HHS Public Access

Author manuscript

Nat Neurosci. Author manuscript; available in PMC 2016 November 09.

Published in final edited form as:

Nat Neurosci. 2016 June ; 19(6): 855–861. doi:10.1038/nn.4300.

Neuronal Remapping and Circuit Persistence in Economic Decisions

Jue Xie[a] and **Camillo Padoa-Schioppa**[a,b,c]

[a]Department of Neuroscience, Washington University in St Louis, St Louis, MO 63110

[b]Department Economics, Washington University in St Louis, St Louis, MO 63110

[c]Department Biomedical Engineering, Washington University in St Louis, St Louis, MO 63110

Abstract

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 functional role 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.

Introduction

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.

Correspondence: Camillo Padoa-Schioppa, Ph.D., Department of Neuroscience, Washington University in St Louis, *Campus Box 8108, Tel: 314-747-2253*, camillo@wustl.edu.

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.

To shed light onto how the decision circuit re-organizes 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 (opposite) decision in one context also supported the same (opposite) decision 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 (see 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 neurons (**Fig. 2d**) and negative for the other neuron (**Fig. 2e**). Note that for *offer value* cells (**Fig. 2ab**) and *chosen juice* cells (**Fig. 2c**), encoding the same variable in both trial blocks 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 (e.g., [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 *untuned* if it was not explained by any variable (see 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 symbols "+" 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. However, some cells fell outside the main diagonal, indicating that our classification had some degree of inaccuracy (see 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 and quite remarkable. Consider neurons encoding the *offer value* in the first block. In principle, when the animal progresses to the second block, the entire circuit could re-organize itself in such a way 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 (see bottom row and rightmost column in **Fig. 3e,f**).

Several control analyses confirmed these results (**Supplementary Figs. 1, 2, 3** and **4**).

Contextual remapping is essentially complete

In the analyses illustrated in **Fig. 3**, the null hypothesis represented the scenario under 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. A separate set of analyses was designed to contrast the results obtained in the experimental condition (A:B, C:D design) with 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 and we obtained 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, chi2 = 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 the Breslow-Day statistic for homogeneous associations[21]. **Fig. 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 conditions. This result implies that the classification patterns measured in the experimental condition were as consistent as one could expect, given the fact 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 functional role 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. As illustrated (**Fig. 6a**), 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. 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 (non-preferred) juice in the first block to the preferred (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 rather 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 a first 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. Of note, 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]. Similar to 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" adpatation[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 significant complexity to that picture, because they show that OFC neurons are not rigidly associated with any given outcome (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 re-organizes when 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.

Online Methods

Experimental design and neuronal recordings

Two adult male rhesus monkeys (*Macaca mulatta*; V, 8.3 kg; C, 11.7 kg) participated in the study. 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 head 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 cutoff: 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 (65/70 μl for monkey V/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 (e.g., [1A:3B]), while a "trial type" was defined by an offer type and one choice (e.g., [1A:3B, B]). In experimental sessions, we used four different juices (A:B, C:D design; 499/219 cells recorded from monkey V/C). In control sessions, we used the same juice pair in both blocks (A:B, A:B design; 225/162 cells recorded from monkey V/C). 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 (e.g., x-axis of **Fig. 2a**). In each block, offer types were pseudo-randomly interleaved and the left/right configurations of the offers were counterbalanced. In all sessions, the two blocks were separated by a 10-15 minute 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 ran for each cell (sample size) were not pre-determined 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 1-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 **Fig. 3** imposing a more conservative ANOVA threshold ($p<0.001$), and we 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 we 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 pitching 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 replicate 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 into consideration the sign of the encoding. 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.

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: $a_{11} = X_{i,j}$, $a_{12} = \Sigma_{n\neq j} X_{i,n}$, $a_{21} = \Sigma_{m\neq i} X_{m,j}$, $a_{22} = \Sigma_{n\neq j,\, m\neq i} X_{m,n}$. 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 Fig. 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 **Fig. 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 compared directly 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 for each element of the table a Breslow-Day test[21]. All the chi-square values are shown in **Supplementary Fig. 7**.

Comparing neuron-juice associations across trial blocks

The analyses illustrated in **Figs. 3**, **4** and **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 we 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 (e.g., [*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 wanted to examine 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 and code availability

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. However, a control variable-selection analysis in which linear regressions were corrected for unequal variance provided similar results ([8], not shown).

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. All analyses were performed using custom code written in Matlab (MathWorks) and available upon request. A Supplementary Methods Checklist is available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

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 (pre-doctoral fellowship to J.X.).

References

1. Rahman S, Sahakian BJ, Cardinal RN, Rogers R, Robbins T. Decision making and neuropsychiatry. Trends Cogn Sci. 2001; 5:271–277. [PubMed: 11390298]
2. Volkow ND, Li TK. Drug addiction: the neurobiology of behaviour gone awry. Nat Rev Neurosci. 2004; 5:963–970. [PubMed: 15550951]
3. Strauss GP, Waltz JA, Gold JM. A review of reward processing and motivational impairment in schizophrenia. Schizophr Bull. 2014; 40(Suppl 2):S107–116. [PubMed: 24375459]
4. Gallagher M, McMahan RW, Schoenbaum G. Orbitofrontal cortex and representation of incentive value in associative learning. J Neurosci. 1999; 19:6610–6614. [PubMed: 10414988]
5. Camille N, Griffiths CA, Vo K, Fellows LK, Kable JW. Ventromedial frontal lobe damage disrupts value maximization in humans. J Neurosci. 2011; 31:7527–7532. [PubMed: 21593337]
6. Rudebeck PH, Saunders RC, Prescott AT, Chau LS, Murray EA. Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. Nat Neurosci. 2013; 16:1140–1145. [PubMed: 23792944]
7. Thorpe SJ, Rolls ET, Maddison S. The orbitofrontal cortex: neuronal activity in the behaving monkey. Exp Brain Res. 1983; 49:93–115. [PubMed: 6861938]
8. Padoa-Schioppa C, Assad JA. Neurons in orbitofrontal cortex encode economic value. Nature. 2006; 441:223–226. [PubMed: 16633341]
9. Raghuraman AP, Padoa-Schioppa C. Integration of multiple determinants in the neuronal computation of economic values. J Neurosci. 2014; 34:11583–11603. [PubMed: 25164656]
10. Howard JD, Gottfried JA, Tobler PN, Kahnt T. Identity-specific coding of future rewards in the human orbitofrontal cortex. Proc Natl Acad Sci U S A. 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. 2015; 35:2717–2730. [PubMed: 25673861]
12. Klein-Flugge MC, Barron HC, Brodersen KH, Dolan RJ, Behrens TE. Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. J Neurosci. 2013; 33:3202–3211. [PubMed: 23407973]
13. Lara AH, Kennerley SW, Wallis JD. Encoding of gustatory working memory by orbitofrontal neurons. J Neurosci. 2009; 29:765–774. [PubMed: 19158302]
14. McDannald MA, et al. Orbitofrontal neurons acquire responses to 'valueless' Pavlovian cues during unblocking. Elife. 2014; 3:e02653. [PubMed: 25037263]
15. Critchley HD, Rolls ET. Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. J Neurophysiol. 1996; 75:1673–1686. [PubMed: 8727405]
16. Gottfried JA, O'Doherty J, Dolan RJ. Encoding predictive reward value in human amygdala and orbitofrontal cortex. Science. 2003; 301:1104–1107. [PubMed: 12934011]
17. Sul JH, Kim H, Huh N, Lee D, Jung MW. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. Neuron. 2010; 66:449–460. [PubMed: 20471357]
18. Padoa-Schioppa C. Neurobiology of economic choice: a good-based model. Annu Rev Neurosci. 2011; 34:333–359. [PubMed: 21456961]
19. Rustichini A, Padoa-Schioppa C. A neuro-computational model of economic decisions. J Neurophysiol. 2015 in press.
20. Padoa-Schioppa C. Neuronal origins of choice variability in economic decisions. Neuron. 2013; 80:1322–1336. [PubMed: 24314733]
21. Agresti, A. An Introduction to Categorical Data Analysis. Wiley; 2007.

22. Gottfried JA, Zelano C. The value of identity: olfactory notes on orbitofrontal cortex function. Ann N Y Acad Sci. 2011; 1239:138–148. [PubMed: 22145883]
23. Pritchard TC, et al. Satiety-responsive neurons in the medial orbitofrontal cortex of the macaque. Behav Neurosci. 2008; 122:174–182. [PubMed: 18298260]
24. Rolls ET. The orbitofrontal cortex and reward. Cereb Cortex. 2000; 10:284–294. [PubMed: 10731223]
25. Small DM. Flavor is in the brain. Physiol Behav. 2012; 107:540–552. [PubMed: 22542991]
26. Verhagen JV, Engelen L. The neurocognitive bases of human multimodal food perception: sensory integration. Neurosci Biobehav Rev. 2006; 30:613–650. [PubMed: 16457886]
27. Cisek P. Making decisions through a distributed consensus. Curr Opin Neurobiol. 2012
28. Glimcher, PW. Foundations of neuroeconomic analysis. Oxford University Press; Oxford; New York: 2011.
29. Rangel A, Hare T. Neural computations associated with goal-directed choice. Curr Opin Neurobiol. 2010; 20:262–270. [PubMed: 20338744]
30. Rushworth MF, Kolling N, Sallet J, Mars RB. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? Curr Opin Neurobiol. 2012
31. Wang XJ. Probabilistic decision making by slow reverberation in cortical circuits. Neuron. 2002; 36:955–968. [PubMed: 12467598]
32. Muller RU, Kubie JL. The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. J Neurosci. 1987; 7:1951–1968. [PubMed: 3612226]
33. Wills TJ, Lever C, Cacucci F, Burgess N, O'Keefe J. Attractor dynamics in the hippocampal representation of the local environment. Science. 2005; 308:873–876. [PubMed: 15879220]
34. Padoa-Schioppa C. Range-adapting representation of economic value in the orbitofrontal cortex. J Neurosci. 2009; 29:14004–14014. [PubMed: 19890010]
35. Kobayashi S, Pinto de Carvalho O, Schultz W. Adaptation of reward sensitivity in orbitofrontal neurons. J Neurosci. 2010; 30:534–544. [PubMed: 20071516]
36. Padoa-Schioppa C, Rustichini A. Rational attention and adaptive coding: a puzzle and a solution. American Economic Review: Papers and Proceedings. 2014; 104:507–513.
37. Tremblay L, Schultz W. Relative reward preference in primate orbitofrontal cortex. Nature. 1999; 398:704–708. [PubMed: 10227292]
38. Schultz W. Neuronal reward and decision signals: from theories to data. Physiol Rev. 2015; 95:853–951. [PubMed: 26109341]
39. Morrison SE, Saez A, Lau B, Salzman CD. Different time courses for learning-related changes in amygdala and orbitofrontal cortex. Neuron. 2011; 71:1127–1140. [PubMed: 21943608]
40. Padoa-Schioppa C, Assad JA. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. Nat Neurosci. 2008; 11:95–102. [PubMed: 18066060]
41. Cai X, Padoa-Schioppa C. Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation. Neuron. 2014; 81:1140–1151. [PubMed: 24529981]
42. Saez A, Rigotti M, Ostojic S, Fusi S, Salzman CD. Abstract context representations in primate amygdala and prefrontal cortex. Neuron. 2015; 87:869–881. [PubMed: 26291167]
43. Neter, J.; Wasserman, W.; Kutner, MH. Applied linear statistical models: regression, analysis of variance, and experimental designs. Irwin, Homewood, IL: 1990.

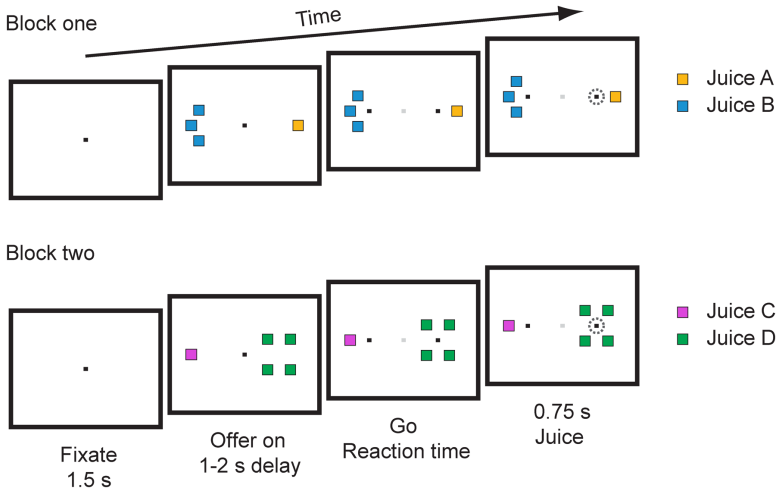

Figure 1.

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 (see Online Methods).

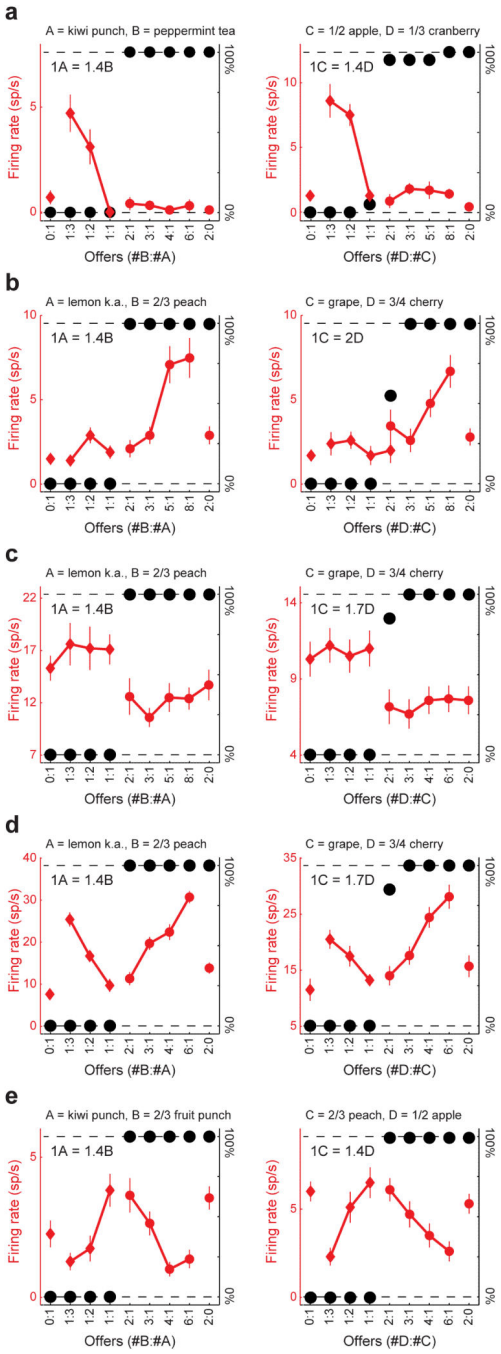

Figure 2.

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, 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/circles indicate trials in which the animal chose the preferred/non-preferred juice, and error bars indicate SEM. (**a**) Cell encoding the

offer value A (kiwi punch) in the first block and the *offer value C* (apple juice) in the second block. (**b**) Cell encoding the *offer value B* (peach juice) in the first block and the *offer value D* (cherry juice) in the second block. (**c**) Cell encoding the *chosen juice A* (lemon Kool-Aid) in the first block and the *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-juice 2, (d) post-juice 2, (e) late delay (see Online Methods). For these neurons, the number of trials in each trial block was as follows: (a) 179, 179, (b) 180, 180, (c) 180, 180, (d) 180, 180, (e) 200, 181.

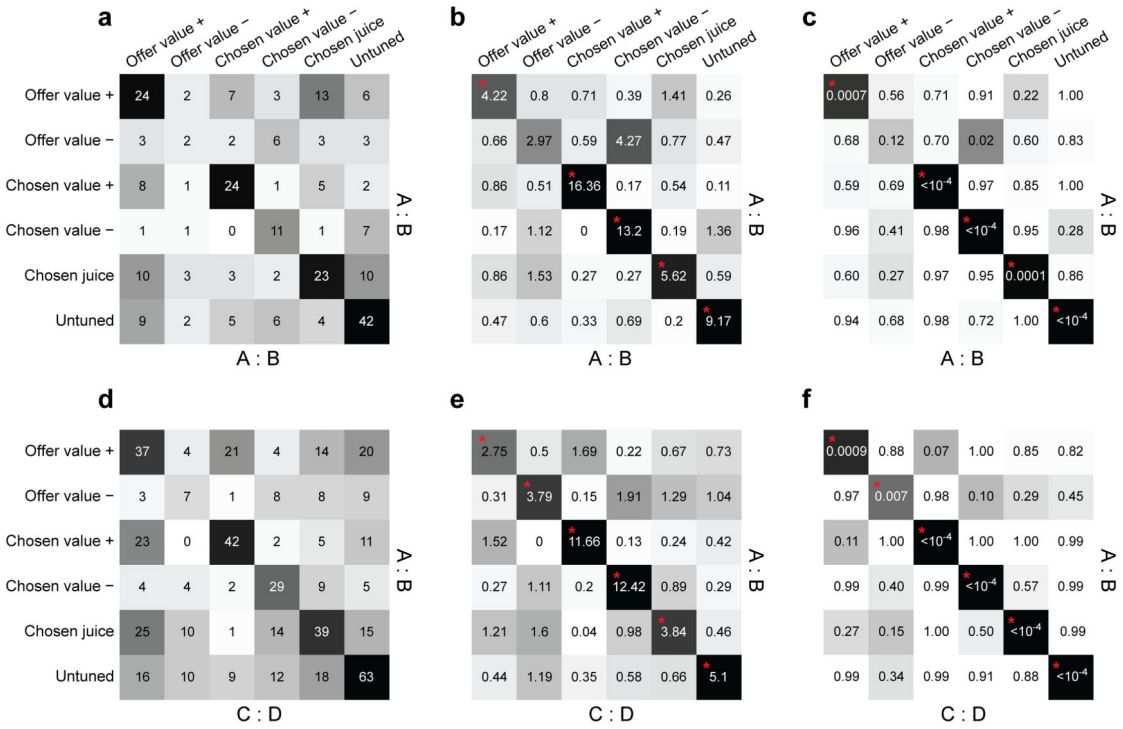

Figure 3.

Comparing classifications across trial blocks. This population analysis was conducted separately for control cells (panels a,b,c) and experimental cells (panels d,e,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 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 represent the odds ratios obtained for each location in panel a. Chance level is 1, and numbers >1 (<1) indicate that the cell count was above (below) that expected by chance. For each location, we performed Fisher's exact test. Red asterisks indicate that the cell count was significantly above chance ($p<0.01$). Cell counts significantly below chance are not indicated here. However, **Supplementary Fig. 6a** shows all the exact p values. (**c**) Bootstrap analysis, control cells. Numbers represent the p values obtained from the bootstrap analysis. Red asterisks indicate that the cell count was significantly above chance ($p<0.01$). (**d**) Contingency table, experimental cells (N = 504). Same format as in panel 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 panel b. **Supplementary Fig. 6b** shows the exact p values. (**f**) Bootstrap analysis, experimental cells. Same format as in panel c.

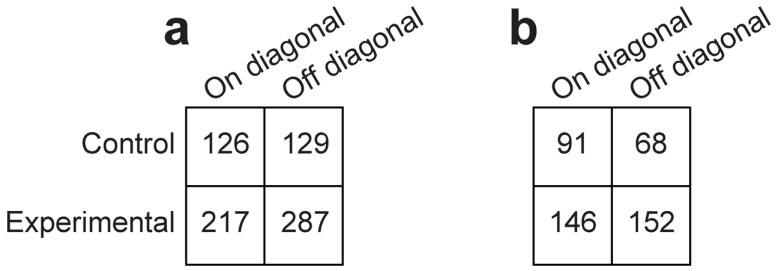

Figure 4.

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 **Fig. 3a** and **Fig. 3d**. (**b**) Cross-classification obtained excluding cells classified as *chosen value* in either block (N = 457). Same format as in panel a.

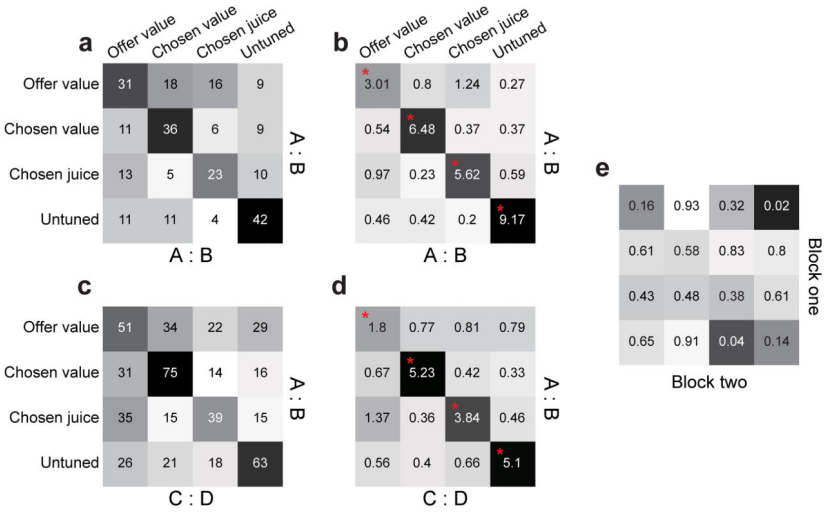

Figure 5.

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 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 represent the odds ratios obtained for the corresponding location in panel a. Numbers >1 (<1) indicate that the cell count was above (below) that expected by chance. Red asterisks indicate that the cell count was significantly above chance ($p<0.01$, Fisher's exact test). Exact p values are shown in **Supplementary Fig. 6c**. (**c**) Reduced contingency table, experimental cells (N = 504). Same format as in panel a. (**d**) Analysis of odds ratios, experimental cells. Same format as in panel b. Exact p values are shown in **Supplementary Fig. 6d**. (**e**) Results of element-wise comparisons of panels b and d. Numbers indicate the p values obtained from the Breslow-Day tests (d.f. = 1). For this test, the null hypothesis represent 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 Fig. 7**.

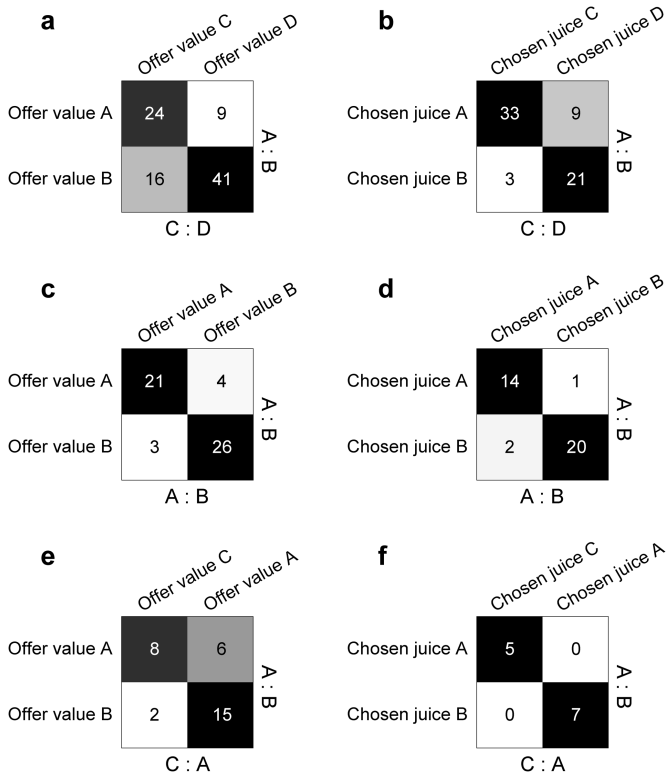

Figure 6.

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, 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, b. In every panel, the Cell count on the main diagonal was significantly higher than expected by chance in all panels. Exact p values were as follows: (a) $p = 5.8\times10^{-5}$, (b) $p = 2.8\times10^{-7}$, (c) $p = 4.2\times10^{-8}$, (d) $p = 2.7\times10^{-7}$, (e) $p = 0.018$, (f) $p = 0.0013$ (Fisher's exact test).