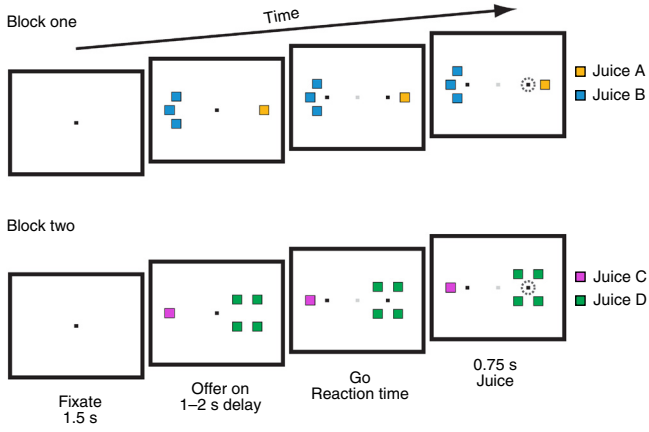


Figure 1 Experimental design. The top and bottom row show the time course of a trial in the first and second block of trials, respectively. The two offers are represented by sets of colored squares, with the color indicating the juice type and the number of squares indicating the juice quantity (Online Methods).

meant that these neurons were associated with a particular juice in the first block and with a different juice in the second block.

Our data set included 718 neurons recorded in experimental sessions (A:B, C:D design) and 387 neurons recorded in control sessions, where the same juice pair was offered in both trial blocks (A:B, A:B design). For a population analysis, we proceeded as follows. Trial types were defined by two offers and a choice (for example, [1A:2B, A]). For each cell and each time window, we conducted a 1-way ANOVA ($P < 0.01$) across trial types separately for each block. Neurons that passed the criterion in at least one block in any time window were identified as task-related and included in further analyses (504 cells in experimental sessions, 255 cells in control sessions). The activity of each cell in each time window was regressed against variables Offer Value, Chosen Value and Chosen Juice. The variable encoded by the cell was identified as that which provided the highest total R^2 across time windows, and the cell was classified as Untuned if it was not explained by any variable (Online Methods). Each neuron in each trial block was thus assigned to one of six groups depending on the encoded variable and on the sign of the encoding: Offer Value+, Offer Value–, Chosen Value+, Chosen Value–, Chosen Juice and Untuned, where "+" and "–" indicate whether the variable was encoded with a positive or negative slope.

Figure 3a depicts the contingency table obtained for the population of control cells (A:B, A:B design). Neurons tend to concentrate on the main diagonal, indicating that they encoded the same variable with the same sign in both trial blocks. This result was expected because the two trial blocks were essentially identical in control sessions.

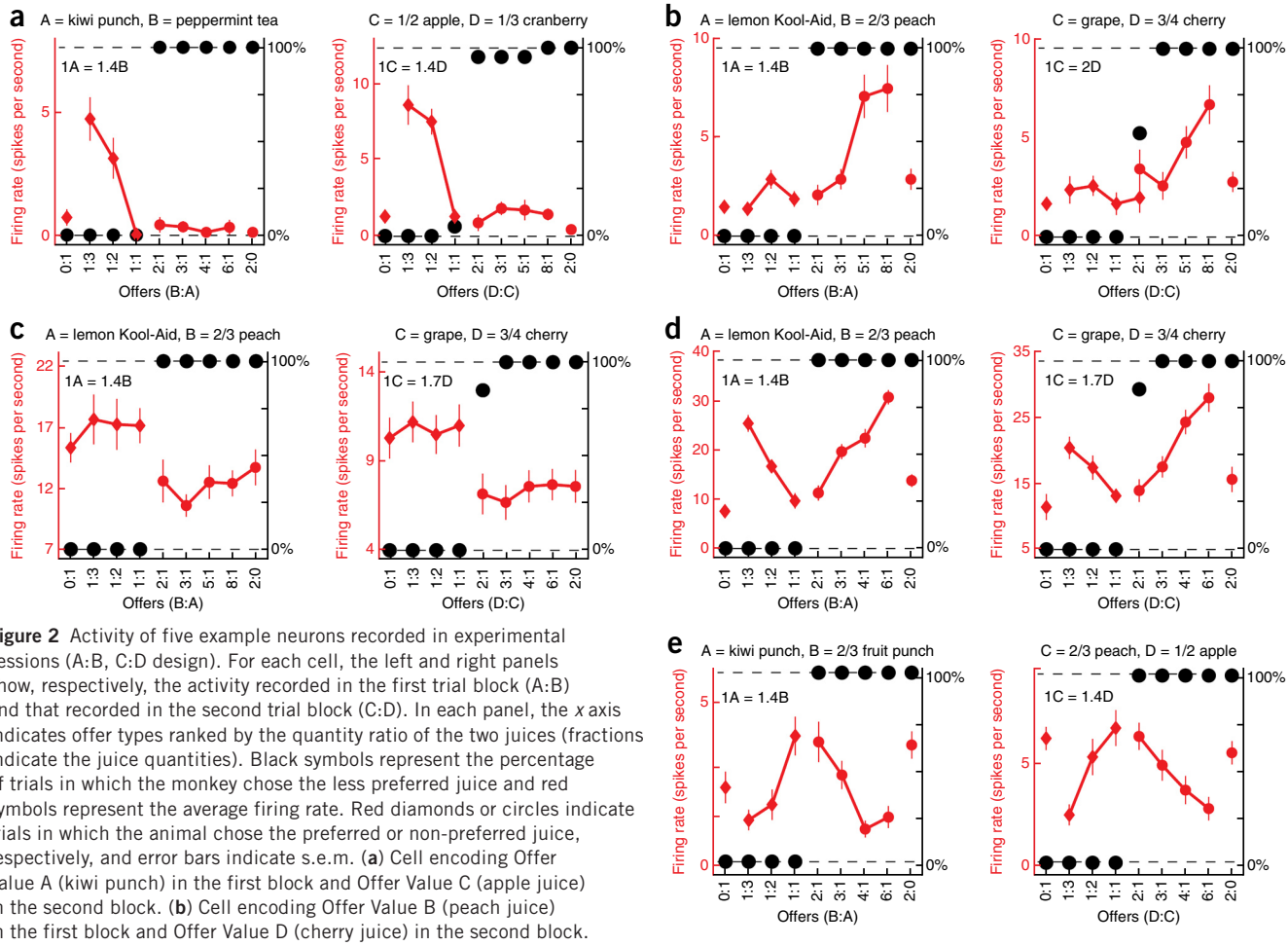


Figure 2 Activity of five example neurons recorded in experimental sessions (A:B, C:D design). For each cell, the left and right panels show, respectively, the activity recorded in the first trial block (A:B) and that recorded in the second trial block (C:D). In each panel, the *x* axis indicates offer types ranked by the quantity ratio of the two juices (fractions indicate the juice quantities). Black symbols represent the percentage of trials in which the monkey chose the less preferred juice and red symbols represent the average firing rate. Red diamonds or circles indicate trials in which the animal chose the preferred or non-preferred juice, respectively, and error bars indicate s.e.m. (**a**) Cell encoding Offer Value A (kiwi punch) in the first block and Offer Value C (apple juice) in the second block. (**b**) Cell encoding Offer Value B (peach juice) in the first block and Offer Value D (cherry juice) in the second block. (**c**) Cell encoding Chosen Juice A (lemon Kool-Aid) in the first block and Chosen Juice C (grape juice) in the second trial block. (**d**,**e**) Cells encoding the Chosen Value in both trial blocks, with a positive slope (**d**) and with a negative slope (**e**). Firing rates shown here are from the following time windows: (**a**) post-offer, (**b**) post-juice, (**c**) post-juice2, (**d**) post-juice2, (**e**) late delay (Online Methods). For these neurons, the numbers of trials in each trial block were as follows: (**a**) 179, 179, (**b**) 180, 180, (**c**) 180, 180, (**d**) 180, 180, (**e**) 200, 181.

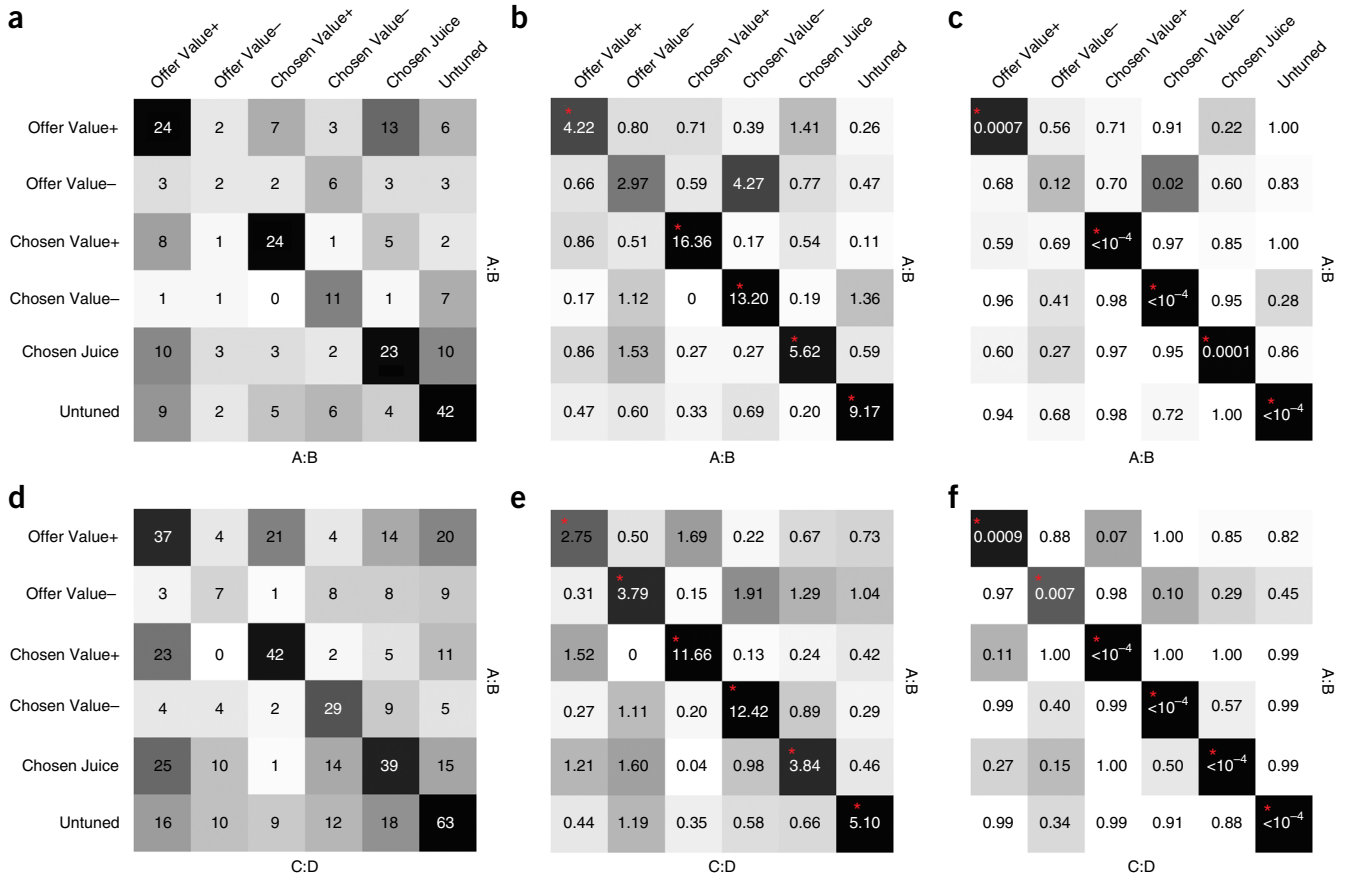


Figure 3 Comparing classifications across trial blocks. This population analysis was conducted separately for control cells (**a**–**c**) and experimental cells (**d**–**f**). (**a**) Contingency table, control cells ($n = 255$). Rows and columns represent, respectively, the classification obtained in the first and in the second block of trials. Numbers in the table and grayscale colors indicate cell counts. For Offer Value and Chosen Value cells, the slope of the encoding could be positive (+) or negative (–). Task-related cells not explained by any variable were classified as Untuned. (**b**) Analysis of odds ratios, control cells. Numbers in the table and grayscale colors represent the odds ratios obtained for each location in **a**. Chance level is 1 and numbers >1 (or <1) indicate that the cell count was above (or below) that expected by chance. For each location, we performed Fisher's exact test. Red * indicate that the cell count was significantly above chance ($P < 0.01$). Cell counts significantly below chance are not indicated here. However, **Supplementary Figure 6a** shows all the exact P values. (**c**) Bootstrap analysis, control cells. Numbers and grayscale colors represent P values obtained from the bootstrap analysis (colors are in log scale). Red * indicate that the cell count was significantly above chance ($P < 0.01$). (**d**) Contingency table, experimental cells ($n = 504$). Same format as in **a**. Rows and columns represent the classification obtained in the A:B block and in the C:D block, respectively. (**e**) Analysis of odds ratios, experimental cells. Same format as in **b**. **Supplementary Figure 6b** shows the exact P values. (**f**) Bootstrap analysis, experimental cells. Same format as in **c**.

However, some cells fell outside the main diagonal, indicating that our classification had some degree of inaccuracy (Online Methods). To assess whether the prevalence of neurons on the main diagonal was statistically significant, we conducted two statistical tests, based on odds ratios (**Fig. 3b**) and on a bootstrap analysis (**Fig. 3c**). Both tests assessed, for each position in the contingency table, whether the cell count deviated from chance level, assuming that the classifications in the two trial blocks were independent. Both tests provided the same results: cell counts for control sessions were significantly above chance ($P < 0.01$) only for positions on the main diagonal.

Figure 3d depicts the contingency table obtained for experimental cells (A:B, C:D design). Again, neurons tend to concentrate on the main diagonal, indicating that they encoded the same variable with the same sign in both trial blocks. In this case, however, the result was not foregone. Consider neurons encoding the Offer Value in the first block. In principle, when the animal progresses to the second block, the entire circuit could reorganize itself such that 'old' Offer Value cells are randomly reassigned to any variable. Alternatively, Offer Value cells associated with a particular juice no longer available (A or B) could exit the pool and become temporarily Untuned, while other cells, previously Untuned, become associated with one of the current juices (C or D). In contrast with these plausible scenarios, OFC neurons generally encoded the same variable in the two trial blocks (an element of stability). Thus Offer Value cells and Chosen Juice cells became associated with new juices in the new trial block (an element of flexibility). The analysis of odds ratios (**Fig. 3e**) and the bootstrap analysis (**Fig. 3f**) both indicated that cell counts were significantly above chance only on the main diagonal ($P < 0.01$). Note that experimental cells classified as Offer Value or Chosen Juice in one block were not more likely than chance to be classified as Untuned in the other trial block (**Fig. 3e**,**f**). Several control analyses confirmed these results (**Supplementary Figs. 1**–**4**).

Contextual remapping is essentially complete

In the analyses illustrated in **Figure 3**, the null hypothesis represented the scenario in which the classifications obtained in the two trial blocks were independent. The results demonstrated that they were not and this observation held true for both experimental and control cells. We designed a separate set of analyses to contrast the results obtained in the experimental condition (A:B, C:D design) with

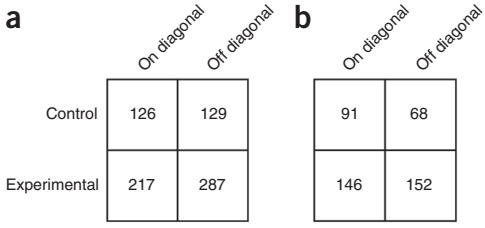


Figure 4 Classification consistency in experimental and control cells. (**a**) Cross-classification of all task-related cells ($n = 759$). Numbers indicate cell counts. For each row, the two columns indicate the total number of neurons on and off the main diagonal in the contingency tables shown in **Figure 3a**,**d**. (**b**) Cross-classification obtained excluding cells classified as Chosen Value in either block ($n = 457$). Same format as in **a**.

those obtained in the control condition (A:B, A:B design). In other words, we examined whether the classification patterns obtained in the experimental condition differed from the benchmark provided by the control condition. We performed several tests.

First, we examined whether the frequency of neurons presenting consistent classifications differed significantly across conditions. To do so, we considered the number of cells located on and off the main diagonal in the contingency table, separately for experimental (**Fig. 3d**) and control (**Fig. 3a**) cells. We thus obtained a 2 × 2 table (**Fig. 4a**). A statistical test failed to find any significant difference between the two conditions (odds ratio = 1.29; $P = 0.11$, Fisher's exact test). It may be argued that Chosen Value cells should respond similarly in the two trial blocks even if changing juices affected the other groups of neurons. Thus we repeated this analysis, excluding cells classified as Chosen Value in either trial block (**Fig. 4b**). The results were essentially identical (odds ratio = 1.39; $P = 0.10$, Fisher's exact test).

Second, we directly compared the contingency tables obtained for the two conditions (experimental and control). To do so, we first constructed for each condition a reduced 4 × 4 contingency table by pooling cells encoding the same variable with opposite signs. Thus cells in each block were assigned to one of four groups: Offer Value, Chosen Value, Chosen Juice and Untuned (**Fig. 5a**,**c**). (This reduction was performed because the chi-square test requires that the expected cell counts be sufficiently large[21].) We then concatenated the reduced contingency tables obtained for experimental and control conditions to obtain a 4 × 4 × 2 table. A chi-square test of joint independence[21] indicated that the first two dimensions of this table were independent of the third dimension ($P = 0.7$, $\chi^2 = 11.77$, d.f. = 15). In other words, the patterns of classification obtained across trial blocks did not differ between control and experimental conditions.

Third, we conducted the odds ratio analysis on the reduced contingency tables described above (**Fig. 5b**,**d**). Replicating the results obtained for the full contingency tables (**Fig. 3**), cell counts were significantly above chance ($P < 0.01$) only on the main diagonal. We then performed element-wise comparisons between the odds ratios obtained in the two conditions[21]. Each odds ratio can be thought of as quantifying the strength of association between two particular variables across the two trial blocks. Thus we examined whether these association strengths depend on the condition (experimental versus control). To do so, we first estimated the common odds ratios across conditions[21]. Then we computed, for each condition, the expected cell counts based on the common odds ratios. Finally, we compared the empirical cell counts with the expected cell counts using Breslow-Day statistics for homogeneous associations[21]. **Figure 5e** illustrates the P values obtained from the Breslow-Day tests. For most locations on the table, the association strength did not differ significantly between the two conditions. Locations corresponding to associations [Offer Value, Untuned] and [Untuned, Chosen Juice] appeared to depart from chance level ($P = 0.02$ and $P = 0.04$, respectively). However, these effects did not reach statistical significance once we accounted for multiple comparisons. An additional analysis of conditional odds ratios[21] confirmed this point (**Supplementary Fig. 5**).

To summarize, we did not find any systematic difference between the classification patterns obtained in the experimental and control

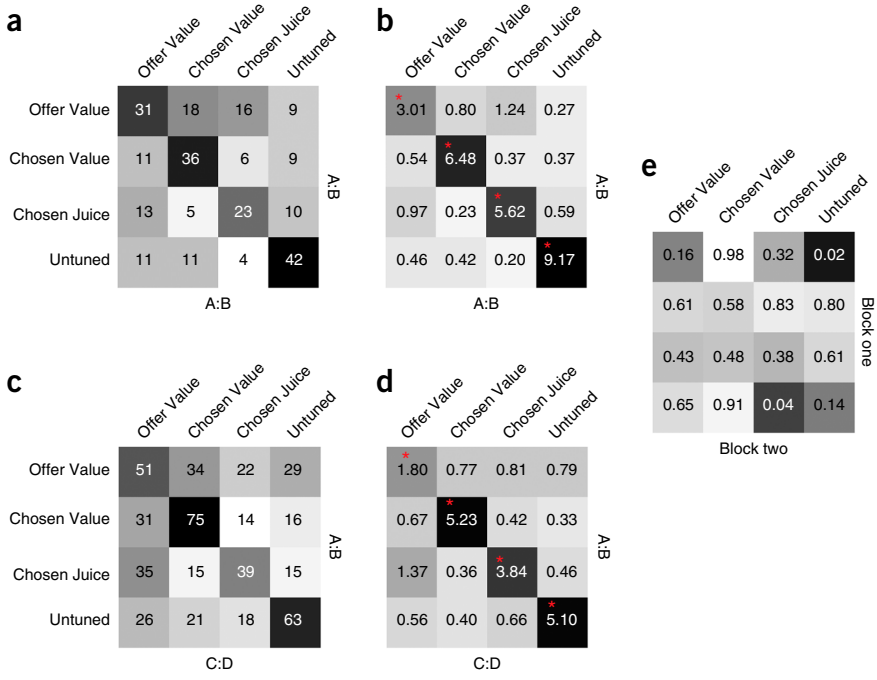


Figure 5 Direct comparison of experimental and control conditions. (**a**) Reduced contingency table, control cells ($n = 255$). Rows and columns represent, respectively, the classification obtained in the first and in the second trial block. Numbers in the table and grayscale colors indicate cell counts. Cells encoding the same variable with opposite signs were pooled (see Results). (**b**) Analysis of odds ratios, control cells. Numbers in the table and grayscale colors represent the odds ratios obtained for the corresponding location in **a**. Numbers >1 (or <1) indicate that the cell count was above (or below) that expected by chance. Red * indicate that the cell count was significantly above chance ($P < 0.01$, Fisher's exact test). Exact P values are shown in **Supplementary Figure 6c**. (**c**) Reduced contingency table, experimental cells ($n = 504$). Same format as in **a**. (**d**) Analysis of odds ratios, experimental cells. Same format as in **b**. Exact P values are shown in **Supplementary Figure 6d**. (**e**) Results of element-wise comparisons of **b** and **d**. Numbers and grayscale colors indicate the P values obtained from the Breslow-Day tests (d.f. = 1; colors are in log scale). For this test, the null hypothesis represents the scenario under which the two odds ratios in corresponding positions are drawn from the same distribution. The test is performed after computing a common odds ratio across conditions[21]. Chi-square values are shown in **Supplementary Figure 7**.

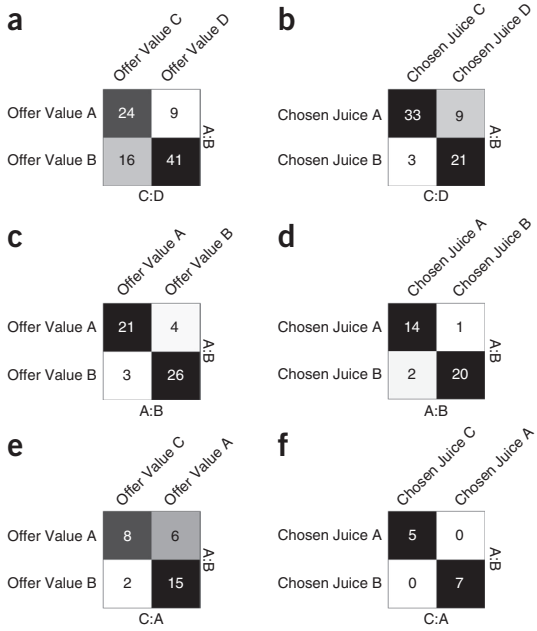


Figure 6 Persistence of neuronal pools. (**a**) Cross-classification of Offer Value cells, experimental condition. For each Offer Value cell, we identified the juice encoded in the A:B block (rows) and that encoded in the C:D block (columns). Numbers in the table indicate cell counts. Most neurons are on the main diagonal. In other words, most cells encoding the Offer Value A came to encode the Offer Value C, while most cells encoding the Offer Value B came to encode the Offer Value D (for single-cell examples, see **Fig. 2a**,**b**). (**b**) Cross-classification of Chosen Juice cells, experimental condition (for a single-cell example, see **Fig. 2c**). (**c**,**d**) Cross-classification obtained in the control condition (**c**, Offer Value cells; **d**, Chosen Juice cells). Same format as in panels **a** and **b**. (**e**,**f**) Cross-classification obtained with the A:B, C:A design (**e**, Offer Value cells; **f**, Chosen Juice cells). Same format as in panels **a** and **b**. In every panel, the cell count on the main diagonal was significantly higher than chance. Exact *P* values were as follows: **a**, $P = 5.8 \times 10^{-5}$; **b**, $P = 2.8 \times 10^{-7}$; **c**, $P = 4.2 \times 10^{-8}$; **d**, $P = 2.7 \times 10^{-7}$; **e**, $P = 0.018$; **f**, $P = 0.0013$ (Fisher's exact test).

conditions. This result implies that the classification patterns measured in the experimental condition were as consistent as one could expect, given that classification procedures bear some degree of inaccuracy. Thus the neuronal remapping from one set of juices to the next was essentially complete.

Persistence of neuronal pools

Consider neurons recorded in experimental sessions (A:B, C:D design). The results discussed so far indicate that Offer Value cells and Chosen Juice cells associated with one particular juice in the first trial block became associated with a different juice in the second trial block. These cells maintained their function in the decision circuit, but they remapped onto one of the goods available in the current behavioral context. Importantly, two different juices were offered in each trial block. If the decision circuit is indeed stable, one would expect not only that individual cells encode the same variable in the two blocks, but also that neuronal pools persists across trial blocks. In other words, two neurons 'supporting' the same decision in the first block should support the same decision also in the second block. Our data confirmed this prediction.

To examine the composition of neuronal pools, we identified for each Offer Value cell the juice encoded in each trial block. Most neurons encoding the Offer Value of the preferred juice (juice A) in the first block also encoded the Offer Value of the preferred juice (juice C) in the second block, while most neurons encoding the Offer Value of the non-preferred juice (juice B) in the first block also encoded the Offer Value of the non-preferred juice (juice D) in the second block (**Fig. 6a**). The analysis of Chosen Juice cells provided similar results (**Fig. 6b**) and similar results were also obtained for control cells (**Fig. 6c**,**d**). In all these cases, most cells remapped from the preferred (or non-preferred) juice in the first block to the preferred (or non-preferred) juice in the second block (all $P < 0.001$, Fisher's exact test). For control sessions, this result was expected. For experimental sessions, however, this result was noteworthy and consequential. It implied that neuronal pools within the decision circuit remained stable across behavioral contexts.

OFC has often been discussed in relation to its sensory inputs, especially gustatory and olfactory inputs[22–26]. If neuronal responses in this area were primarily driven by sensory features, we would expect relatively stable associations between individual neurons and specific juices in our task. In contrast, our results indicate a high degree of flexibility in the neural representation. To further explore this point, we conducted a follow-up experiment in which one juice was offered in both trial blocks but with different preference ranking (A:B, C:A design; with A preferred to B and C preferred to A). We recorded and analyzed the activity of 329 cells (161 task-related cells). Surprisingly, Offer Value cells recorded in these conditions typically remapped according to the preference ranking of the juices rather than to the juice identities. In other words, neurons encoding the Offer Value A in the first block typically came to encode the Offer Value C in the second block, while neurons encoding the Offer Value B in the first block came to encode the Offer Value A in the second block (**Fig. 6e**, $P < 0.02$, Fisher's exact test). Similar results were observed for Chosen Juice cells (**Fig. 6f**, $P < 0.01$, Fisher's exact test). Thus neuron–juice associations in our experiments were not driven by the sensory features of the juices.

DISCUSSION

The neuronal mechanisms through which subjective values are compared during economic decisions are not well understood, but there is general consensus that decisions between goods take place in an abstract representation[27–30]. Importantly, the three groups of neurons identified in OFC are computationally sufficient to generate choices[19], suggesting that good-based decisions emerge from a neural circuit within this region[18]. In this framework, the present study provides an assessment of how the decision circuit adapts to the goods available at any given time. This adaptation presents two complementary aspects. On the one hand, the association between individual cells and particular goods is highly flexible, as neurons remap to encode the identities and subjective values of the goods currently available. On the other hand, the overall organization of the circuit, including the variable encoded by each neuron and the composition of neuronal pools, persists across behavioral contexts. Taken together, circuit persistence and neuronal remapping make it possible for the same circuit to generate decisions involving a potentially infinite variety of goods.

In principle, the fact that individual cells encode the same variable in different contexts might be dictated by intrinsic properties of the neuronal populations. For example, Offer Value cells (capturing the decision input) and Chosen Juice cells (capturing the decision output) might be preferentially found in different cortical layers and/or might present different patterns of local and long-distance connectivity. Along similar lines, different groups of cells might preferentially correspond to pyramidal cells versus interneurons. A neural network in

which economic decisions emerge from a balance of recurrent excitation and pooled inhibition suggests that Chosen Juice cells are primarily excitatory while Chosen Value cells are primarily inhibitory[19,31]. Future research should examine in greater detail the neuronal circuitry within the OFC, including the excitatory or inhibitory nature of different cell groups and their location in cortical layers.

We use the term 'remapping' in analogy to hippocampal place cells, whose place fields relocate when animals move to a new arena[32]. Like the mapping between hippocampal neurons and the Euclidean space, the mapping between orbitofrontal neurons and the space of goods may be viewed as a collective property of the neural assembly[33]. The rules that govern this mapping—i.e., how a given pool of cells becomes assigned to one particular good—remain unclear. In this study, assignments appeared driven by the preference ranking, but we cannot rule out alternative schemes. Notably, the preferred juice was always offered in a smaller range in our experiments. Previous work indicates that unequal value ranges impose a cascade of neuronal and synaptic adaptations, with Offer Value cells adapting to the range of offered values[34,35] and downstream populations accounting for this adaptation[36]. In this light, the remapping patterns observed here might have been the most efficient for the system in our conditions. Future work will investigate the mechanisms of remapping, including the rules governing more complex situations where choices are made between multiple goods.

The experiment in which neurons remapped according to the preference ranking as opposed to the juice identity (A:B, C:A design) makes it clear that neuron–juice associations were not driven by the sensory features of the goods. This finding resonates with the results of an earlier study, in which monkeys were delivered one of three foods and different foods were paired in different trial blocks[37]. An important difference between the present finding and earlier results is that the previous study did not distinguish between different groups of neurons. Consequently, the results afforded at least two interpretations. One possibility is that neurons encoded the value of a particular food in the first block and then remapped to encode the value of a different food in the second block. Another possibility is that neurons encoded the outcome value independently of the food identity in both blocks, while adapting to the expected value of the outcome in each trial block. Recent critiques[35,38] generally favor the latter interpretation, which corresponds to a 'quantitative' adaptation[34]. In contrast, the present results for Offer Value and Chosen Juice cells can only be interpreted in terms of remapping, which corresponds to a 'qualitative' adaptation[34].

The present results also resonate with previous work on reversal learning showing that neurons in the primate OFC are associated with a given outcome, independently of how that outcome is signaled to the animal[7,39]. However, our results add considerable complexity to that picture because they show that OFC neurons are not rigidly associated with any given outcome (i.e., a particular juice). Rather, OFC neurons remap onto one of the possible outcomes available in any behavioral context. While previous reports could have justified the opposite prediction, the core concept shared by present and past observations is that neuronal representations in the OFC are highly flexible.

The remapping of neuron–juice associations described here is a form of neuronal plasticity. Some insight into the time course of this process come from a previous study, in which we examined decisions between three juices (A, B, C in decreasing order of preference) offered pairwise[40]. In that study, trials with the three juice pairs (A:B, B:C, A:C) were randomly interleaved and neuron–juice associations did not change from trial to trial. For example, some neurons encoded the Offer Value B in both A:B trials and B:C trials and were Untuned in A:C trials. Taken together, earlier and present result indicate that neuronal remapping in the OFC takes place over the time scale of minutes, not seconds.

In conclusion, we examined how the neuronal representation of good identities and subjective values reorganizes when the economic goods available for choice change. In accord with the fact that choices in different contexts may involve a potentially infinite variety of goods, neurons in the OFC remapped to represent the goods available at any given time. Most remarkably, the overall organization of the decision circuit and the composition of neuronal pools remained stable across behavioral contexts. Thus the same neural circuit may generate decisions between any two goods.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank H. Schoknecht for help with animal training, A. Raghuraman for help with recording and X. Cai, K. Conen, E. Han, I. Monosov and L. Snyder for comments on the manuscript. This work was supported by the National Institutes of Health (grant numbers R01-DA032758 and R01-MH104494 to C.P.-S.) and by the McDonnell Center for Systems Neuroscience (predoctoral fellowship to J.X.).

AUTHOR CONTRIBUTIONS

J.X. and C.P.-S. designed the study; J.X. collected and analyzed the data; J.X. and C.P.-S. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interest.

1. Rahman, S., J Sahakia, B., N Cardinal, R., Rogers, R. & Robbins, T. Decision making and neuropsychiatry. *Trends Cogn. Sci.* **5**, 271–277 (2001).
2. Volkow, N.D. & Li, T.K. Drug addiction: the neurobiology of behaviour gone awry. *Nat. Rev. Neurosci.* **5**, 963–970 (2004).
3. Strauss, G.P., Waltz, J.A. & Gold, J.M. A review of reward processing and motivational impairment in schizophrenia. *Schizophr. Bull.* **40** (suppl. 2), S107–S116 (2014).
4. Gallagher, M., McMahan, R.W. & Schoenbaum, G. Orbitofrontal cortex and representation of incentive value in associative learning. *J. Neurosci.* **19**, 6610–6614 (1999).
5. Camille, N., Griffiths, C.A., Vo, K., Fellows, L.K. & Kable, J.W. Ventromedial frontal lobe damage disrupts value maximization in humans. *J. Neurosci.* **31**, 7527–7532 (2011).
6. Rudebeck, P.H., Saunders, R.C., Prescott, A.T., Chau, L.S. & Murray, E.A. Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nat. Neurosci.* **16**, 1140–1145 (2013).
7. Thorpe, S.J., Rolls, E.T. & Maddison, S. The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* **49**, 93–115 (1983).
8. Padoa-Schioppa, C. & Assad, J.A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
9. Raghuraman, A.P. & Padoa-Schioppa, C. Integration of multiple determinants in the neuronal computation of economic values. *J. Neurosci.* **34**, 11583–11603 (2014).
10. Howard, J.D., Gottfried, J.A., Tobler, P.N. & Kahnt, T. Identity-specific coding of future rewards in the human orbitofrontal cortex. *Proc. Natl. Acad. Sci. USA* **112**, 5195–5200 (2015).
11. Zhou, J., Jia, C., Feng, Q., Bao, J. & Luo, M. Prospective coding of dorsal raphe reward signals by the orbitofrontal cortex. *J. Neurosci.* **35**, 2717–2730 (2015).
12. Klein-Flügge, M.C., Barron, H.C., Brodersen, K.H., Dolan, R.J. & Behrens, T.E. Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. *J. Neurosci.* **33**, 3202–3211 (2013).
13. Lara, A.H., Kennerley, S.W. & Wallis, J.D. Encoding of gustatory working memory by orbitofrontal neurons. *J. Neurosci.* **29**, 765–774 (2009).
14. McDannald, M.A. *et al.* Orbitofrontal neurons acquire responses to 'valueless' Pavlovian cues during unblocking. *eLife* **3**, e02653 (2014).
15. Critchley, H.D. & Rolls, E.T. Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. *J. Neurophysiol.* **75**, 1673–1686 (1996).

16. Gottfried, J.A., O'Doherty, J. & Dolan, R.J. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* **301**, 1104–1107 (2003).
17. Sul, J.H., Kim, H., Huh, N., Lee, D. & Jung, M.W. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron* **66**, 449–460 (2010).
18. Padoa-Schioppa, C. Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* **34**, 333–359 (2011).
19. Rustichini, A. & Padoa-Schioppa, C. A neuro-computational model of economic decisions. *J. Neurophysiol.* **114**, 1382–1398 (2015).
20. Padoa-Schioppa, C. Neuronal origins of choice variability in economic decisions. *Neuron* **80**, 1322–1336 (2013).
21. Agresti, A. *An Introduction to Categorical Data Analysis* (Wiley, 2007).
22. Gottfried, J.A. & Zelano, C. The value of identity: olfactory notes on orbitofrontal cortex function. *Ann. NY Acad. Sci.* **1239**, 138–148 (2011).
23. Pritchard, T.C. *et al.* Satiety-responsive neurons in the medial orbitofrontal cortex of the macaque. *Behav. Neurosci.* **122**, 174–182 (2008).
24. Rolls, E.T. The orbitofrontal cortex and reward. *Cereb. Cortex* **10**, 284–294 (2000).
25. Small, D.M. Flavor is in the brain. *Physiol. Behav.* **107**, 540–552 (2012).
26. Verhagen, J.V. & Engelen, L. The neurocognitive bases of human multimodal food perception: sensory integration. *Neurosci. Biobehav. Rev.* **30**, 613–650 (2006).
27. Cisek, P. Making decisions through a distributed consensus. *Curr. Opin. Neurobiol.* **22**, 927–936 (2012).
28. Glimcher, P.W. *Foundations of Neuroeconomic Analysis* (Oxford University Press, 2011).
29. Rangel, A. & Hare, T. Neural computations associated with goal-directed choice. *Curr. Opin. Neurobiol.* **20**, 262–270 (2010).
30. Rushworth, M.F., Kolling, N., Sallet, J. & Mars, R.B. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Curr. Opin. Neurobiol.* **22**, 946–955 (2012).
31. Wang, X.J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955–968 (2002).
32. Muller, R.U. & Kubie, J.L. The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *J. Neurosci.* **7**, 1951–1968 (1987).
33. Wills, T.J., Lever, C., Cacucci, F., Burgess, N. & O'Keefe, J. Attractor dynamics in the hippocampal representation of the local environment. *Science* **308**, 873–876 (2005).
34. Padoa-Schioppa, C. Range-adapting representation of economic value in the orbitofrontal cortex. *J. Neurosci.* **29**, 14004–14014 (2009).
35. Kobayashi, S., Pinto de Carvalho, O. & Schultz, W. Adaptation of reward sensitivity in orbitofrontal neurons. *J. Neurosci.* **30**, 534–544 (2010).
36. Padoa-Schioppa, C. & Rustichini, A. Rational attention and adaptive coding: a puzzle and a solution. *Am. Econ. Rev.* **104**, 507–513 (2014).
37. Tremblay, L. & Schultz, W. Relative reward preference in primate orbitofrontal cortex. *Nature* **398**, 704–708 (1999).
38. Schultz, W. Neuronal reward and decision signals: from theories to data. *Physiol. Rev.* **95**, 853–951 (2015).
39. Morrison, S.E., Saez, A., Lau, B. & Salzman, C.D. Different time courses for learning-related changes in amygdala and orbitofrontal cortex. *Neuron* **71**, 1127–1140 (2011).
40. Padoa-Schioppa, C. & Assad, J.A. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat. Neurosci.* **11**, 95–102 (2008).

ONLINE METHODS

Experimental design and neuronal recordings. Two adult male rhesus monkeys (*Macaca mulatta*; monkey V, 8.3 kg; monkey C, 11.7 kg) participated in the study. All experimental procedures conformed to the NIH *Guide for the Care and Use of Laboratory Animals* and were approved by the Animal Studies Committee at Washington University in St Louis. Before training, a head restraining device and an oval recording chamber (main axes 50 × 30 mm) were implanted on the skull under general anesthesia. The chamber was centered on stereotaxic coordinates (A30, L0) and allowed access to bilateral OFC with penetrations on a nearly coronal plane. During the experiments, monkeys sat in an electrically insulated enclosure with their heads restrained and performed a juice choice task. Visual stimuli were presented on a computer monitor placed 57 cm in front of the monkey. Eye positions were monitored with an infrared video camera (Eyelink, SR research). The behavioral task was controlled through custom-written software (http://www.monkeylogic.net) based on Matlab (MathWorks).

Structural MRI scans were obtained before and after the surgery. Coronal images parallel to the paths of electrodes were reconstructed from the two MRIs to provide a reference for recording locations. Recordings were performed in the orbital gyrus, in a region roughly corresponding to area 13m. Neuronal data were collected from the left hemisphere of monkey V (A 31:35, L −7: −11) and both hemispheres of monkey C (A 33:38, L −6: −10, left hemisphere; A 33:38, L 6:11, right hemisphere). Most of the recordings were performed using 4 or 6 individual tungsten electrodes (125 µm diameter, Frederick Haer) simultaneously. The electrodes, placed ≥1 mm apart from each other, were advanced using a custom-built motorized micro-drive (step size: 2.5 µm). A subset of data was collected using 8- or 16-channel linear arrays (U-probe; Plexon). U-probes had a diameter of 185 µm and contacts were 100 µm from each other. U-probes were advanced with the same micro-drive system used for single electrodes. Electrical signals were amplified (gain: 10,000), band-pass filtered (low-pass cut-off: 300 Hz, high-pass cut-off: 6 kHz; Lynx 8, Neuralynx) and recorded at 25 kHz resolution (Power 1401, Cambridge Electronic Design). Action potentials were detected online and saved to disk for spike sorting, which was performed using standard software (Spike 2, Cambridge Electronic Design). Only neurons that remained stable and well-isolated during two consecutive blocks were included in the analysis.

Figure 1 illustrates the experimental design. In each trial, the animal chose between two juices offered in variable amounts. The trial began with the monkey gazing a fixation point in the center of the monitor (center fixation window: 2°). After 1.5 s, two sets of squares representing the two offers appeared on the two sides of the fixation point. For each offer, the color indicated the juice type and the number of squares indicated juice quantity, with each square representing a juice quantum (65 µl for monkey V, 70 µl for monkey C). The animal maintained center fixation for a randomly variable delay (1–2 s, uniform distribution) followed by a go signal, which was indicated by the extinction of the center fixation point and the appearance of two saccade targets. The monkey had 1.5 s to indicate its choice with a saccade and was required to maintain target fixation (peripheral fixation window: 3°) for 0.75 s before juice delivery.

Each session included two blocks of trials, and two different juices were used in each block. A, B, C and D indicate different juice types, and "X:Y" indicates choices between juices X and Y, with juice X preferred to juice Y. An "offer type" was defined by two offers (for example, [1A:3B]), while a "trial type" was defined by an offer type and one choice (for example, [1A:3B, B]). In experimental sessions, we used four different juices (A:B, C:D design; 499 and 219 cells recorded from monkey V and C, respectively). In control sessions, we used the same juice pair in both blocks (A:B, A:B design; 225 and 162 cells recorded from monkey V and C, respectively). Experimental and control sessions were interleaved in the experiments, but the order was not determined using a formal randomization. Blinding was not used. We also conducted a follow-up experiment in which the preferred juice of the first block was used as a non-preferred juice in the second block (A:B, C:A design; 329 cells recorded from monkey V). Trial blocks lasted 200–250 trials and typically included 9–11 offer types (**Fig. 2a**). In each block, offer types were pseudorandomly interleaved and the left/right configurations of the offers were counterbalanced. In all sessions, the two blocks were separated by a 10–15 min break. At the beginning of each block, we let the monkeys practice for 20 ± 2 trials to adapt to the current juice pairs. Practice trials were excluded from the analysis. The number of cells collected for each condition and the number of trials run for each cell (sample size) were not predetermined using a statistical method, but were comparable to those of previous studies[8,40].

Behavioral analysis and neuronal classification. Behavioral data were analyzed separately for each trial block, as previously described[8,40]. The 'choice pattern' was defined as the percent of trials in which the animal chose the non-preferred juice as a function of the log quantity ratio of the two juices. In the analysis, we fitted the choice pattern with a normal cumulative distribution function (probit). The flex of the sigmoid, corresponding to the indifference point, provided a measure for the relative value between the two juices.

The analysis performed for each cell in each trial block followed the procedures used in previous studies[8,20]. Neuronal activity was examined in the following time windows: pre-offer (0.5 s before offer), post-offer (0.5 s after the offer), late delay (0.5–1.0 s after the offer), pre-go (0.5 s before the go), reaction time (from the go to the saccade), pre-juice (0.5 s before the juice), post-juice (0.5 s after juice) and post-juice2 (0.5–1.0 s after juice). Firing rates were obtained by averaging spike counts over each time window and across trials for each trial type. A 'neuronal response' was defined as the firing rate of one cell, in one block, in one time window, as a function of the trial type. The two responses recorded for the same cell, in the same time window, in the two blocks defined a 'response pair'. For each neuron and each time window, we performed a one-way ANOVA (factor: trial type) separately in each block. We imposed a $P < 0.01$ threshold and admitted to subsequent analyses only response pairs that met this criterion in at least one block. Neurons that passed this ANOVA criterion in at least one time window were identified as 'task-related'. For a control, we repeated the analysis of **Figure 3**, imposing a more conservative ANOVA threshold ($P < 0.001$), and obtained very similar results.

This study builds on previous work showing that under similar behavioral conditions neurons in the OFC encode three variables, namely Offer Value, Chosen Value and Chosen Juice[8]. Thus as a preliminary step, we repeated on the current data set the analyses conducted previously. Specifically, we pooled all the available neurons (protocols A:B, A:B; A:B, C:D and A:B, C:A). Data from each trial block were considered separately, so that the entire data set included 2,868 cells. Since data from each trial block were analyzed separately, each actual neuron contributed two 'cells' to this analysis. Each cell was analyzed in eight time windows and we imposed the ANOVA criterion. We then computed the same 19 variables examined previously and applied the same two procedures for variable selection[8]. Confirming earlier observations, both procedures selected variables Offer Value A, Offer Value B, Chosen Value and Chosen Juice. We also performed a *post hoc* analysis pitting each selected variable against other, non-selected variables. We found that the marginal explanatory power of each selected variable was statistically higher than that of any other, competing variable (all $P < 0.02$; binomial test). These results closely replicated previous findings[8,40,41].

The neuronal classification was performed separately for each trial block. Here we describe the procedures referring to the A:B block. Previous work (and the preliminary analysis described above) showed that the vast majority of neuronal responses encoded one of four variables: Offer Value A, Offer Value B, Chosen Value and Chosen Juice[8]. For each variable, the encoding was linear and either positive (increasing firing rate for increasing values) or negative (increasing firing rate for decreasing values)[8,20]. Importantly, neurons typically encoded the same variable with the same sign across time windows[20]. To classify any given neuron, we proceeded as in previous studies[8,20]. We only considered time windows that passed the ANOVA criterion, and we performed a linear regression of each response on each variable. A variable was said to 'explain' the response if the regression slope differed significantly from zero ($P < 0.05$). The regression also provided an R^2, which was set equal to zero if the variable did not explain the response. For each variable we then computed the total R^2 across time windows, taking the encoding sign into consideration. Specifically, we set R^2 equal to zero if the sign of the encoding was opposite to the sign of the variable under consideration. We identified the signed encoded variable as that which provided the highest total R^2. Thus each neuron in each trial block was assigned to one of the following classes: Offer Value+, Offer Value−, Chosen Value+, Chosen Value− and Chosen Juice. If none of the regression slopes differed significantly from zero, the neuron was classified as Untuned.

The study was designed to examine whether and how the variable encoded by any given cell varied from block to block in experimental sessions (A:B, C:D design). Since the four variables are generally correlated[8] and since neuronal firing rates are highly variable, we expected that our classification procedure might bear some degree of inaccuracy. Recordings in control session (A:B, A:B design) were conducted to obtain a benchmark for the encoding consistency.

doi:10.1038/nn.4300

Comparing neuronal classifications across trial blocks. To compare the classifications across blocks at the population level, we constructed contingency tables separately for data collected in control sessions and experimental sessions (**Fig. 3a**,**d**). The rows and columns of these tables represent the classifications of neurons in the two blocks respectively, while the numbers represent cell counts. Cells located on the main diagonal of these tables encoded the same variables in both trial blocks. Our purpose was to assess, for each location in the contingency table, whether the cell count deviated significantly from chance level, assuming independent classifications in the two trial blocks. To this end, we used two statistical tests.

The first test was based on odds ratios. We indicate with $X_{i,j}$ the number of cells encoding variable i in the first block and variable j in the second block. For each location (i, j) in the table, we computed a 2×2 matrix with elements

$$a_{11} = X_{i,j},\ a_{12} = \sum_{n \neq j} X_{i,n},\ a_{21} = \sum_{m \neq i} X_{m,j} \text{ and } a_{22} = \sum_{n \neq j, m \neq i} X_{m,n}$$

The odds ratio for location (i, j) was defined as $OR_{i,j} = (a_{11}/a_{12})/(a_{21}/a_{22})$. If the likelihood of a neuron encoding variable i in the first block is independent of the likelihood of it encoding variable j in the second block, the expected value of $OR_{i,j}$ equals 1. In other words, the chance level for odds ratio is 1. Conversely, $OR_{i,j} > 1$ ($OR_{i,j} < 1$) indicated that the cell count in location (i, j) was above (below) chance level. To assess whether departures from chance level were statistically significant, we used Fisher's exact test (two tails)[21]. Exact P values are shown in **Supplementary Figure 6**.

The second test was based on a bootstrap analysis, in which we estimated the distribution of cell count for each element of the contingency tables. The classifications of neurons in the two blocks constituted two separate data sets. We sampled with replacement from the two data sets independently and randomly paired the two samples. We thus constructed a contingency table with the same number of cells recorded in the experiments. We repeated the procedure 10,000 times. For each location in the contingency table, we thus obtained a bootstrap distribution of cell counts, to which we compared the cell counts obtained experimentally. Numbers indicated in **Figure 3c**,**f** represent the fraction of bootstrap repetitions for which the cell count was equal or larger than that obtained experimentally (i.e., the P values).

In some cases, the representations provided by a population of cells in different contexts can be compared using a linear decoder analysis[42]. However, this approach was not possible in our study because offer types and relative values varied from session to session. Hence, neurons recorded in different sessions could not be pooled. Conversely, the results obtained in control sessions provided us with a valuable benchmark. Thus in a series of analyses we directly compared the classification results obtained for experimental and control cells (see Results). For each condition, we constructed a reduced contingency table (**Fig. 5a**,**c**). To compare the two tables, we estimated the common odds ratios and the expected cell counts, and we ran a Breslow-Day test for each element of the table[21]. All chi-square values are shown in **Supplementary Figure 7**.

Comparing neuron–juice associations across trial blocks. The analyses illustrated in **Figures 3–5** indicated that neurons generally encoded the same variable (Offer Value, Chosen Value, Chosen Juice) with the same slope sign (+ or −) in both trial blocks. In subsequent analyses we assumed that this was the case and reclassified neurons accordingly. We imposed that each cell encode the same variable (Offer Value, Chosen Juice, Chosen Value) with the same sign (+ or −) in both trial blocks. Critically, Offer Value cells and Chosen Juice cells could be associated with either juice in each block. This resulted in 14 possible combinations (for example, [Offer Value A+, Offer Value C+], [Offer Value A+, Offer Value D+], etc.). Each cell was assigned to a particular combination based on the total R^2 summed across time windows and trial blocks. Neurons that could not be explained by the same variable in both trial blocks were excluded from this analysis.

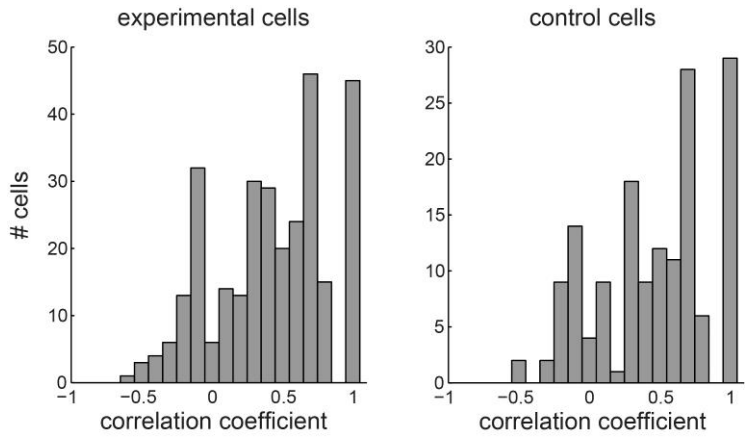
Focusing on Offer Value cells and Chosen Juice cells, we sought to assess whether the neuron–juice association recorded in the first trial block was generally correlated with the neuron–juice association recorded in the second trial block. Specifically, we examined whether neuron–juice associations were typically dictated by the preference ranking of the juices. We thus constructed 2×2 contingency tables, separating neurons according to whether they were associated with the preferred juice or to the non-preferred juice in each trial block (**Fig. 6**). Each contingency table was examined using Fisher's exact test, which established whether the two neuron–juice associations were significantly correlated across the population.

Statistical notes. The primary results of this study were based on analyses of categorical data. Fisher's exact test and the bootstrap procedure had minimal assumptions and the Breslow-Day test had a cell count requirement[21] that was met by our reduced contingency tables. Importantly, these analyses followed neuronal classification procedures based on least-squares linear regressions, which assume normality and equal variance. As previously discussed[8], data in our experiments approximately satisfied normality because individual data points in the regressions were averages of 10–20 trials. In contrast, variances were generally unequal, with high-variance data points closer to the behavioral indifference point (where choices were split). Because correcting for unequal variance would effectively reduce the weight of these data points[43], which were in many respects the most informative, we deemed it preferable to use uncorrected data.

A **Supplementary Methods Checklist** is available.

Code availability. All analyses were performed using custom code written in Matlab (MathWorks) and available upon request.

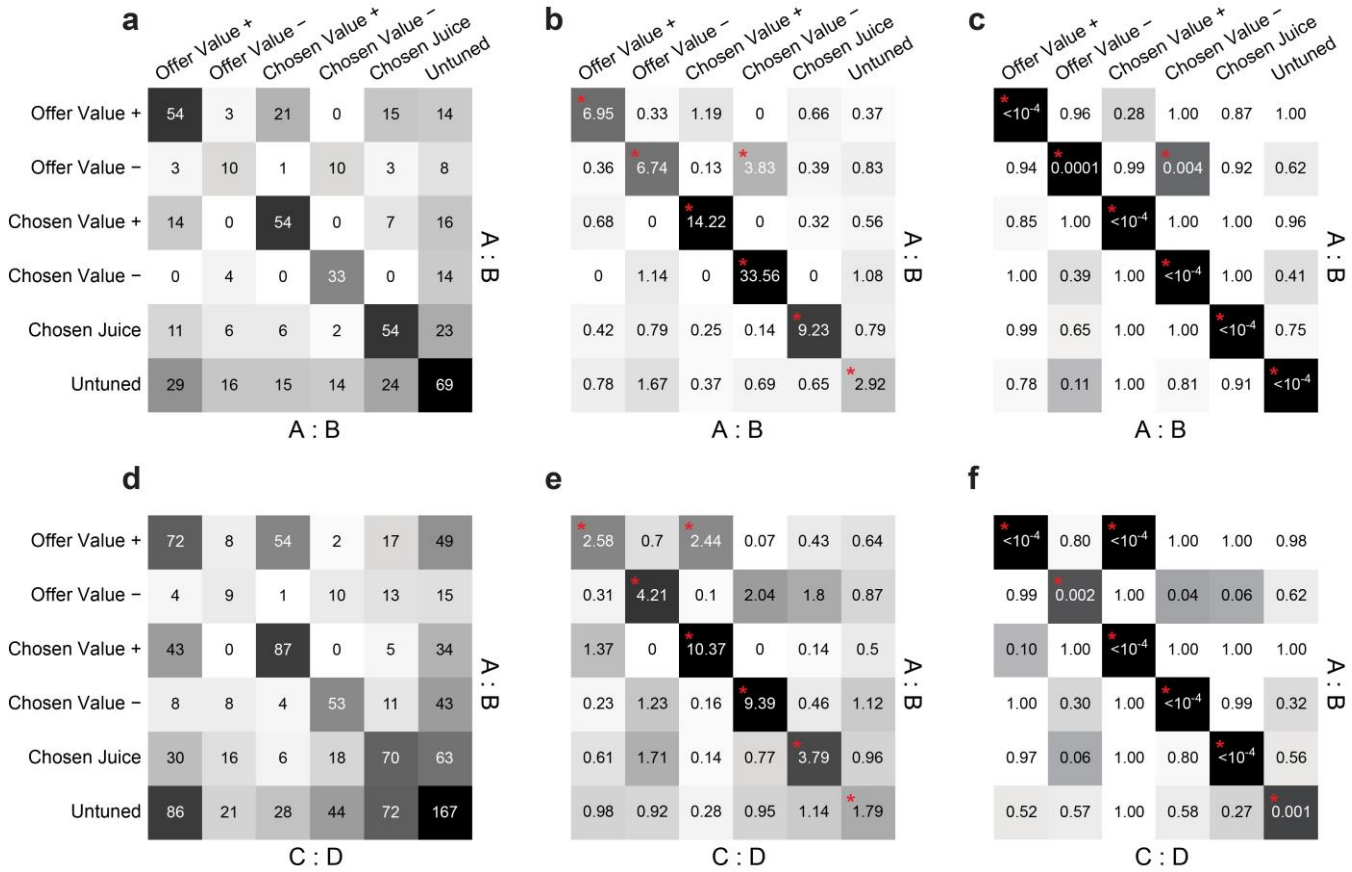
41. Cai, X. & Padoa-Schioppa, C. Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation. *Neuron* **81**, 1140–1151 (2014).
42. Saez, A., Rigotti, M., Ostojic, S., Fusi, S. & Salzman, C.D. Abstract context representations in primate amygdala and prefrontal cortex. *Neuron* **87**, 869–881 (2015).
43. Neter, J., Wasserman, W. & Kutner, M.H. *Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs* (Irwin, Homewood, Illinois, USA, 1990).



Supplementary Figure 1

Correlation between task-related activity across trial blocks.

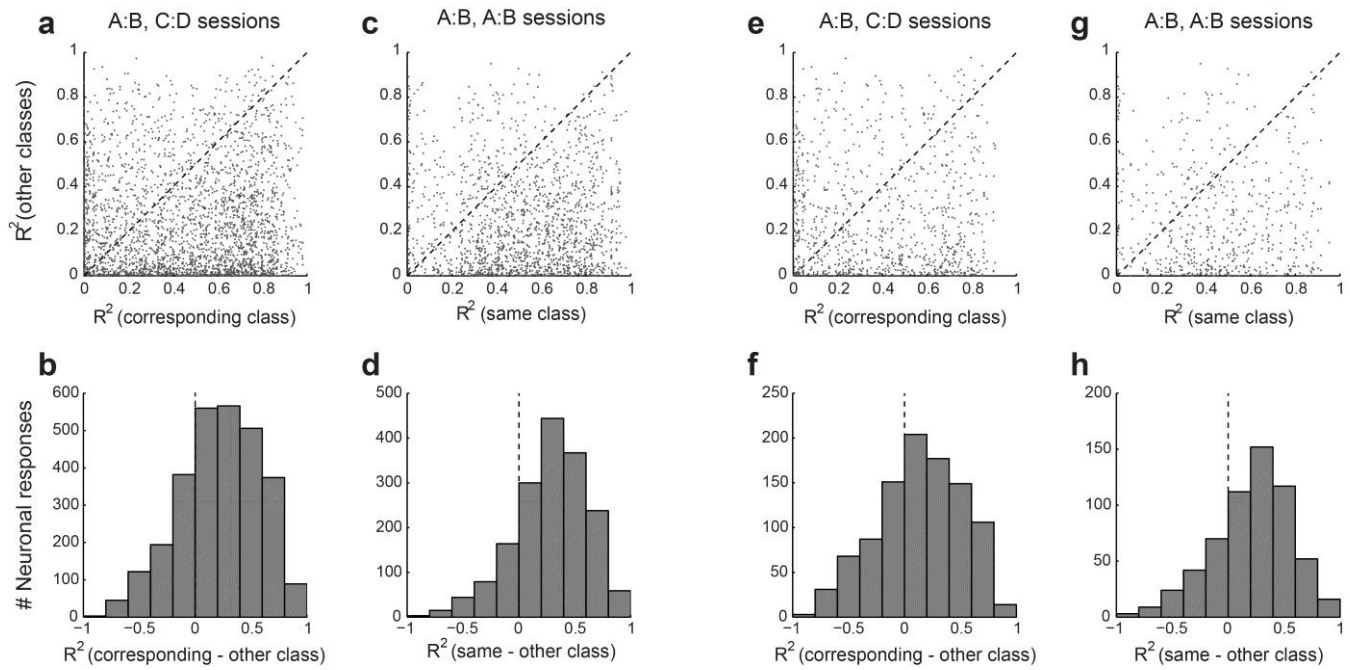
In this analysis, we examined whether individual neurons typically met the ANOVA criterion in the same time windows in both trial blocks. For each neuron in each block, we obtained a vector of binary variables indicating whether the cell passed the criterion in different time windows. We then calculated the correlation coefficient between the two vectors obtained for the two trial blocks. Finally, we examined the distribution of correlation coefficients across the population, separately for experimental and control cells. Correlation coefficients were typically positive, indicating that neurons were generally task-related in the same time windows. This result held true both in the experimental condition (median = 0.45) and in the control condition (median = 0.49). Moreover, the distributions obtained in the two conditions did not differ significantly from one another ($p = 0.83$, Kolmogorov-Smirnov test).



Supplementary Figure 2

Analysis based on neuronal responses.

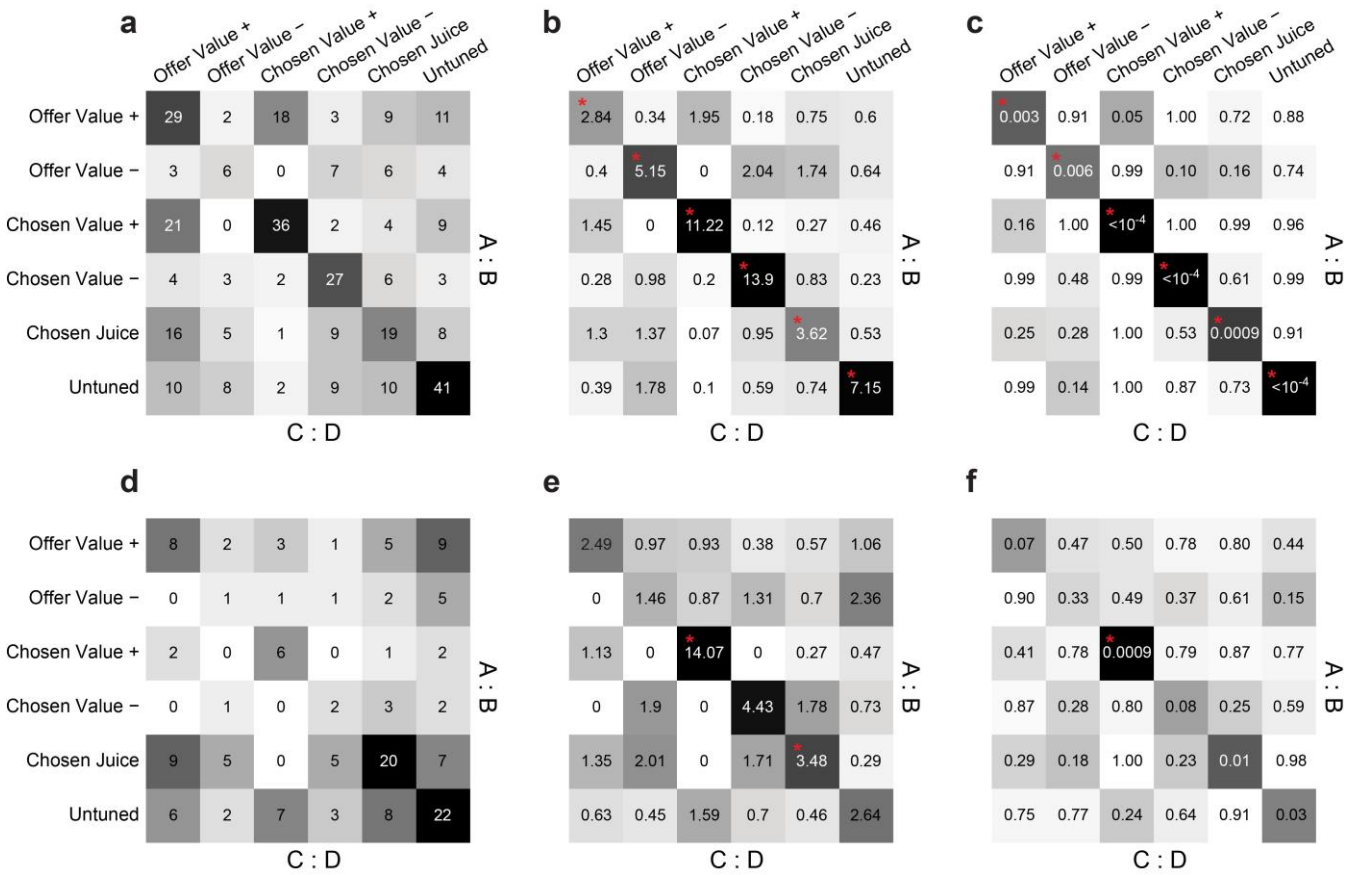
The analyses presented in **Fig. 3** were based on a classification involving multiple time windows (see Online Methods). For a control, we repeated the analysis comparing the classifications obtained for neuronal responses (as opposed to neurons) across trial blocks. As a reminder, a neuronal response was defined as the activity of one cell, in one time window, in one trial block, as a function of the trial type. In this analysis, we compared the two responses obtained from the same cell, in the same time window, for the two trial blocks. (Thus each cell might contribute more than one response to the figure.) (**a,b,c**) Control sessions (N = 553 responses). (**d,e,f**) Experimental sessions (N = 1,171 responses). Data are presented in the same format used for **Fig.** 3. Panels illustrate the response counts (a, d), the analysis of odds ratios (b, e) and the results of the bootstrap analysis (c, f). All statistical criteria are as in **Fig.** 3. For odds ratios, exact p values are shown in **Supplementary Figs. 6e** (relative to panel b) and **6f** (relative to panel e). The results obtained in this analysis confirmed those obtained for neurons (**Fig. 3**). In other words, response counts were significantly above chance level for all positions on the main diagonal, and this results held true both in the control condition and in the experimental condition.



Supplementary Figure 3

Analysis of response similarity across trial blocks based on the R^2.

We conducted a control analysis to examine whether the R^2 obtained for one response in one block predicted the R^2 obtained for the corresponding response in the other block. In essence, this approach avoided the classification procedure. The analysis included all pairs of responses that met the ANOVA criterion in both blocks. We first examined experimental cells. For each block, we obtained four R^2 from the linear regressions on each of the 4 variables. We identified the variable providing the highest R^2 in one block, and we considered the "corresponding" variable in the other block. For variables *offer value A* and *offer value B*, the corresponding variables were *offer value C* and *offer value D*, respectively. Independently of the value measured for the R^2 in the A:B block, we predicted that the R^2 measured for the corresponding variable in the C:D block would be typically higher than that measured for the other, non-corresponding variables (and vice versa). This is indeed what we observed. (**a**) Neuronal population recorded in experimental sessions (A:B, C:D). The x-axis and y-axis in the scatter plot represent the R^2 obtained in the second block for the corresponding variable and other variables, respectively. While data are quite scattered, the population as a whole lies below the identity line, indicating that the variable encoded in one block is predictive of the variable encoded in the other block. Each response provided 3 data points to this plot, one for each of the "other" variables. Furthermore, we repeated the analysis using for reference the second block, and we pooled the results obtained with the two procedures. (**b**) For each data point in panel a we computed the difference in R^2 obtained for the corresponding and other variables. The distribution obtained for the population was significantly shifted towards positive values (median = 0.24, $p<10^{-10}$, Wilcoxon signed rank test). (**c,d**) Same analysis as in panels ab performed for the population recorded in control sessions (A:B, A:B). In panel d, the distribution of ΔR^2 obtained for the population was significantly shifted towards positive values (median = 0.32, $p<10^{-10}$, Wilcoxon signed rank test). (**e,f,g,h**) Analysis restricted to *offer value* responses. The four panels illustrate the results obtained when the analysis was restricted to neuronal responses classified as *offer value* in the reference block. The results were very similar to those obtained for the entire population. The distribution of ΔR^2 was significantly shifted towards positive values for both experimental (median = 0.15, $p<10^{-10}$, Wilcoxon signed rank test) and control sessions (median = 0.24, $p<10^{-10}$, Wilcoxon signed rank test).



Supplementary Figure 4

Analysis of individual animals (experimental cells).

(**a,b,c**) Monkey V (N = 353 task-related cells). (**d,e,f**) Monkey C (N = 151 task-related cells). Data are presented in the same format used for **Fig. 3**. Panels illustrate the cell counts (a, d), the analysis of odds ratios (b, e) and the results of the bootstrap analysis (c, f). All statistical criteria are as in **Fig. 3**. For odds ratios, exact p values are shown in **Supplementary Figs. 6g** (relative to panel b) and **6h** (relative to panel e). For monkey V, odds ratios were significantly above chance for all locations on the main diagonal. For monkey C, odds ratios on the main diagonal were all above chance, but the effects reached significance level only for *chosen value* + cells and *chosen juice* cells (lower statistical power). Notably, cells classified as *offer value* or *chosen juice* in one block were not more likely to be classified as *untuned* in the other block, for either monkey.

a

	Untuned	Other
Control	9	65
Experimental	29	107

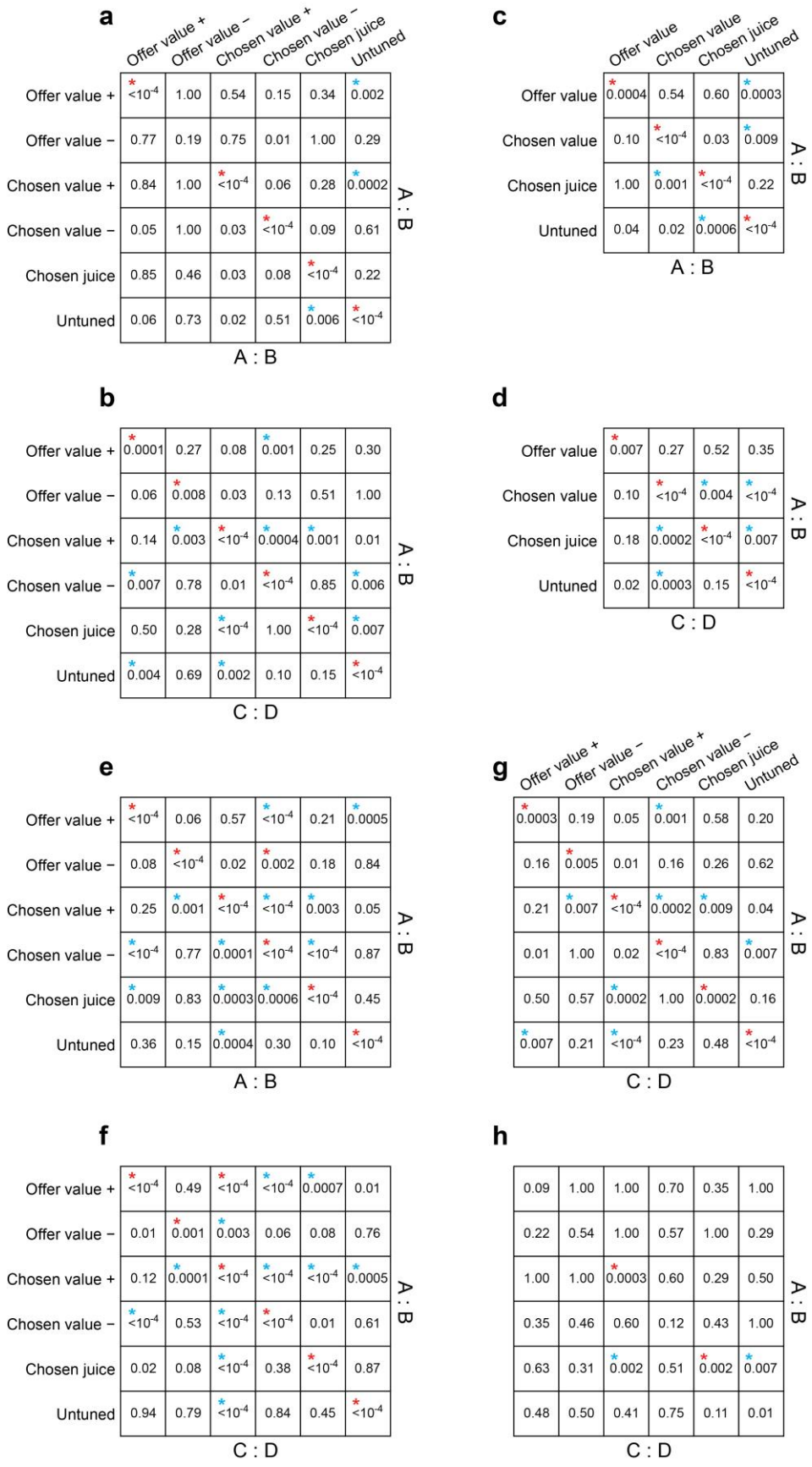
b

	Chosen Juice	Other
Control	4	64
Experimental	18	110

Supplementary Figure 5

Analysis of conditional odds ratios.

In the analysis illustrated in **Fig. 5e** and based on the Breslow-Day statistics, two association strengths, namely [*offer value*, *untuned*] and [*untuned*, *chosen juice*], approached significance level. We thus conducted an additional analysis on these two associations. Specifically, we computed the conditional odds ratios for fixed levels of one dimension (Agresti, *An Introduction to Categorical Data Analysis*, 2007). (**a**) Association [*offer value*, *untuned*]. Numbers in the table indicate cell counts. Only neurons classified as *offer value* in the first trial block enter this table. The two columns indicate the results of the classification obtained in the second block (*untuned* or any of the three other classes). For this 2x2 table, odds ratio = 0.51 (p = 0.13, Fisher's exact test). In other words, the strength of the association [*offer value*, *untuned*] did not differ between conditions. Conditioning cell counts on the classification obtained in the second block provided similar results (odds ratio = 0.48; p = 0.09, Fisher's exact test). (**b**) Association [*untuned*, *chosen juice*]. Same format as in panel a. Only neurons classified as *untuned* in the first trial block enter this table. For this table, odds ratio = 0.38 (p = 0.10, Fisher's exact test). In other words, the strength of the association [*untuned*, *chosen juice*] did not differ between conditions. Conditioning cell counts on the classification obtained in the second block provided similar results (odds ratio = 0.37; p = 0.09, Fisher's exact test). In conclusion, the association strengths measured in the experimental condition were statistically indistinguishable from those measured in the control condition.

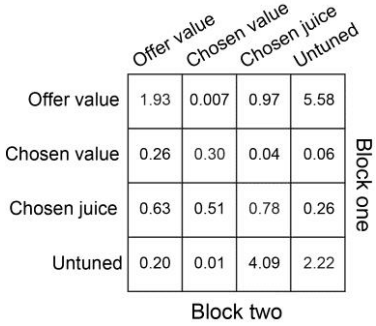


Nature Neuroscience: doi:10.1038/nn.4300

Supplementary Figure 6

Exact *P* values for odds ratios.

The figure shows the p values obtained using Fisher's exact test (two tails) on odds ratios throughout the paper. The correspondence between the panels of this figures and those of other figures is as follows: (**a**) **Fig. 3b**. (**b**) **Fig. 3e**. (**c**) **Fig. 5b**. (**d**) **Fig. 5d**. (**e**) **Supplementary Fig. 2b**. (**f**) **Supplementary Fig. 2e**. (**g**) **Supplementary Fig. 4b**. (**h**) **Supplementary Fig. 4e**. In each panel, red/cyan asterisks indicate that the cell counts were significantly above/below chance level ($p<0.01$).



Supplementary Figure 7

Chi-square values.

The table indicates the chi2 values obtained for each of the Breslow-Day tests performed for **Fig. 5e**.