Orbitofrontal cortex neurons code utility changes during natural reward consumption as correlates of relative reward-specific satiety

Alexandre Pastor-Bernier, Arkadiusz Stasiak and Wolfram Schultz*

Department of Physiology, Development and Neuroscience
University of Cambridge
Cambridge CB2 3DY
United Kingdom

Corresponding Author: Wolfram Schultz
Email: Wolfram.Schultz@protonmail.com

Author Contributions: APB and WS designed the experiment, APB conducted experiments, APB and AS analyzed data, APB and WS wrote the paper.

Competing Interests: The authors declare no competing interests.

Keywords: Stochastic choice, psychophysics, indifference curve, revealed preference, multi-component choice option

Abstract

Natural, on-going reward consumption can differentially reduce the subjective value ('utility') of specific rewards, which indicates relative, reward-specific satiety. Two-dimensional choice indifference curves (IC) represent the utility of choice options with two distinct reward components ('bundles') according to Revealed Preference Theory. We estimated two-dimensional ICs from stochastic choices and found that natural on-going consumption of two bundle rewards induced specific IC distortions that indicated differential reduction of reward utility indicative of relative reward-specific satiety. Licking changes confirmed satiety in a mechanism-independent manner. Neuronal signals in orbitofrontal cortex (OFC) that coded the value of the chosen option followed closely the consumption-induced IC distortions within recording periods of individual neurons. A neuronal classifier predicted well the changed utility inferred from the altered behavioral choices. Neuronal signals for more conventional single-reward choice options showed similar relationships to utility alterations from on-going consumption. These results demonstrate a neuronal substrate for the differential, reward-specific alteration of utility by on-going reward consumption reflecting reward-specific satiety.

Significance

Repeated delivery reduces the subjective value ('utility') of rewards to different degrees depending on their individual properties, a phenomenon commonly referred to as sensory-specific satiety. We tested monkeys during economic choice of two-component options. On-going consumption differentially reduced reward utility in a way that suggested relative reward-specific satiety between the two components. Neurons in the orbitofrontal cortex (OFC) changed their responses in close correspondence to the differential utility reduction, thus representing a neuronal correlate of relative reward-specific satiety. Control experiments with conventional single-component choice showed similar satiety-induced differential response reductions. These results are compatible with the notion of OFC neurons coding crucial decision variables robustly across different satiety levels.

Introduction

Consumption of rewards can reduce their attraction. A classic case is food consumption. While eating a meal of vegetables and meat, we may soon come to prefer the vegetables and stop eating the meat. From these choices, we infer that the vegetables have lost less value for us than the meat. Such consumption-induced specific value loss is often referred to as sensory-specific satiety. Experimentally, sensory-specific satiety is achieved by satiation with one particular reward without altering the intake of another, control reward. Satiation can be induced in several ways. Explicit tests consist of rapidly and abundantly feeding with a test reward. Its controlled nature makes this method popular for studying the effects of general or sensory-specific satiety on neuronal and behavioral functions of orbitofrontal cortex (OFC) and midbrain (Rolls et al., 1989; Critchley and Rolls, 1996; Small et al., 2001; Kringelbach et al., 2003; Gottfried et al., 2003; Izquierdo et al. 2004; Bouret & Richmond, 2010; Rudebeck et al., 2013; Murray et al., 2015). Another explicit but less artificial test consists of repeatedly feeding smaller quantities while conducting intervening tests (Yaxley et al., 1985; Rolls et al. 1988). Opposite to intentional induction, spontaneous variations in thirst and hunger change the subjective reward value while keeping physical properties unchanged and serve for identifying subjective reward value signals in monkey OFC (Padoa-Schioppa & Assad, 2006). Despite the multitude of these heterogenous data, satiety tests that combine controllability with natural satiation are scarce. Such tests should involve neither bolus administration nor spontaneous thirst or hunger variations.

Testing reward-specific satiety requires comparison between a reward on which an animal is sated and another reward on which the animal is less or not sated. This two-reward requirement matches the fact that choice options can have multiple components. For example, a meal is composed of vegetables and meat, and choosing a particular meal concerns both food stuffs. The multi-component nature is conceptualized in Revealed Preference Theory; its two-dimensional indifference curves (IC) graphically display reward preferences that are revealed by measurable choice (Fisher, 1892; Samuelson, 1937; Samuelson, 1938). The preferences may be fixed, as the theory assumes, or they may be constructed on the fly at the time of choice (Payne, Bettman, & Schkade, 1999; Simonson, 2008; Dhar & Novemsky, 2008; Kivetz, Netzer & Schrift, 2008; Warren, McGraw & Van Boven, 2011). We estimated ICs in rhesus monkeys that represented their revealed preferences for multi-component reward bundles in an orderly manner. The animals' choices satisfied necessary requirements for rationality, including completeness (preference for one or the other option, or indifference), transitivity, and independence of option set size (Pastor-Bernier et al., 2017). These tests with two-component choice options seem appropriate for testing reward-specific satiety with two differentially sated rewards. Specifically, as ICs represent the integrated economic utility of all bundle rewards, how would a change in the utility of one bundle reward relative to the other reward change the IC shape? And how would neuronal responses reflect these relative utility changes? During typical experimental sessions, the natural on-going consumption of the two bundle rewards would allow to study differential utility changes reflecting relative reward-specific satiety, without requiring artificial bolus administration or spontaneous variations in thirst or hunger.

Here we used the IC scheme to investigate the influence of natural, on-going reward consumption on OFC neurons. We built on our earlier study on OFC neurons whose responses followed the IC scheme; monotonic change with increasing utility irrespective of specific bundle composition, and equal response with equally preferred but differently composed bundles (Pastor-Bernier et al., 2019). We now report that neuronal reward signals in OFC followed the systematically changed ICs that reflected differential reward value changes from natural, on-going reward consumption during the recording period of individual neurons, compatible with the notion of relative, reward-specific satiety. In doing so, OFC neurons coded the integrated economic utility of all bundle rewards in a systematic and conceptually defined manner.

Results

We presented the monkey simultaneously with two composite stimuli on a horizontally mounted touch screen (binary choice task with two discrete, mutually exclusive and collectively exhaustive options; Figure 1A, B). Two rectangles in each stimulus represented a bundle with two reward components whose individual amounts were indicated by a vertical bar (higher was more). The two components were blackcurrant juice or blackcurrant juice with added monosodium glutamate (MSG) in all bundle types as Reward A, and grape juice, strawberry juice, mango juice, water, apple juice, peach juice or grape juice with added inosine monophosphate (IMG) as Reward B.

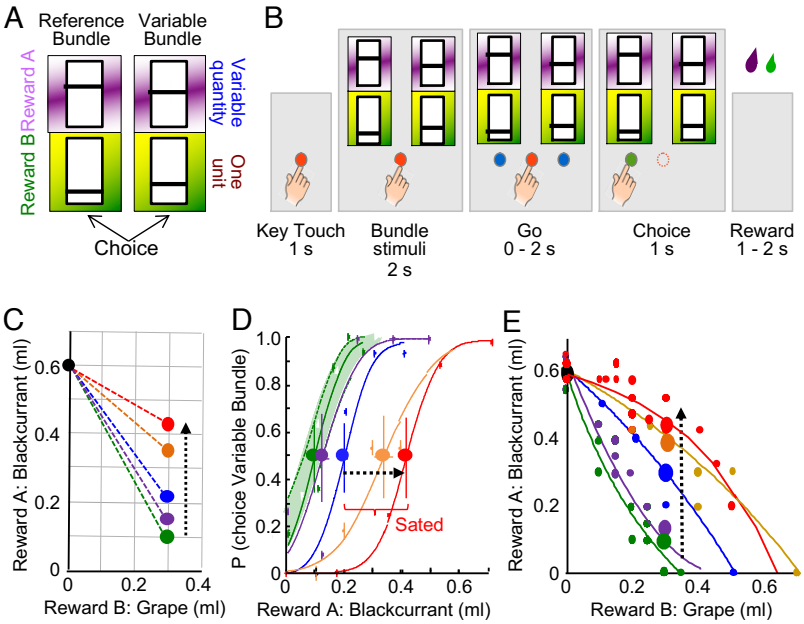

Figure 1. Task, design and behavior

(A) Choice options. Each bundle contained two rewards (A, B) with independently set amounts indicated by the vertical bar position within each rectangle (higher was more). The Reference Bundle contained two preset reward amounts. The Variable Bundle contained a specific amount of one reward and an experimentally varied amount of the other reward.

(B) Task sequence: In each trial the animal contacted a central touch key for 1.0 s; then the two choice options appeared on a computer monitor. After 2.0 s, two blue spots appeared on the monitor, and the animal touched one option within 2.0 s. After touching the target for 1.0 s, the blue spot underneath the chosen bundle turned green as feedback for successful selection, and the blue spot disappeared. The computer-controlled liquid solenoid valve delivered Reward A at 1.0 s after the choice, and Reward B 0.5 s later.

(C) Simplified test scheme: relative reward-specific satiety indicated by decreasing trade-off. With on-going consumption of both juices, the animal gave up progressively less blackcurrant juice for obtaining the same amount (0.3 ml) of grape juice while maintaining choice indifference between the black and one of the colored bundles (from green to red), suggesting utility loss of grape juice relative to blackcurrant juice.

(D) Psychophysical assessment of choice between constant Reference Bundle (0.6 ml blackcurrant juice, 0.0 ml grape juice) and Variable bundle (varying reward A, blackcurrant juice, from 0 to 0.7 ml, while holding grape juice constant at 0.3 ml) (same bundles as in C). Green and violet curves inside green ±95% confidence intervals (CI): initial choices; blue, orange and red curves: on-going consumption; heavy dots: indifference points (IP). Satiety was defined by IPs exceeding CIs. Each curve and IP were estimated from 80 trials in a single block (Weibull fits, Monkey A).

(E) Gradual changes in slope and curvature of ICs between pre-satiety (green, violet) and during increasing satiety (blue, orange, red). Each IC was estimated from fitting to about 35 IPs (Eq. 1), with 80 trials/IP (Monkey A). Small dots indicate IPs, large dots indicate IPs estimated from a single psychophysical test sequence (as shown in (D) with same color convention but from different session).

Basic behavioral design

Our study followed the notions that subjective reward value (utility) can be inferred from observable economic choice, that altered choice would indicate a change in utility, and that a reduction of utility from natural, on-going consumption reflects satiety. The assessment of differential, reward-specific utility change requires at least two rewards. We tested choices between bundles that each had two liquid rewards whose independently variable amounts were represented at the axes and interior of two-dimensional graphs (Figure 1C). We investigated neuronal activity in repeated trials for reasons of statistics and thus tested stochastic, rather than single-shot, choices that are often used on humans.

Pilot tests of all rewards had indicated that blackcurrant juice was least prone to satiety, possibly reflecting taste and/or sugar content differences. Therefore, we designated blackcurrant juice as Reward A for the y-axis of the two-dimensional graph, whereas all other liquids constituted Reward B and were plotted on the x-axis. This convention allowed us to estimate the relative value of all rewards in the common currency of blackcurrant juice at choice indifference.

In choice between two bundles, relative reward utility is inferred from the amount of the bundle reward the animal gives up in order to gain one unit of the other reward of the same bundle, without change in bundle utility (Marginal Rate of Substitution, MRS); unchanged bundle utility is evidenced by maintained equal preference in the trade-off between the old bundle and the new bundle (choice probability of $P = 0.5$ for each option) (Figure 1C, black dot vs. colored dots). By contrast, a binary choice between a single reward and its alternative does not amount to a trade-off in the stricter sense of giving up something one already has for obtaining something one does not yet have; in this more simple binary choice, either one or the other reward is obtained but nothing already owned is given up. With the true trade-off during multi-component bundle choice, only parts of each bundle are exchanged, and any relative utility change with on-going reward consumption is manifested as altered trade-off slope between the two bundles being chosen (black dot vs. colored dots; MRS change). In addition to allowing a true trade-off, the design with two bundle components allows to test bundles with intermediate values between the x- and y-axes.

Consumption-induced relative utility reduction

At the onset of a daily experiment, the black and green bundles of Figure 1C were chosen with equal probability. When choosing the green bundle, the animal gave up 0.5 ml of blackcurrant juice (from 0.6 ml to 0.1 ml) to gain 0.3 ml of grape juice. With on-going consumption of both juices the value ratio between the rewards (trade-off amount) changed: to gain the same 0.3 ml amount of grape juice, the animal gave up progressively less blackcurrant juice, from 0.45 ml via 0.38 ml and 0.25 ml to finally only 1.8 ml (Figure 1C; upward arrow, from violet via blue and orange to red). Thus, the slope between the two bundles on the two-dimensional graph changed as the animal 'payed' progressively less blackcurrant juice for the same amount of grape juice.

We set both rewards in the Reference Bundle, and one reward of the Variable Bundle, to specific amounts, varied psychophysically the amount of the other Variable Bundle reward over the whole testing range, and then fitted a Weibull function to the choice probabilities in order to estimate the amount of the variable reward at which both bundles were chosen with equal probability. For example, in choices between a Reference Bundle, which contained only blackcurrant juice, and a Variable Bundle with fixed amount of grape juice and variable amount of blackcurrant juice, on-going consumption of both juices required increasing amounts of blackcurrant juice for choice indifference (Figure 1D, heavy dots). The rightward shift of the indifference point (IP) from green via violet, blue and orange to red indicated that the animal became gradually more reluctant to give up blackcurrant juice for obtaining the same amount of grape juice; apparently, grape juice had lost more value compared to blackcurrant juice. As each IP was estimated psychophysically in 80 trials, satiety as studied here progressed in test blocks rather than on a trial-by-trial basis. The initial two IPs were close together (green and violet within green 95% confidence interval, CI), suggesting initially maintained reward value, whereas the next IPs outside the CI were considerably higher and indicated substantial value loss (blue, yellow and red

IPs). In other words, the MRS declined with on-going consumption, as schematized in Figure 1C. We assumed that the value change inferred from IP positions outside the CI indicated satiety.

At choice indifference between the two bundles, the amounts of the two Variable Bundle rewards defined an IP (Figure 1E). A new IP was obtained by setting the Reference Bundle to a previously estimated IP position, then setting one reward of the Variable Bundle to a specific amount, varying its other reward psychophysically and estimating choice indifference from curve fitting. Repetition of this procedure, in pseudorandomly alternating directions to avoid local distortions (Knetsch, 1989), resulted in a series of equally preferred IPs. We used these IPs to fit two-dimensional indifference curves (IC) whose slope and curvature reflected the utility of one bundle reward relative to the other bundle reward (Figure 1E; see Methods; Eq. 1). Thus, on-going reward consumption resulted not only in slope change (Figure 1C) but in more informative monotonic IC curvature change from convex (green) via near-linear (blue) to concave (red), which provided systematic evidence for the animal's increasing reluctance to give up blackcurrant juice unless receiving more substantial amounts of grape juice. Both IC changes characterized in a systematic manner the differential reduction of utility of grape juice relative to blackcurrant juice during on-going consumption of both juices, which suggested relative reward-specific satiety for grape juice. These two-dimensional changes were measured during recording periods of individual neurons and constituted our test scheme for behavioral and neuronal correlates of satiety.

For more simple numeric value assessment, we positioned single-component bundles on the x-and y-axes and studied only the ratio between equally preferred rewards, which was graphically represented as two-dimensional slope change (anchor trials). We held blackcurrant juice constant and psychophysically varied grape juice to obtain an IP (Figure S1A-C). With on-going reward consumption, the animal gave up the same constant blackcurrant juice amount only when gaining monotonically increasing grape juice amounts at IP. This change reduced the ratio blackcurrant:grape juice required for choice indifference and suggested relative value reduction of grape juice. The IC curvature showed similar flattening and frequent transition from convex to concave as with the original testing scheme (Figure S1D). The ICs with Monkey B showed similar slope flattening (Figure S1E, F). These tests demonstrate robust value reduction of grape juice with on-going consumption irrespective of the test scheme employed.

Consistency across different bundles

Two rhesus monkeys performed 74,659 trials with the eight bundle types (Figure 2). Given that relative reward-specific satiety would change the ratio of reward amounts at IPs, and the observation that animals sated least on blackcurrant juice, we defined the boundary between pre-sated and sated states by the CI of the initial, left-most choice function between blackcurrant juice and any reward (green in Figures 1D, S1A and S1E); any IP outside this interval indicated utility reduction.

Before satiety, we used a total of 38,443 trials to estimate 56 IPs for fitting 5 ICs with the bundle (blackcurrant juice, grape juice), 68 IPs for 4 ICs with bundle (blackcurrant juice, strawberry juice), 58 IPs for 4 ICs with bundle (blackcurrant juice, water), 38 IPs for 5 ICs with bundle (blackcurrant juice, mango juice) (Monkey B), 65 IPs for 5 ICs with bundle (blackcurrant+MSG, grape+IMP), 55 IPs for 5 ICs with bundle (blackcurrant juice, mango juice), 45 IPs for 3 ICs with bundle (blackcurrant juice, apple juice), and 40 IPs for 2 ICs with bundle (blackcurrant juice, peach juice) (Monkey B).

During satiety, we used 36,216 trials to estimate 52 IPs for 3 ICs with bundle (blackcurrant juice, grape juice), 37 IPs for 4 ICs with bundle (blackcurrant juice, strawberry juice), 63 IPs for 4 ICs with bundle (blackcurrant juice, water), 48 IPs for 5 ICs with bundle (blackcurrant juice, mango juice) (Monkey B), 49 IPs for 4 ICs with bundle (blackcurrant+MSG, grape+IMP), 52 IPs for 4 ICs with bundle (blackcurrant juice, mango juice), 55 IPs for 3 ICs with bundle (blackcurrant juice, apple juice), and 44 IPs for 2 ICs with bundle (blackcurrant juice, peach juice) (Monkey B).

On-going reward consumption induced IC shape changes with all eight bundles in both animals (Figure 2). Stronger satiety for 6 of the 8 liquids (x-axis) relative to blackcurrant (y-axis)

resulted in flattening of IC slopes and transition from convex to linear and concave curvature (Figure S1G, H). However, monkey B seemed to become less sated on peach juice compared to blackcurrant juice, as suggested by steeper ICs (Figure 2H). Together, these IC changes demonstrated robust relative utility loss with natural, on-going liquid consumption across a variety of bundle types.

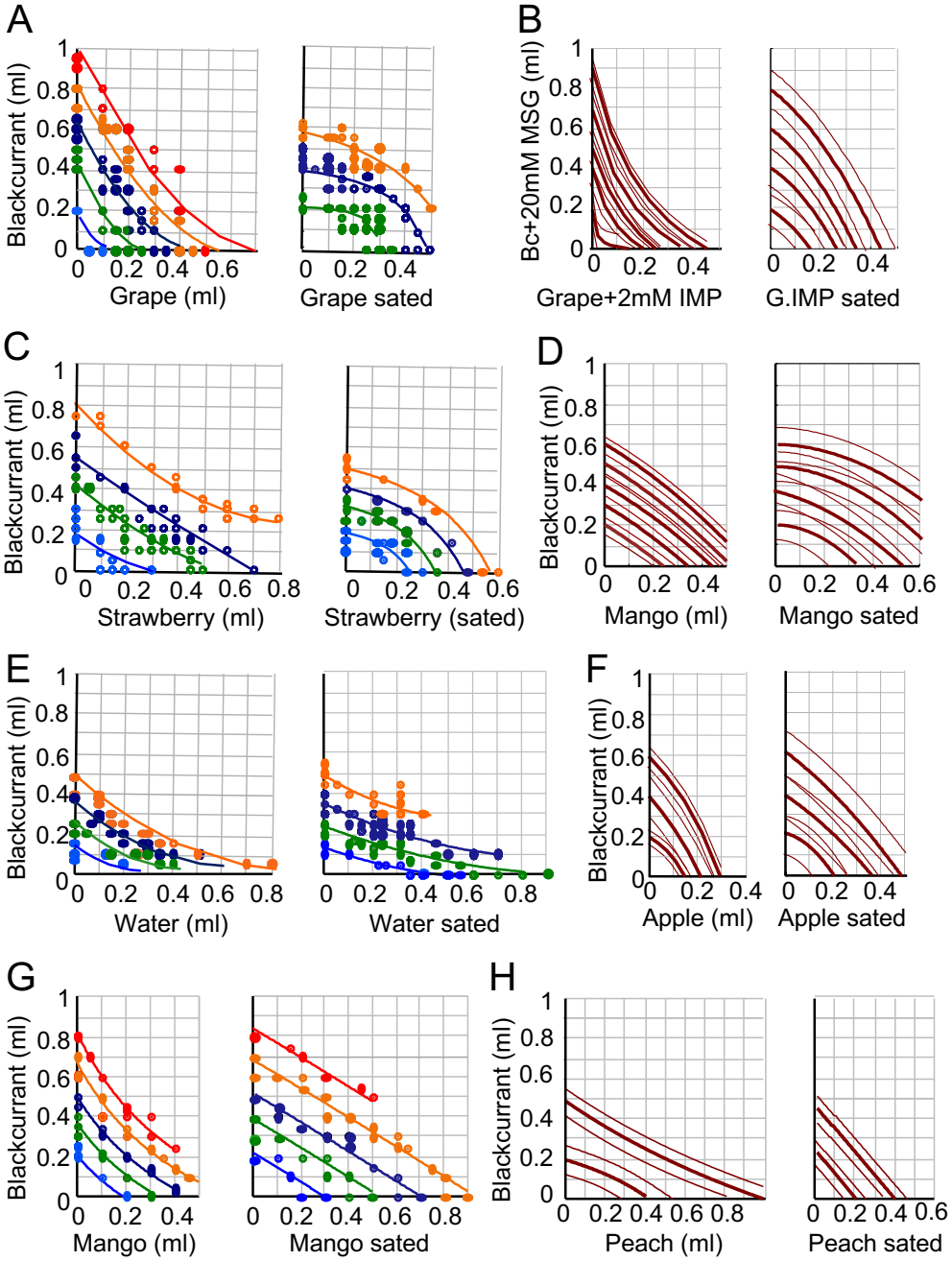

Figure 2. indifference curves reflect relative reward-specific satiety for different bundle types

(A) - (F) Behavioral indifference curves (ICs) for all bundle types used in the current experiment with Monkey A. Lines show ICs fitted hyperbolically to indifference points (IP) of same color (Eq. 1). Dots in A, C, E show measured IPs (choice indifference between all bundles of same color). Dotted lines in B, D, F show ± 95% confidence intervals. Reward A is plotted on the y-axis, Reward B on the x-axis. Bc, blackcurrant juice; MSG, monosodium glutamate; IMP, inosine monophosphate. Same color convention in (A), (C), (E) and (G) as in Figure 1C, D, E.

(G), (H) as (A) but for Monkey B.

Control for other choice variables

A logistic regression served to confirm that bundle choice varied only with the bundle rewards rather than unrelated variables with on-going consumption (Eq. 2). As before satiety (Pastor-Bernier et al. 2019), the probability of choosing the Variable Bundle continued to correlate positively with the amounts of both of its rewards, and inversely with the amounts of both Reference Bundle rewards (Figure S1I; VA, VB vs. RA, RB). Further, choice probability for the Variable Bundle was anticorrelated with the accumulated consumption of blackcurrant juice (MA) and positively correlated with grape juice consumption (MB). This asymmetry is explained by the trade-off at IPs; as grape juice lost more utility than blackcurrant juice during satiety, the animal consumed more grape juice and gave up less blackcurrant juice. Trial number within individual trial blocks (CT) and spatial choice CL) did not explain the choice. Thus, even with on-going consumption, the animals based their choice on the reward amounts of the bundles and the actually consumed rewards according to the experimental design; unrelated variables kept having no significant influence.

Licking and liquid consumption

Lick durations are crude measures for subjective reward value but could serve as mechanism-independent confirmation for the utility changes seen with the choices. Trial-by-trial time courses of lick durations with on-going consumption showed gradual and asymmetric decreases for the bundle rewards (Figure 3A, B). Lick durations remained nearly constant for blackcurrant juice (slope = -2.86 deg, $R^2 = 0.56$; linear regression) but decreased strongly for grape juice (slope = -20.6 deg, $R^2 = 0.50$), suggesting stronger utility loss for grape juice compared to blackcurrant juice. Cumulative licking was significantly shorter in the sated state (green) compared to the pre-sated state (pink) with the main liquids tested in both monkeys (Figure 3C-G). Thus, the reward value changes inferred from lick durations corresponded in direction to those inferred from IC slope and curvature changes (Figure 2).

The IC flattening with on-going consumption indicated that the animal required increasing amounts of the more devalued Reward B for giving up the same amount of the less devalued blackcurrant juice (Reward A) at trade-off. This change was particularly evident in choice between the constant Reference Bundle containing only blackcurant juice (Reward A) and the Variable Bundle containing only one of the other liquids (Reward B) (Figure S1A, B; increase on x-axis). With on-going consumption, the animal gave up the same amount of the less sated blackcurrant juice only if it received increasingly more of the sated Reward B at choice indifference. As the animal had no control over the constant Reference Bundle that defined the IP, the animal ended up consuming more of the devalued reward as the session advanced. For example, with the bundle (blackcurrant juice, water), water consumption increased with beginning satiety more than that of blackcurrant juice (Figure 3H; blue vs. red; $P = 5.0979 \times 10^{-7}$; Kolmogorov-Smirnov test; N = 7,160 trials), thus decreasing the ratio of consumed blackcurrant:water amounts at IP (Figure 3I). Consumption in bundles containing blackcurrant juice together with grape juice, water, strawberry juice or mango juice correlated significantly with their consumption ratios (Rho = 0.3859; P = 0.0056; Pearson). At the end of daily sessions, animals lost complete interest in water and mostly chose the Reference Bundle containing blackcurrant juice alone (right in Figure 3H, not included in Figure 3I).

Thus, the licking changes confirmed in a mechanism-independent manner the relative reward-specific utility changes inferred from bundle choices.

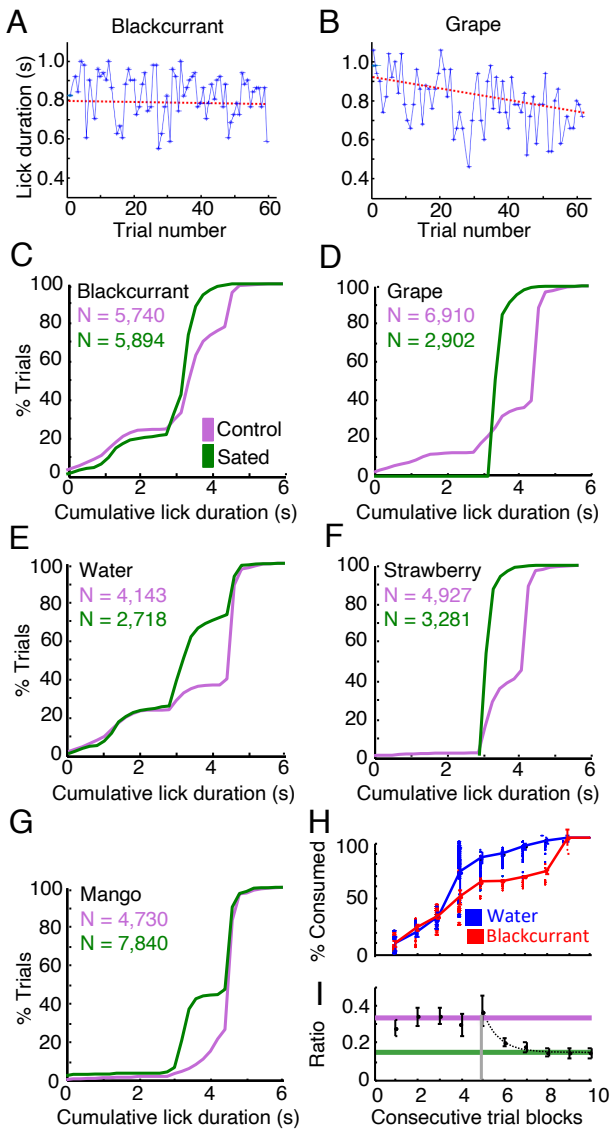

Figure 3. Anticipatory licking and differential juice consumption

(A), (B) Anticipatory licking with bundles (blackcurrant juice, grape juice) with advancing reward consumption within single test sessions (N = 69 and 65 trials, respectively; Monkey A). Red lines show linear regressions of lick duration across trials. Lick durations remained nearly constant for blackcurrant juice, but decreased for grape juice, indicating relative utility loss for grape juice.

(C) - (G) Cumulative distributions of lick durations between bundle appearance and reward delivery for several bundles. Both animals showed significantly more trials with longer lick durations before (pink) than during satiety (green). Monkey A, blackcurrant juice: $P = 5.46 \times 10^{-4}$; Kolmogorov-Smirnov test; N = 5,740 / 5,894 pre-sated/sated trials) grape juice: $P = 2.59 \times 10^{-9}$; N = 6,910 / 2,902, water: $P = 3.60 \times 10^{-3}$; N = 4,143 / 2,718, strawberry juice: $P = 8.66 \times 10^{-6}$; N = 4,920 / 3,281; Monkey B, mango juice: $P = 2.41 \times 10^{-9}$; N = 4,730 / 7,840.

(H) Cumulative consumption of water and blackcurrant juice during 10 advancing blocks and 7,160 anchor trials (each bundle contained only one liquid; see Figure S1A, B for test scheme). For constant blackcurrant amounts (red), the animal consumed significantly more water than blackcurrant as the session advanced (Monkey A).

(I) Exponential reduction of blackcurrant:water ratio from 0.32 (1:3) to 0.15 (1:6) after initial trials (vertical grey line). Single exponential function $f(\beta, x)$: $\beta_1 + \beta_2 e^{(\beta 3x)}$; $[\beta_1, \beta_2, \beta_3] = [0.15, 254.78, -1.41]$ (β_1: final ratio, green line; β_2: decay constant). Consecutive 10 trial blocks for fitting included last block with stable ratio (N= 5,520 trials; Monkey A).

Neuronal test design

We used the IC changes with on-going reward consumption observed in a large variety of bundles to investigate altered value coding in OFC reward neurons. Given the shallower slopes and the less convex and more concave curvatures, we placed bundles on specific segments of the ICs that would change with on-going consumption, such that the physically unaltered bundles would end up on different ICs or IC parts. We subjected most neurons to two tests: (i) during choice over zero-bundle; both rewards were set to zero in one bundle, and the animal unfailingly chose the alternative, non-zero bundle; (ii) during choice between two non-zero bundles; at least one reward was set to non-zero in both bundles, and the animal chose either bundle.

All satiety-tested neuronal responses followed the basic scheme of ICs: monotonic increase with bundles placed on different ICs (testing bundles with different utility), and insignificant response variation with bundles positioned along same ICs (testing equally preferred bundles with equal utility) (Pastor-Bernier et al. 2019). We assessed these characteristics with a combination of multiple linear regression (Eq. 3), Spearman rank-correlation, and 2-way Anova (see Methods). All tested responses belonged to the subgroup of OFC neurons that were sensitive to multiple rewards and coded the value of the bundle the animal chose ('chosen value', as defined by Eqs. 4 and 5). Our task design aimed for maximal similarity between the two choice options and therefore used quantitative bundle stimuli that were visually not unequivocally identifiable; therefore, we could not test object value or offer value that indicate the value of an identifiable choice option.

We tested the influence of on-going reward consumption during the recording period of individual neurons, which allowed us to compare responses in non-sated vs. sated states for the same neuron, as defined by IPs inside vs. outside 95% CIs, respectively (Figure 1D, green zone). As these tests required several tens of minutes with each neuron, neurons not coding chosen value were not further investigated. Our test involved two bundle placements that considered the IC properties: variation of blackcurrant juice while holding grape juice constant, and variation of grape juice while holding blackcurrant juice constant. Comparison of the x-y plots between the pre-sated state (Figure 4A and B) and the sated state (C and D) illustrates this test scheme. The IC flattening with satiety moved the bundle positions relative to the ICs substantially for grape juice variation (compare B and D) but very little for blackcurrant juice variation (compare A with C). Thus, tests following this design should be sensitive for detecting neuronal changes with satiety.

Single-neuron value-coding follows IC changes

At the beginning of daily testing, neuronal responses during choice over zero-bundle followed monotonically the increase of both bundle rewards, confirming value coding (Figure 4A, B). The ICs changed with on-going reward consumption. Despite the change, bundles aligned according to increasing blackcurrant juice were still positioned on different ICs, and the neuronal responses correspondingly continued to distinguish reward value (although in this case only between the top two utilities) (Figure 4C; red vs. blue-green). By contrast, as the ICs flattened and became concave, the three physically unaltered bundles aligned with increasing grape juice were now positioned on or near only one IC (Figure 4D, left), indicating similar utility despite increasing grape juice. The IC concavity indicated that the animal was only ready to give up meaningful amounts of blackcurrant juice (on which it was less sated) when higher amounts of grape juice were offered (right, descending part of IC). Correspondingly, OFC responses failed to vary with grape juice amounts on the flat part of the sated IC (Figure 4D, right), and the response peak for the largest grape juice amount dropped by 75%. Thus, with on-going consumption of both juices, neuronal coding was maintained but reflected the utility reduction of grape juice relative to blackcurrant juice indicated by the corresponding IC changes.

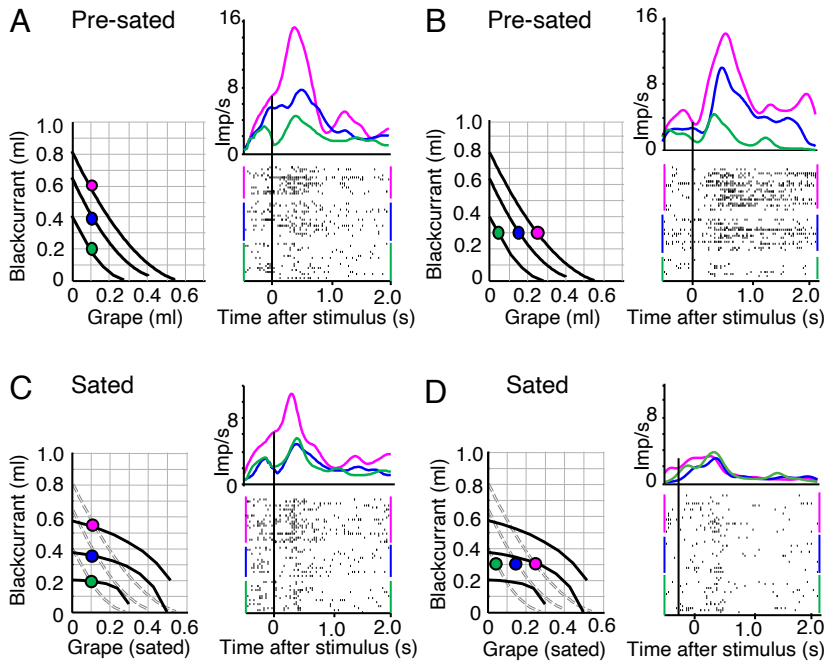

Figure 4. Response change in single OFC neuron reflecting relative reward-specific satiety

(A) Monotonic response increase across three indifference curves (IC) with increasing blackcurrant juice before satiety during choice over zero-bundle. Each colored dot indicates a bundle with specific amounts of blackcurrant and grape juice located on a specific IC. Responses varied monotonically and significantly across ICs with increasing blackcurrant juice (grape juice remained constant) (P = 0.0053, F = 8.88, 36 trials; 1-way Anova).

(B) As (A) but significant response variation with grape juice across ICs (blackcurrant juice remained constant) (P = 1.97141 x 10^{-6}, F = 39.73, 25 trials). Same colors as in (A).

(C) After on-going consumption of both bundle rewards while recording from same neuron: lack of effect for unsated blackcurrant juice. Despite IC change, the three bundles remained on their three original and separate ICs, and neuronal coding of blackcurrant juice remained significant (P = 0.0029, F = 10.28, 36 trials). Note 29% reduction of peak response, from 15.5 to 11 impulses/s (red), and indiscriminate responses between intermediate and low bundles. Grey dotted lines repeat the ICs before satiety shown in (A).

(D) Neuronal response change for sated grape juice: response reduction by 75% (from 15.2 to 3.8 imp/s at peak, red), and loss of significant variation (P = 0.1116, F = 2.68, 34 trials). After the consumption-induced slope and curvature change of the ICs (from convex to concave), the three physically unchanged bundles lie now on or close to the same, intermediate IC, indicating similar utility among them and reflecting satiety for grape juice. Dotted ICs are from pre-sated state. Thus, while continuing to code reward value (C), the responses followed the satiety-induced IC change.

The same consumption-induced neuronal changes occurred in choice between two non-zero bundles (Figure S2). Bundles varying only in blackcurrant juice remained on similarly increasing ICs as before; correspondingly, the chosen value OFC responses continued to increase with blackcurrant juice, confirming basically unaltered coding of blackcurrant juice (Figure S2A, C). By contrast, the three physically unaltered bundles varying only in grape juice were now on lower and narrower spaced ICs, indicating lower and less different values; the neuronal responses decreased correspondingly and became less differential (Figure S2B, D; red, blue, green). Further, the responses to the physically unaltered bundle whose position had changed from intermediate to highest IC (hollow blue) now dominated all other responses (Figure S2D right, dotted blue line). Finally, before satiety the bundle containing only 0.6 ml blackcurrant juice had similar utility as the bundle with only 0.4 ml grape juice (Figure S2B; hollow and solid blue dots on same IC), and correspondingly drew similar neuronal responses (dotted and solid blue lines), whereas with satiety

the physically same two bundles were positioned on different ICs (Figure S2D; hollow vs. solid blue dot) and correspondingly drew different responses (dotted vs. solid blue line). Thus, the differential neuronal response changes with on-going reward consumption occurred irrespective of choice over zero-bundle or choice between two non-zero bundles.

Taken together, OFC neurons continued to code reward value with on-going reward consumption. The responses continued to discriminate well the amount of blackcurrant juice whose utility had changed relatively less (Figures 4A, 4C, S2A, S2C) but were altered for grape juice whose relative utility had dropped more (Figures 4B, 4D, S2B, S2D). The altered OFC signals reflected the reward-specific relative utility change induced by on-going consumption as inferred from the altered ICs.

Neuronal population

We investigated the effects of on-going reward consumption in a total of 272 task-related OFC neurons in area 13 at 30-38 mm anterior to the interaural line and lateral 0-19 mm from the midline; these neurons were parts of the population reported previously (Pastor-Bernier et al., 2019). Responses in 98 of these OFC neurons coded chosen value (defined by Eqs. 4 and 5) and followed the IC scheme in any of the four task epochs (Bundle stimulus, Go, Choice or Reward) during choice over zero-bundle or choice between two non-zero bundles (Table 1). Of the 98 chosen value neurons, 82 showed satiety-related changes with bundles composed of blackcurrant juice (component A) and grape juice, water or mango juice (component B) (Table 2).

We tested averaged z-scored neuronal population responses with the same scheme of bundle alignment on ICs as with single neurons; the scheme is shown in Figures 4 and S2. Bundles aligned with blackcurrant juice (component A) remained on the same three ICs during satiety; by contrast, with consumption-induced IC flattening, bundles aligned with grape juice, water or mango juice (component B) that were on different ICs before satiety were now very close to a single, intermediate IC with little utility variation (see left x-y maps in Figures 4 and S2). The population of 101 positive value coding responses in 31 neurons continued to vary with blackcurrant juice amount during satiety in any task epoch (Bundle stimulus, Go, Choice or Reward), although with 12% peak reduction (Figure 5A, B). By contrast, neuronal coding of reward amounts of component B in the same neurons went from significant before satiety to insignificant during satiety, with 43% peak reduction (Figure 5C, D). Thus, the neuronal population responses showed similar alterations as single neuron responses.

Numeric quantification of individual responses demonstrated satiety-induced significant response reduction with positive value coding neurons (higher response with higher value) and significant response increase with negative (inverse) coding neurons (lower response with higher value) during choice over zero-bundle (Figure 5E and F, red) and during choice between two non-zero bundles (Figure 5G and H, red; Table 2). These responses changes reflected the differential reduction in reward value from on-going reward consumption. By contrast, a minority of neurons showed either inverse changes that were difficult to reconcile with the changes in value (black in Figure 5E-H), or no significant changes at all.

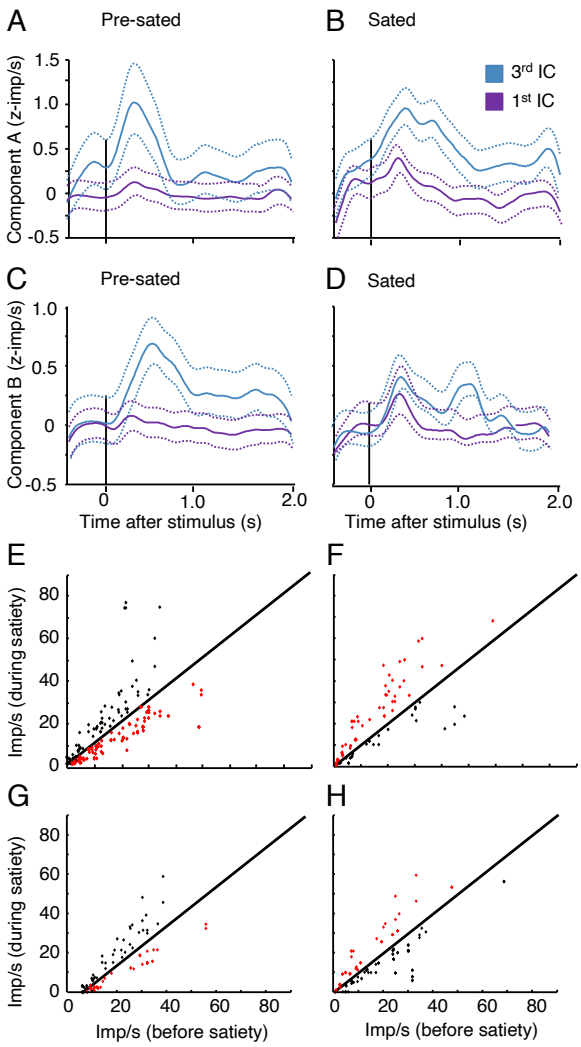

Figure 5. Population responses

(A) - (D) Averaged z-scored population responses from 31 positive coding neurons showing response reduction during satiety. Each part shows responses to bundles on lowest and highest of three indifference curves (IC) during choice over zero-bundle. Data are from choice over zero-bundle, both animals, four bundle types (component A: blackcurrant juice, component B: grape juice, water or mango juice). The response differences between lowest and highest ICs were statistically significant both before satiety (P = 1.53862 x 10^{-5}, F = 19.28, 1-way Anova) and during satiety (P = 2.96646 x 10^{-16}, F = 72.18), but degraded and lost statistical significance with component B (before satiety: P = 4.39918 x 10^{-16}, F = 73.24; during satiety: P = 0.6796, F = 0.17). Dotted lines show ± 95% confidence intervals.

(E) Response changes in positively coding neurons in any of four task epochs (Bundle stimulus, Go, Choice and Reward; Table 2) during choice over zero-bundle. Red: significant response decrease in population reflecting satiety-induced value reduction (P = 7.15 x 10^{-4}; 101 responses in 31 neurons; 1-tailed t-test). Black: significant response increase (P = 0.0014; 69 responses in 21 neurons). Imp/s: impulses/second).

(F) As (E) but for negative (inverse) value coding neurons. Red: significant response increase reflecting satiety-induced value reduction (P = 0.0013; 54 responses in 15 neurons). Black: insignificant response decrease (P = 0.1274; 33 responses in 14 neurons).

(G) As (E) but for choice between two non-zero bundles. Red: response decrease (P = 0.0156; 54 responses in 16 neurons; 1-tailed t-test). Black: response increase (P = 0.0101; 57 responses in 16 neurons). Imp/s: impulses/second).

(H) As (F) but for choice between two non-zero bundles. Red: significant response increase (P = 0.0242; 31 responses in 9 neurons). Black: insignificant response decrease (P = 0.1939; 36 responses in 14 neurons).

Neuronal satiety-induced changes indicated by classification accuracy

Next we used a neuronal classifier as another means for demonstrating how much on-going reward consumption changed neuronal reward coding. We first established the accuracy with which neuronal responses distinguished bundles on different ICs before satiety set in; satiety was defined by IPs exceeding their CIs (Figure 1D). Then we tested the accuracy with which initial neuronal bundle responses distinguished the physically same bundles after on-going reward consumption had changed the ICs. If the neuronal responses had changed substantially with on-going reward consumption, classification accuracy should be low when a classifier trained on bundle responses before satiety was tested for bundle discrimination after satiety. To this end, we trained a support vector machine (SVM) classifier on neuronal responses to randomly selected bundles positioned on the lowest and highest of three ICs, respectively.

The classifier trained on neuronal responses to bundle stimuli before satiety showed decent bundle discrimination with as few as five neurons during choice over zero-bundle; classifier performance was intuitively meaningful as it increased with added neurons (Figure 6A). However, accuracy dropped dramatically when the same classifier trained before satiety was tested for bundle distinction between different ICs during satiety; the maintained accuracy increase with added neurons demonstrated valid classification. Inversely, accuracy was high when training and testing the classifier during satiety (Figure 6B), but lower when training during satiety and testing for bundle distinction before satiety, thus confirming the neuronal changes with satiety.

These accuracy differences were seen during choice over zero-bundle with neuronal responses to Bundle stimuli (Figure 6) and during the Go epoch (Figure S3A), but not during Choice and Reward epochs (Figure S3B, C). The changes were not explained by baseline changes during the 1 s Pretrial control epoch (Figure S3D). Similar substantial accuracy differences were seen in choice between two non-zero bundles during Bundle stimuli, Go epoch and Choice epoch but not during the Reward epoch (Figure S3E-H), again not explained by baseline changes (Figure S3I). The accuracy differences were consistent across on-going consumption steps (Figure S3J).

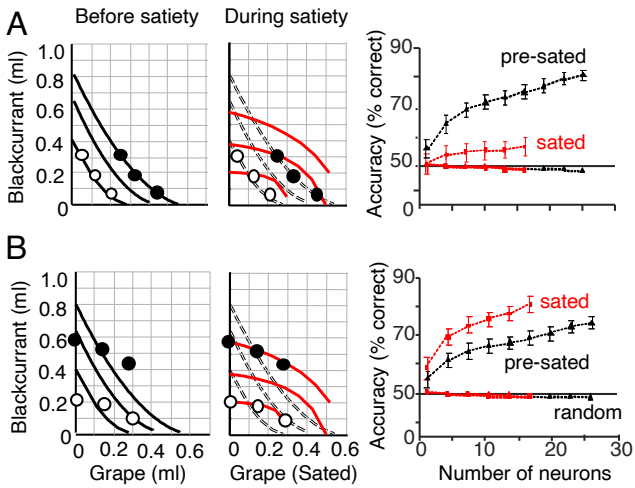

Figure 6. Bundle classification demonstrates satiety-induced change of neuronal value coding

(A) Bundle classification by support vector machine using neuronal responses to stimuli of bundles positioned on the lowest and highest indifference curve, respectively (choice over zero-bundle). The classifier was trained on neuronal responses before satiety and tested for bundle distinction before satiety (black) and during satiety (red). Left: identical bundle positions on two-dimensional map but IC change with on-going consumption, indicating satiety-induced relative utility change (red). Right: classifier accuracy increase with neuron numbers before satiety (black), but drop when tested for bundle distinction during satiety (red). Error bars indicate standard errors of the mean (SEM).

(B) As (A), but reverse order: classifier trained on neuronal responses during satiety and tested before satiety.

In demonstrating substantial accuracy changes, these tests suggested that the neuronal responses followed the substantial IC changes that reflected the utility changes from on-going reward consumption indicative of satiety.

Neuronal satiety changes with single-reward bundles

Using choice options with two reward components differs in several ways from previous studies using single rewards (Tremblay & Schultz 1999; Padoa-Schioppa & Assad 2006) and thus requires controls and additional analyses. We used the same two visual component stimuli but set only one, but different, reward in each bundle to a non-zero amount, which positioned the bundles graphically along the x-axis and y-axis but not inside the IC map (anchor trials; Figure S1B). These degenerated bundles were equivalent to single-reward choice options tested earlier (Padoa-Schioppa & Assad 2006).

First we used single-reward bundles for confirming the results with our two-component bundles. The responses of the neuron shown in Figure 7A, B distinguished both rewards during choice over zero-bundle before satiety. With on-going consumption of both rewards, the ICs flattened, preserving the blackcurrant juice positions on the ICs but changing the physically unchanged position of the two water amounts relative to the ICs (Figure 7C, D). The neuron kept discriminating blackcurrant juice amounts during satiety (Figure 7C). However, with the satiety-induced IC change, the large water amount was now positioned much more below the highest IC than before (Figure 7D, red on x-axis) and on about the same IC as the small blackcurrant amount (blue on y-axis). Correspondingly, the neuronal activity with the large water amount lost its peak (reduction by 50%) and was now very similar to the activity with the small blackcurrant amount (Figure 7C, D, red dotted vs. blue solid arrows). Further, the position of the small water amount was now below its original IC (blue on x-axis), and the neuron, with its lost response, failed to distinguish between the two water amounts. Thus, the neuronal changes with single-reward bundles followed the satiety-induced IC changes, demonstrating that the neuronal satiety changes reported above occurred also with single rewards (degenerated bundles).

Next, we used single-reward bundles to quantify neuronal response changes with on-going reward consumption in relation to utility changes inferred from behavioral choices. We established vector plots that display the ratio of reward weights (β's) for behavioral choice (Eq. 1a; Figure 7E-I, green) and z-scored neuronal population responses (Eq. 3; red). The inequality of utilities of the two rewards was manifested as deviation of these vectors from the diagonal. On-going reward consumption increased the elevation angle of the behavioral vector, indicating loss of utility for component B (grape juice, water or mango juice) relative to component A (blackcurrant juice). The neuronal vector changed correspondingly (Figure 7E-I, green vs. red). For example, during choice of the bundle (blackcurrant juice, grape juice) over zero-reward bundle, the elevation angle of the behavioral vector increased from 40 deg before satiety to 65 deg during satiety, and the neuronal population vector increased correspondingly from 35 deg to 62 deg (Figure 7E, green, red). Similarly, during choice between two non-zero bundles, the behavioral vector increased from 40 deg to 52 deg and the neuronal vector increased correspondingly from 38 deg to 45 deg (Figure 7F). Further, the shorter neuronal vectors during satiety indicated general reduced responding (red). Bundles containing water or mango juice showed similar changes (Figure 7G-I). Thus, both before and during satiety, the neuronal vectors (red) were within the CIs of the behavioral vectors (green), indicating intact neuronal value coding that followed the utility changes with on-going reward consumption.

In addition to the vector analysis, IC slopes confirmed the close neuronal-behavioral correspondence during satiety, with satiety being defined by the IPs exceeding the initial, pre-sated IPs (Figure S1A, E). As estimated from regression coefficient ratios ($-\beta_2 / \beta_1$) (Eq. 3) and ($-b / a$) (Eq. 1), the slopes of the linear neuronal ICs of single-reward bundles correlated well with the slopes of linear behavioral ICs (Figure 7J). These results from testing single-reward bundles with on-going reward consumption compared well with the results from the earlier OFC study on single rewards with spontaneously varying subjective reward value (Padoa-Schioppa & Assad 2006).

Taken together, the population changes indicated the influence of relative, reward-specific satiety on neuronal reward value coding. They confirmed the single-neuron changes with single-reward bundles (Figure 7A-D) and multi-reward bundles (Figures 4, S2).

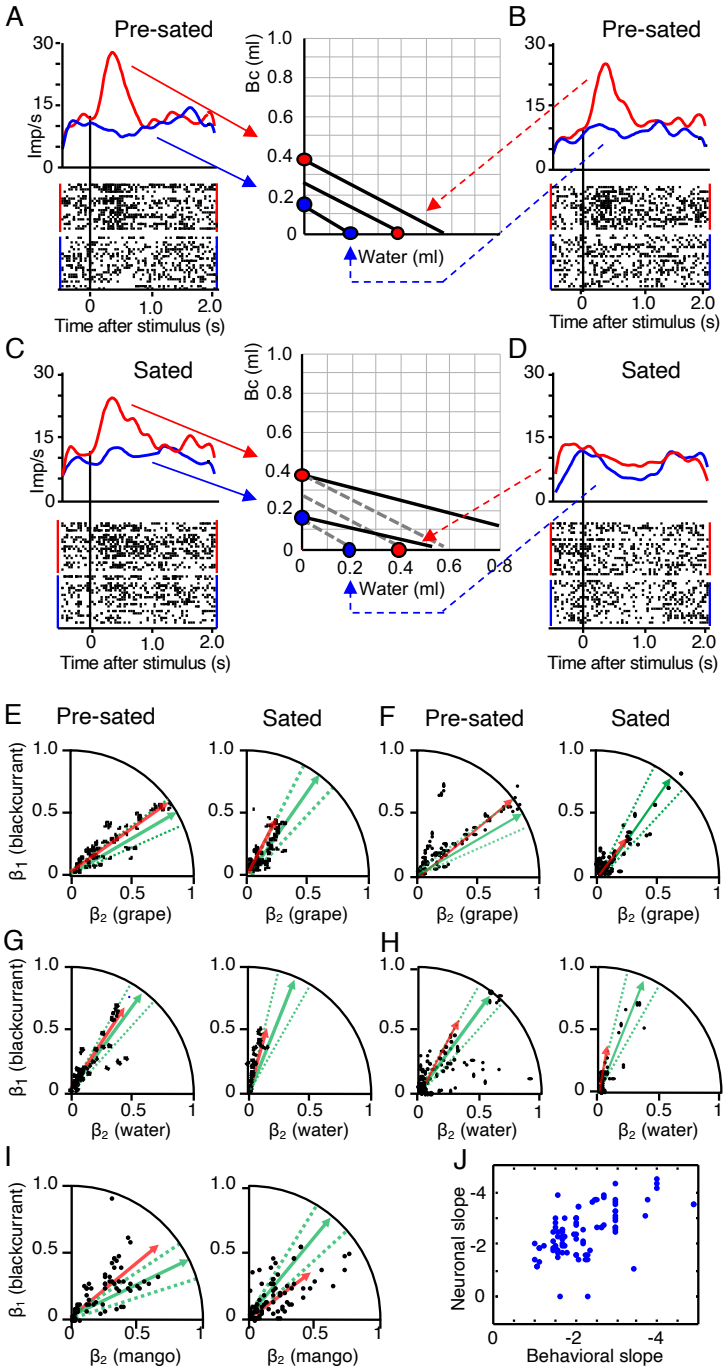

Figure 7. Reward-specific satiety with single-reward bundles

(A-D) Responses of same single neuron before and during satiety. Each bundle contained specific non-zero amounts of only blackcurrant juice or only water (colored dots on indifference curves, ICs) and was tested during choice over zero-bundle.

(A) Significant response increase across two ICs with increasing blackcurrant juice (Bc) before satiety (water remained zero) (red vs. blue; P = 0.0091, F = 6.92, 23 trials; 1-way Anova).

(B) As (A) but significant response variation with increasing water across two ICs (blackcurrant juice remained zero) (P = 0.0113, F = 7.32, 31 trials). Same colors as (A).

(C) Despite IC flattening after on-going reward consumption, the two bundles with blackcurrant juice variation remained on the same two ICs, and the neuronal response variation remained significant (P = 0.002, F = 11.04, 40 trials), and the peak response was only slightly reduced (red). Dotted ICs are from pre-sated state.
(D) IC flattening after on-going reward consumption indicates relative utility reduction of water. The two unchanged bundles with water variation were now located below and at the IC. The neuronal response was substantially reduced by 50% (red) and had lost significant variation (P = 4337, F = 0.64, 40 trials). Further, the large-water bundle (dashed red line) elicited now a similar response as the low-blackcurrant bundle that is now located on the same IC (solid blue line). Thus, while continuing to code reward value (C), the responses followed the satiety-induced IC change.
(E) Vector plots for behavioural choice of bundle (blackcurrant juice, grape juice) over zero-bundle (green) and corresponding z-scored neuronal population responses (black, red). Neuronal vector slopes were 35 deg before satiety and 62 deg during satiety, using all significantly positive and normalized negative (inverse) coding responses from all four task epochs; all included responses followed the IC scheme. Dots refer to neuronal responses, vectors represent averages from behavioral choices (green; dotted lines: 95% confidence intervals) and neuronal responses (red), based on Eqs. 1a and 3, respectively (see Methods). Neuronal slope regression coefficients (β's) on axes refer to Eq. 3.
(F) As for (C) but for choice between two non-zero bundles. Neuronal vector slopes were 38 deg before and 45 deg during satiety.
(G), (H) As (E, F) but for bundle (blackcurrant juice, water).
(I) As (E) but for bundle (blackcurrant juice, mango juice).
(J) Correlation between rectified neuronal and behavioral IC slopes (β's from Eqs. 3 and 1a, respectively) during satiety in all tested neurons (rho = 0.604; P = 8×10^{-6}, Pearson correlation; rho = 0.595, P = 2×10^{-5}, Spearman rank-correlation; N = 90 responses during choice between two non-zero bundles).

Discussion

This study tested binary choice between bundles of two rewards and found response changes in OFC reward neurons that suggested a differential loss of reward utility indicative of relative reward-specific satiety from on-going reward consumption. The choices were captured by graphic ICs that represented the relative utilities of the two bundle rewards in a conceptually rigorous manner. The ICs changed in an orderly and characteristic manner with on-going reward consumption, without requiring unnatural reward bolus administration (Figures 1, 2, S1). The ICs flattened progressively and showed gradual curvature changes from convexity to concavity, which indicated gradual utility loss for one bundle reward (blackcurrant juice, plotted on the y-axis) relative to the other bundle reward (all other bundle rewards except peach juice, x-axis). This IC change suggested that the animal became increasingly reluctant to give up blackcurrant juice for the same increment of the other reward. The specific and asymmetric IC changes make alternative explanations unlikely, such as passage of time, general satiety, loss of motivation, or proximity of return to home cage, all of which would have affected all rewards in a similar manner. Licking behavior supported the notion of differential reward satiety in a mechanism-independent manner (Figure 3).

Our preceding study had established neuronal chosen value responses in OFC that were sensitive to multiple rewards and followed the animal's rational choice of two-reward bundles, including completeness, transitivity and independence from option set size (Pastor-Bernier et al., 2019). The current study tested the effects of on-going reward consumption during the recording period of individual neurons. We found OFC value responses that matched the consumption-induced IC changes. The responses became weaker for the more devalued reward, as indicated by slope and curvature changes of ICs (Figures 4, 5, S2). Most impressively, neuronal responses failed to distinguish between bundles that had landed on the flat parts of ICs because of the ICs' curvature change to concave (Figure 4D). Classifiers predicting bundle discrimination from neuronal responses confirmed accurate reward value coding both before and during satiety and demonstrated the substantial nature of the neuronal changes (Figures 6, S3). Neuronal response vectors of

conventional single-reward choice options correlated well with behavioral choice vectors; these correlations were maintained with the utility change from on-going reward consumption (Figure 7). Taken together, these particularly sensitive reward utility tests, informed by Revealed Preference Theory, demonstrate good correlation between OFC responses and the differential utility alterations induced by on-going reward consumption. As the physical reward properties did not change with satiety, these results also confirm the subjective value (utility) coding of OFC neurons demonstrated with choices between two-component bundles (Pastor-Bernier et al., 2019) and single rewards (Padoa-Schioppa & Assad, 2006).

The observed increase in consumption of sated liquids like water (Figure 3H) seemed to contradict earlier findings and the general intuition that satiety would rather reduce consumption of rewards on which an animal is sated (Rolls et al. 1989; Critchley & Rolls 1996). Differences in study design might explain these discrepancies. When an animal has the choice between a sated and a non-sated reward, or the choice between accepting and not accepting a reward, it would naturally prefer the non-sated reward. This was the case in the cited earlier studies. By contrast, in our study, the animal chose between two bundles that each had two differently sated rewards. As the animal was still interested to obtain the less sated bundle reward, it would inadvertently also receive the other, more sated reward that was a part of the bundle. The animal had no control over the setting of the Reference Bundle against which it would choose the alternative bundle. At the IP, the animal had the choice to give up some of the non-sated reward in order to receive more of the sated reward. If the animal had still a limited interest in the less sated reward, maybe because it was still somewhat thirsty, it might give up a limited amount in order to receive a lot more of the other reward (as long as it would not outrightly reject it, which was not the case). This trade-off was indicated by the increasing IC concavity with on-going consumption, which demonstrated that really large amounts of the more devalued Reward B were required for giving up the less devalued Reward A (Figures 1E, 2, S1D). Outright rejection of Reward B would be represented not by a downward sloped IC but by an upward sloped IC, which had been observed with lemon juice, yoghourt and saline (Pastor-Bernier et al., 2017) but not with the currently used rewards; such upward sloped ICs indicate that an animal needed to be 'bribed' with more reward for accepting these normally rejected rewards. By contrast, in the current satiety experiment, the animal inadvertently consumed more of the sated reward during satiety compared to before, and the maintained downward IC slope indicated that the animal was not entirely averse to the sated reward.

The current study of systematically altered reward value coding with reward-specific satiety builds on previous studies on monkey OFC neurons that investigated satiety in a more basic manner. There are notably the studies from Rolls' laboratory in which monkeys were presented with syringes or tubes containing various fruit juices; rating scales were used to assess behavioral acceptance or rejection of these juices after bolus administration (Rolls et al. 1989; Critchley & Rolls 1996). The studies report OFC neurons that responded to several juices and lost the responses only for the particular juice on which the animal was sated. The response reduction with sensory-specific satiety in OFC contrast with Rolls' studies on earlier stages of the gustatory system, including the nucleus of the solitary tract, the frontal opercular taste cortex, and the insular taste cortex, where no such satiety-related changes were found (Yaxley et al. 1985; Yaxley et al. 1988; Rolls et al. 1988). However, it is unknown whether the studied OFC neurons coded subjective reward value inferred from choices in the absence of satiety or covaried with other crucial aspects of subjective reward value, such as reward amount and behavioral preference that formed the basis for our study. A subsequent study with bolus water reward administration found even stronger general satiety effects in ventromedial prefrontal cortex compared to OFC (Bouret & Richmond, 2010), suggesting widespread satiety effects in the ventral frontal cortex. Our results are compatible with the relationship between spontaneous choice variations and chosen value coding in monkey OFC (Padoa-Schioppa & Assad 2006); the choice variations likely reflected changes in thirst level that were synonymous with satiety variations during the course of each day's experimentation. These value changes were instrumental in distinguishing subjective value coding from the coding of purely physical properties during economic choice. Thus, the current experiment brings together a

number of heterogeneous arguments in favor of OFC coding of subjective value and presents a conceptually coherent argument for economic utility coding according to Revealed Preference Theory.

While reward-specific satiety concerns the specific utility of individual rewards, on-going consumption induces also a reduction of general reward value manifested as changes in arousal, attention and motivation. General satiety effects cannot be distinguished from reward-specific satiety when testing only a single reward, and the effects may be attributed to loss of motivation, as in the case of reduced dopamine responses in mice that received food pellets for extended periods of time (Rossi et al., 2013). The loss of motivation may be associated with a loss of pleasure and development of aversion; neural indices may consist of reduced human midbrain responses, as shown with on-going consumption of Swiss chocolate (Small et al., 2001). In our results, the shorter neuronal population vectors might indicate an effect of general satiety on neuronal responses (Figure 7E-I), in addition to the reward-specific satiety suggested by the changes in vector angle. However, general satiety cannot explain our asymmetric neuronal changes that correlate with relative reward-specific utility changes.

Materials and Methods

The study used the same 2 male adult rhesus monkeys as previously (Pastor-Bernier et al., 2019) and was licensed by the UK Home Office (for details, see Supplementary Information). The animals chose between two bundles that contained the same two rewards with independently varying amounts. We estimated psychophysically multiple choice indifference points (IP; Figure 1D, S1A, E) to which we fitted indifference curves (IC) along which all bundles were equally preferred, using a hyperbolic function d:

$$d = ay + bx + cxy \qquad \text{(Eq. 1)}$$

with y and x as milliliter amount of Rewards A and B (Figures 1C, E, S1B, D, F), a and b as weights of the influence of the two Reward amounts, and c as curvature. Eq. 1 can be equivalently re-written as regression in analogy to the regression used for analysing neuronal responses:

$$y = \beta_0 + \beta_1 A + \beta_2 B + \beta_3 AB + \varepsilon \qquad \text{(Eq. 1a)}$$

with A and B as milliliter amount of Reward A (plotted at y-axis) and Reward B (x-axis), respectively, β_0 as offset coefficient, β_1 and β_2 as behavioral regression coefficients, and ε as compound of errors err_0, err_1, err_2, err_3 for offset and regressors 1-3.

To test whether the animal's choice reflected the amount of the bundle rewards during satiety, rather than other, unintended variables such as spatial bias, we used the logistic regression:

$$P\,(V) = \beta_0 + \beta_1 CT + \beta_2 RA + \beta_3 RB + \beta_4 VA + \beta_5 VB + \beta_6 CL + \beta_7 MA + \beta_8 MB + \varepsilon \qquad \text{(Eq. 2)}$$

with P (V) as probability of choice of Variable Bundle, β_0 as offset coefficient, β_1 - β_7 as correlation strength (regression slope) coefficients indicating the influence of the respective regressor, CT as trial number within block of consecutive trials, RA as amount of Reward A of Reference Bundle, RB as amount of Reward B of Reference Bundle, VA as amount of Reward A of Variable Bundle, VB as amount of Reward B of Variable Bundle, CL as choice of any bundle stimulus presented at the left, MA as consumed amount of Reward A, MB as consumed amount of Reward B, and ε as compound error for offset and all regressors.

Following behavioral training and surgical preparation for single neuron recordings, we identified neuronal task relationships with the paired Wilcoxon-test. We identified changes of task-related neuronal responses across ICs with a linear regression:

$$y = \beta_0 + \beta_1 A + \beta_2 B + \beta_3 AB + \varepsilon \quad \text{(Eq. 3)}$$

with y as neuronal response in any of the four task epochs, measured as impulses/s and z-scored normalized to the Pretrial control epoch of 1.0 s, A and B as milliliter amounts of Reward A (plotted at y-axis) and Reward B (x-axis), respectively, β_0 as offset coefficient, β_1 and β_2 as neuronal regression coefficients, and ε as compound error. In addition, all significant neuronal response changes across ICs identified by Eq. 3 needed to be also significant in a Spearman rank-correlation test ($P < 0.05$).

To assess neuronal compliance with the two-dimensional IC scheme, we used a two-factor Anova on each task-related response that was significant for both regressors in Eq. 3. Neuronal responses following the IC scheme were significant across-ICs (factor 1: $P < 0.05$) but insignificant within-IC (factor 2).

Chosen value (CV) was defined as:

$$CV = A + k_1 B \quad \text{(Eq. 4)}$$

weighting parameter k_1 served to adjust for differences in subjective value between rewards A and B, such that the quantity of Reward B entered the regression on a common-currency scale defined by Reward A. We assessed neuronal coding of chosen value in all neurons that followed the revealed preference scheme, using the following regression:

$$y = \beta_0 + \beta_1 CV + \beta_2 UCV + \varepsilon \quad \text{(Eq. 5)}$$

with UCV as value of the unchosen option that was not further considered here, and ε as compound error.

References

Bouret, S., and Richmond, B.J. (2010). Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. *J Neurosci* 30: 8591-8601.

Critchley, H.D., and Rolls, E.T. (1996). Olfactory neuronal responses in the primate orbitofrontal cortex: analysis in an olfactory discrimination task. J. Neurophysiol. 75, 1659-1672.

Dhar, R., & Novemsky, N. (2008). Beyond rationality: the content of preferences. J Consum Psychol 18, 175-178.

Fisher, I. (1892). Mathematical Investigations in the theory of value and prices. Trans Connecticut Acad 9, 1–124.

Gottfried JA, O'Doherty J, Dolan RJ (2003) Encoding predictive reward value in human amygdala and orbitofrontal cortex. Science 301, 1104-1107.

Izquierdo A, Suda RK, Murray EA (2004). Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. J Neurosci 24, 7540-7548.

Kivetz, R., Netzer, O., & Schrift, R. (2008) The synthesis of preference: Bridging behavioral decision research and marketing science. Journal of Consumer Research 18, 179-186.

Kobayashi, S., Pinto de Carvalho, O., and Schultz, W. (2010). Adaptation of reward sensitivity in orbitofrontal neurons. J. Neurosci. 30, 534–544.

Knetsch, J. L. (1989). The endowment effect and evidence of nonreversible indifference curves. Am Econ. Rev. 79, 1277-1288.

Kringelbach, M.L., O'Doherty, J., Rolls, E.T., and Andrews, C. (2003). Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. Cereb. Cortex 13, 1064-1071.

Merrill, E. G., and Ainsworth, A. (1972). Glass-coated platinum-plated tungsten microelectrodes. Med. Biol. Eng. 10, 662–672.

Murray EA, Moylan EJ, Saleem KS, Basile BM, Turchi J (2015). Specialized areas for value updating and goal selection in the primate orbitofrontal cortex. eLife 4: e11695

Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. J. Neurosci. 29, 14004-14014.

Pastor-Bernier, A., Stasiak, A., and Schultz, W. (2019). Orbitofrontal signals for two-component choice options comply with indifference curves of Revealed Preference Theory. Nat. Comm. 10, 4885.

Pastor-Bernier, A., Plott, C.R., and Schultz, W. (2017). Monkeys choose as if maximizing utility compatible with basic principles of revealed preference theory. Proc Natl AcadSci U S A. 114, E1766-E1775.

Paxinos, G., Huang, X.-F. and Toga, A. W. (2000). The Rhesus Monkey Brain in Stereotaxic Coordinates (Academic Press, San Diego).

Payne, J. W., Bettman, J. R., & Schkade, D. A. (1999). Measuring constructed preferences: towards a building code. J Risk Uncert 19, 243-270.

Rolls, E.T., Sienkiewicz, Z.J., and Yaxley, S. (1989). Hunger Modulates the Responses to Gustatory Stimuli of Single Neurons in the Caudolateral Orbitofrontal Cortex of the Macaque Monkey. Eur. J. Neurosci. 1, 53-60.

Rolls, E.T., Scott, T.R., Sienkiewicz, Z.J., and Yaxley, S. (1988). The responsiveness of neurons in the frontal opercular gustatory cortex of the macaque monkey is independent of hunger. J Physiol. 397, 1-12.

Ross, M.A., Fan, D., Barter, J.W. and Yin, H.H. (2013). Bidirectional modulation of substantia nigra activity by motivational state. PLoS ONE 8: e71598.

Rudebeck PH, Saunders RC, Prescott AT, Chau, LS, Murray EA (2013). Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. Nat Neurosci 16: 1140-1145.

Rustichini, A., Conen, K.E., Cai, X., Padoa-Schioppa, C. (2017). Optimal coding and neuronal adaptation in economic decisions. Nat. Comm. 8, 1208.

Samuelson, P. A. (1937). A note on measurement of utility. Rev Econ Stud 4, 155-161.

Samuelson, P. A. (1938). A note on the pure theory of consumer's behavior. Economica 5, 61-71.

Simonson, I. (2008). Will I like a “medium” pillow? Another look at constructed and inherent preferences. J Consum Psychol 18, 155-169.

Small, D.M., Zatorre, R.J., Dagher, A., Evans, A.C., and Jones-Gotman, M. (2001). Changes in brain activity related to eating chocolate: from pleasure to aversion. Brain 124, 1720-33.

Tremblay, L., and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. Nature 398, 704–708.

Tsutsui, K. I., Grabenhorst, F., Kobayashi, S. & Schultz, W. (2016). A dynamic code for economic object valuation in prefrontal cortex neurons. Nat. Comm. 7, 12554.

Warren, C., McGraw, A. P., & Van Boven, L. (2011). Values and preferences: defining preference construction. WIREs Cogni Sci 2, 193-205.

Yaxley, S., Rolls, E.T., Sienkiewicz, Z.J., and Scott, T.R. (1985). Satiety does not affect gustatory activity in the nucleus of the solitary tract of the alert monkey. Brain Res. 347, 85-93.

Yaxley, S., Rolls, E.T., and Sienkiewicz, Z.J. (1988). The responsiveness of neurons in the insular gustatory cortex of the macaque monkey is independent of hunger. Physiol. Behav. 42, 223-229.

Table 1. Numbers of neurons tested with on-going reward consumption

Bundle type	Choice over zero-bundle		Choice between two non-zero bundles	
	Neurons tested	IPs tested	Neurons tested	IPs tested
Blackcurrant, grape	21+11=32	28	7+12=19	38
Blackcurrant, water	20+13=33	39	22+12=34	58
Blackcurrant, mango	14+7=21	11	9+8=17	10
SUM	55+31=86	78	38+32=70	106

The bundle types (blackcurrant, grape) and (blackcurrant, water) were tested in Monkey A (81 and 138 neurons, respectively), whereas bundle type (blackcurrant, mango) was tested in Monkey B (53 neurons). Of these neurons, the neuron and response numbers given above coded chosen value (as indentified by Eqs. 4 and 5) and followed the IC scheme, as defined previously (Pastor-Bernier et al., 2019): monotonic increase or monotonic decrease with bundles compared across ICs, insignificant response variation with bundles compared along individual ICs. Such neurons were recorded only during choice over zero-bundle (N = 28 neurons), only during choice between two non-zero bundles (N = 12 neurons), or both (N = 58 neurons) (total of 98 neurons). In table cells with multiple entries, the first two numbers refer respectively to positive and negative (inverse) relationships to increasing reward quantity, as inferred from the neuronal regression slope (β's in Eq. 3). IP; bundle at choice indifference point at specific x-y coordinate.

852 **Table 2. Neuronal changes with on-going reward consumption**
853

Choice over zero-bundle						
Task epoch	Neurons tested	Neurons	Responses	Neurons	Responses	Neurons
		Response decreases		Response increases		No effects
Positive coding						
Bundle stimulus			30		21	
Go			28		17	
Choice			25		15	
Reward			18		16	
Subtotal	55	31	101	21	69	3
Negative coding						
Bundle stimulus			10		15	
Go			8		15	
Choice			8		11	
Reward			7		13	
Subtotal	31	14	33	15	54	2
Choice between two non-zero bundles						
Task epoch	Neurons tested	Neurons	Responses	Neurons	Responses	Neurons
		Response decreases		Response increases		No effects
Positive coding						
Bundle stimulus			16		16	
Go			15		15	
Choice			13		16	
Reward			10		10	
Subtotal	38	16	54	16	57	6
Negative coding						
Bundle stimulus			11		9	
Go			9		8	
Choice			8		6	
Reward			8		8	
Subtotal	32	14	36	9	31	9

854
855 This table includes data from chosen value responses, separated according to the four task epochs
856 (Bundle stimulus, Go, Choice and Reward) and all bundles tested for satiety (component A:
857 blackcurrant juice, component B: grape juice, water or mango juice). Positive coding refers to
858 response increase with higher value before satiety, whereas negative coding refers to response
859 decrease with higher value. Most neurons were tested both in choice over zero-bundle and in choice
860 between two non-zero bundles.

Suppplementary information for

Orbitofrontal cortex neurons code utility changes during natural reward consumption as correlates of relative reward-specific satiety

Alexandre Pastor-Bernier, Arkadiusz Stasiak and Wolfram Schultz*

Corresponding author: Wolfram Schultz
Email: Wolfram.Schultz@Protonmail.com

This Supplmentary Information includes:
Supplementary Methods
Supplementary Figures S1 to S3 and their legends

SI METHODS

Animals

Two adult male macaque monkeys (*Macaca mulatta*; Monkey A, Monkey B), weighing 11.0 kg and 10.0 kg, respectively, were used in these experiments that had already yielded behavioral and neuronal data without satiety (Pastor-Bernier et al., 2017; Pastor-Bernier et al., 2019). Neither animal had been used in any other study.

Ethical approval

This research has been ethically reviewed, approved, regulated and supervised by the following institutions and individuals in the UK and at the University of Cambridge (UCam): the Minister of State at the UK Home Office, the Animals in Science Regulation Unit (ASRU) of the UK Home Office implementing the Animals (Scientific Procedures) Act 1986 with Amendment Regulations 2012, the UK Animals in Science Committee (ASC), the local UK Home Office Inspector, the UK National Centre for Replacement, Refinement and Reduction of Animal Experiments (NC3Rs), the UCam Animal Welfare and Ethical Review Body (AWERB), the UCam Governance and Strategy Committee, the Home Office Establishment License Holder of the UCam Biomedical Service (UBS), the UBS Director for Governance and Welfare, the UBS Named Information and Compliance Support Officer, the UBS Named Veterinary Surgeon (NVS), and the UBS Named Animal Care and Welfare Officer (NACWO).

General behavior

The animals were habituated during several months to sit in a primate chair (Crist Instruments) for a few hours each working day. They were trained in a specific, computer-controlled behavioral task in which they contacted visual stimuli on a horizontally mounted touch-sensitive computer monitor (Elo) located 30 cm in front of them. The animal's eye position in the horizontal and vertical plane were monitored with a non-invasive infrared oculometer (Iscan). Matlab software (Mathworks) running on a Microsoft Windows XP computer controlled the behavior and collected, analyzed and presented the data on-line. A solenoid valve (ASCO, SCB262C068) controlled by the same Windows computer served to deliver specific liquid amounts. A Microsoft SQL Server 2008 Database served for Matlab off-line data analysis. Following task training for about 6 months, animals were surgically implanted with a recording chamber for electrophysiological recordings, which typically lasted for another 6-10 months.

Stimuli, task and rewards

A computer touch monitor presented the subject with two visual stimuli (4° apart) representing two bundles, a Reference Bundle and a Variable Bundle (Figure 1A). Each bundle contained two

rewards (Component Reward A: violet rectangle, and component Reward B: green rectangle) with independently set amounts indicated by the vertical bar position within each rectangle (higher was more). The Reference Bundle contained two preset Reward amounts that were fixed for a given block of trials. The Variable Bundle contained a specifically set amount of one Reward and an experimentally varied amount of the other reward. The task sequence (Figure 1B) has been described in detail (Pastor-Bernier et al., 2017; Pastor-Bernier et al., 2019) and are summarized as follows. Reward A in all bundles was blackcurrant juice, or blackcurrant juice with added monosodium glutamate (MSG), Reward B was grape juice, strawberry juice, mango juice, water, apple juice, peach juice, or grape juice with added inosine monophosphate (IMG).

Each trial began when the animal contacted a centrally located touch sensitive key for 1.0 s after a pseudorandom inter-trial interval of 1.6 ± 0.25 s. Then two bundles appeared and remained on the screen for 2.0 s, after which two blue spots appeared as GO stimulus underneath the bundles, upon which the animal released the touch key and touched the blue spot of its choice within 2.0 s. After a hold time of 1.0 s, the chosen blue spot turned green and the unchosen blue spot disappeared. Simultaneously a white frame around the chosen bundle appeared providing feedback for successful choice. The computer-controlled liquid solenoid valve delivered liquid A at 1.0 s after the choice, followed 0.5 s later by liquid B (except when using peach juice as Reward B; here the sequence was reversed: liquid B was delivered first, then 0.5 s later liquid A, blackcurrant juice).

Estimation of behavioral ICs

The behavioral method used to obtain an IP from stochastic choice has been presented in full detail (Pastor-Bernier et al., 2017; Pastor-Bernier et al., 2019). With two bundle options, the animal chose between the pre-set Reference Bundle (left in Figure 1A) and the Variable Bundle (right) in repeated trials. Thus, the constant Reference Bundle provided a stable reference against the changing bundle composition in the Variable Bundle. We set one reward in the Variable Bundle to one unit ($\geq$ 0.1 ml) above the amount of the same reward in the Reference Bundle, while pseudorandomly varying the amount of the other reward widely. The variation of the animal's repeated choice with that single varying Reward allowed us to construct a full psychophysical function and estimate the IP from a Weibull fit (point of subjective equivalence; $P = 0.5$ choice of each bundle).

As in our previous study (Pastor-Bernier et al. 2017), we used the Matlab function GLMFIT for psychophysical fitting. This function returns a number called 'Deviance' between 0 to infinity that can be used to compare fitting between Weibull and logit. The Deviance is the difference between the log-likelihood of the fitted model and the maximum possible log-likelihood. Lower values are better. The estimated Deviance for psychophysics for the first 5,000 trials and 2 monkeys was 1.0415 for the Weibull model and 1.6009 for the logit model, suggesting that the Weibull fitted better the data. Hence, we used Weibull fitting for all psychophysical fitting.

We obtained each IP from a total of 80 trials (2 left-right stimulus positions with 5 equally spaced Reward amounts in 8 trials). To avoid known adaptations in OFC neurons (Tremblay and Schultz, 1999; Padoa-Schioppa, 2009; Kobayashi et al., 2010; Rustichini et al., 2017), we always tested the full reward range of the experiment.

To obtain an IC, we fit a series of IPs with a hyperbolic function d using weighted least mean squares:

$$d = ay + bx + cxy \qquad \text{(Eq. 1)}$$

with y and x as milliliter amount of Reward A (plotted at y-axis on 2D graph, Figure 1C and 1E) and Reward B (plotted at x-axis), a and b as weights of the influence of the Reward amounts plotted on the y- and x-axes, respectively, and c as curvature. A potent reward that contributes strongly to the choice of the bundle would have a large weight (high coefficient a or b), whereas a less potent reward would have lower weight coefficients. Thus, with the potent (more weight) reward plotted

on the x-axis, and the less potent (less weight) reward plotted on the y-axis, choice indifference between them (IC) would occur with smaller milliliter amounts on the x-axis compared to the y-axis. Hence, the IC slope would be steeper than the diagonal line (see Figure 1C, D). By resolving Eq. 1 as y = -(b / a) * x, the IC slope would be the ratio of the coefficients that reflect the weights of the rewards: -b / a. With a higher potency of Reward B (x-axis) compared to Reward B (y-axis), the rectified IC slope would be larger than 1. Relatively stronger satiety for Reward B (x-axis) compared to Reward A (y-axis) would reduce the weight of Reward B, reduce the absolute value of the ratio -b / a, and flatten the IC slope. Thus, the IC slope -b / a describes the relative impact of the two bundle rewards (reflecting the value ratio between the two rewards), whereas the weights (a and b) describe the influence of the Reward amounts.

The hyperbolic function can be re-written in an equivalent form to the regression with interaction used for analysing neuronal responses (see Eq. 3 below):

$$y = \beta_0 + \beta_1 A + \beta_2 B + \beta_3 AB + \varepsilon \qquad \text{(Eq. 1a)}$$

with A and B as milliliter amount of Reward A (plotted at y-axis) and Reward B (x-axis), respectively, β_0 as offset coefficient, β_1 and β_2 as behavioral regression coefficients, and ε as compound of errors err_0, err_1, err_2, err_3 for offset and regressors 1-3.

Definition and criteria for pre-sated and sated states

With on-going reward consumption, the changes of psychophysical choice functions exceeding the confidence intervals (CI) of initial tests suggested a changed value relationship between the two bundle rewards suggestive of relative, reward-specific satiety (see Figures 1D, S1A, S1E). More specifically, the gradual effect of satiety on choice preference was identified by tracking the IPs as consumption advanced across blocks of 80 trials. Importantly, these changes occurred fast enough to be studied during the recording durations of single neurons, thus allowing us to compare responses between non-sated and sated states in the same neuron. The Weibull-fitted IPs were obtained psychophysically for fixed and equally spaced amounts of Reward B. Changes in relative value of the two bundle rewards were assessed with interleaved anchor trials in choices between bundles with only one non-zero reward: bundle (fixed non-zero blackcurrant juice; no Reward B) vs. bundle (no blackcurrant juice; variable non-zero Reward B), using any Reward B (Figure S1B). To aggregate IP data across sessions and compensate for across-session variability, we normalized the reward value ratio to the first titration block in all sessions. We then compared the normalized distributions of IPs within the CI of the first block with the distributions of IPs exceeding the CI of the first block.

Control regressions for behavioral choice

To test whether the animal's choice reflected the amount of the bundle rewards during satiety, rather than other, unintended variables such as spatial bias, we used the logistic regression

$$P\,(V) = \beta_0 + \beta_1 CT + \beta_2 RA + \beta_3 RB + \beta_4 VA + \beta_5 VB + \beta_6 CL + \beta_7 MA + \beta_8 MB + \varepsilon \qquad \text{(Eq. 2)}$$

with P (V) as probability of choice of Variable Bundle, β_0 as offset coefficient, β_1 - β_7 as correlation strength (regression slope) coefficients indicating the influence of the respective regressor, CT as trial number within block of consecutive trials, RA as amount of Reward A of Reference Bundle, RB as amount of Reward B of Reference Bundle, VA as amount of Reward A of Variable Bundle, VB as amount of Reward B of Variable Bundle, CL as choice of any bundle stimulus presented at the left, MA as consumed amount of Reward A, MB as consumed amount of Reward B, and ε as compound error for offset and all regressors. We used a binomial fit with logit link function to obtain standardized β coefficients. Choices over zero-reward bundles were excluded in the regression to avoid internal correlation between value and consumption.

Licking

Licking was monitored with an infrared optosensor positioned below the juice spout (V6AP; STM Sensors). Anticipatory licking durations were measured between the appearance of the bundle stimuli and delivery of the first reward liquid (approximate duration 5 - 6 s) in bundles containing only one non-zero component reward with advancing trials in satiety and within single working sessions. Licking data were collected with four different bundles, namely (blackcurrant juice, grape juice), (blackcurrant juice, water), (blackcurrant juice, strawberry juice) and (blackcurrant juice, mango juice).

Surgical procedures and electrophysiology

As described before for the same animals (Pastor-Bernier et al., 2019), a head-restraining device and a recording chamber (40 x 40 mm, Gray Matter) were implanted on the skull under full general anesthesia and aseptic conditions. The stereotactic coordinates of the chamber enabled neuronal recordings of the orbitofrontal cortex (OFC) (Paxinos et al., 2000). We located the OFC from bone marks on coronal and sagittal radiographs taken with a guide cannula inserted at a known coordinate in reference to the implanted chamber, using a medio-lateral vertical and a 20º degree forward directed approach aiming for area 13. Monkey A provided data from the left hemisphere, Monkey B from the right hemisphere, via a craniotomy in each animal ranging from Anterior 30 to 38, and Lateral 0 to 19. We conducted single-neuron electrophysiological recordings using both custom made glass-coated tungsten electrodes (Merrill & Ainsworth, 1972), and commercial electrodes (Alpha Omega, Israel) (impedance of about 1 MOhm at 1 kHz). Electrodes were inserted into the cortex with a multi-electrode drive (NaN drive, Israel) with the same angled approach as used for the radiography. Neuronal signals were collected at 20 kHz, amplified using conventional differential amplifiers (CED 1902 Cambridge Electronics Design) and band-passed filtered (high: 300 Hz, low: 5 kHz). We used a Schmitt-trigger to digitize the analog neuronal signal online into a computer-compatible TTL signal. However, we did not use the Schmitt-trigger to separate simultaneous recordings from multiple neurons, in which case we searched for another recording from only a single neuron, or we stored occasionally the data in analog form for off-line separation by dedicated software (Plexon offline sorter). An infrared eye tracking system monitored eye position (ETL200; ISCAN), with temperature check on an experimenter's hand at the approximate position of the animal's head.

Definition for neurons following the revealed preference scheme

We analysed single-neuron activity during four task epochs vs. Pretrial control (1 s): visual Bundle stimulus (2 s), Go signal (1 s), Choice (1 s) and Reward (2 s, starting with Reward A, followed 0.5 s later by Reward B, thus covering both rewards). To establish neuronal relationships to these task epochs, we compared the activity in each neuron during each task epoch separately against the Pretrial control epoch using the paired Wilcoxon test ($P < 0.01$). A neuron was considered task-related if its activity in at least one of the four task epochs differed significantly from the activity during the Pretrial control epoch.

Responses of individual neurons should follow the scheme of two-dimensional ICs that characterizes revealed behavioral preferences for two-dimensional bundles. Specifically, the responses should comply with three characteristics defined previously (Pastor-Bernier et al., 2019).

(Characteristic 1) Neuronal responses should change monotonically with increasing behavioral preference *across behavioral ICs*, irrespective to bundle composition. Such monotonic neuronal response changes should reflect increasing amounts of one or both bundle rewards, assuming a positive monotonic subjective value function on Reward amount.

(Characteristic 2) Neuronal responses should vary insignificantly for all equally preferred bundles positioned *along a same behavioral IC*, despite different physical bundle composition.

(Characteristic 3) Neuronal responses should follow the IC slope and the non-linear curvature of behavioral ICs. The IC slope reflects the value relationship between the two bundle rewards,

indicating the revealed preference relation between the two rewards of a bundle, and thus the value of one reward relative to a common reference reward.

We used a combination of three statistical tests to assess these characteristics.

Characteristic 1: To capture the change *across ICs* in the most conservative, assumption-free manner possible, we used a simple linear regression on each Wilcoxon-identified task-related response:

$$y = \beta_0 + \beta_1 A + \beta_2 B + \beta_3 AB + \varepsilon \qquad \text{(Eq. 3)}$$

with y as neuronal response in any of the four task epochs, measured as impulses/s and z-scored normalized to the Pretrial control epoch of 1.0 s (z-scoring of neuronal responses applied to all regressions listed below), A and B as milliliter amount of Reward A (plotted at y-axis) and Reward B (x-axis), respectively, β_0 as offset coefficient, β_1 and β_2 as neuronal regression coefficients, and ε as compound error for offset and all regressors.

The coefficients β_1 and β_2 needed to be either both positive (indicating positive neuronal relationship, higher neuronal activity reflecting more reward quantity) or both negative (inverse neuronal relationship) to reflect the additive nature of the individual bundle components giving rise to revealed preference ($P < 0.05$, unless otherwise stated; t-test).

This linear regression assessed the degree of linear monotonicity of neuronal response change across ICs ($P < 0.05$ for β coefficients; t-test). Further, all significant positive or negative response changes identified by Eq. 3 needed to be also significant in a Spearman rank-correlation test that assessed ordinal monotonicity of response change across ICs without assuming linearity and numeric scale ($P < 0.05$).

Characteristics 1 and 2: To assess the two-dimensional *across/along IC* scheme in a direct and intuitive way, and without assuming monotonicity, linearity and numeric scale, we used a two-factor Anova on each Wilcoxon-identified task-related response that was significant for both regressors in Eq. 3; the factors were *across-IC* (ascending rank order of behavioral ICs) and *along-IC* (same rank order of behavioral IC). To be a candidate for following the IC scheme of Revealed Preference Theory, changes across-ICs should be significant ($P < 0.05$), changes within-IC should be insignificant, and their interaction should be insignificant.

Characteristic 3: Whereas the regression defined by Eq. 3 estimated neuronal responses across ICs, a full estimation of neuronal ICs for comparison with behavioral ICs would require inclusion of the IC slope and curvature, both of which depended on both rewards. By simplifying Eq. 3 by setting to zero both the β_3 coefficient and the constant neuronal response along the IC, the neuronal IC slope would be the ratio of coefficients ($-\beta_2 / \beta_1$). Note the different meanings of the slope term: the neuronal IC slope ($-\beta_2 / \beta_1$) describes the relative coding strength of the two bundle rewards (reflecting the neuronal ratio of the two rewards), whereas each neuronal regression slope alone (β) describes the coding strength of neuronal response (correlation with the specific regressor). The neuronal IC curvature was estimated from the β_3 coefficient of the interaction term AB (all β's $P < 0.05$; t-test).

Neuronal chosen value coding. As stated before (Pastor-Bernier et al. 2019), chosen value (CV) was defined as the value of a choice option the animal considered, would obtain or had obtained by its choice. As each option consisted of two components, we used a linear combination of the quantity of the two component rewards A (blackcurrant juice) and B (any of the other five rewards):

$$CV = A + k_1 B \qquad \text{(Eq. 4)}$$

Weighting parameter k_1 served to adjust for differences in subjective value between rewards A and B, such that the quantity of Reward B entered the regression on a common-currency scale defined by Reward A. We established parameter k_1 during neuronal recording sessions from behavioral choice IPs using quantitative psychophysics in anchor trials (80 trials per test, see above Trial types

for neuronal tests), rather than reading it from fitted ICs. Thus, k_1 equals the ratio of coefficients β_2 / β_1 of Eq. 3.

We established a common-currency scale in ml for all tested rewards by defining blackcurrant juice or blackcurrant-MSG (Reward A) as reference (numeraire); the subjective value of any reward is expressed as real-number multiple k_1 of the quantity of the numeraire at choice indifference. Specifically, the animal chose between the Variable Bundle that contained a psychophysically varied quantity of blackcurrant juice (the other bundle reward being set to 0 ml) and the Reference Bundle that contained a fixed quantity of the other reward (blackcurrant juice being set to 0 ml). At choice indifference, the quantity of blackcurrant juice (Reward A) in the Variable Bundle defined the common-currency value of the other reward, from which we calculated parameter k_1 as A / B. A k_1 of < 1 indicated that more quantity was required for choice indifference against blackcurrant juice; thus, $k_1 < 1$ suggested that the tested reward had lower subjective value than blackcurrant juice. By contrast, $k_1 > 1$ suggested higher subjective value, as less quantity was required for choice indifference.

We assessed the coding of chosen value and unchosen value in all neurons that followed the revealed preference scheme, using the following regression:

$$y = \beta_0 + \beta_1 CV + \beta_2 UCV + \varepsilon \qquad \text{(Eq. 5)}$$

with UCV as value of the unchosen option that was not further considered here, and ε as compound error for offset and all regressors.

Vector plots of OFC reward sensitivity. The purpose of this analysis was to provide quantitative and graphic information about satiety-induced behavioral and neuronal changes that would allow comparison with previous OFC studies that had not used two-component choice options with individually varying reward amounts and therefore did not establish ICs (Padoa-Schioppa & Assad 2006). This simplified analysis addressed monotonic response increase or decrease with increasing amounts of bundle rewards across ICs (characteristic 1 above), but did not address other IC characteristics such as trade-off, slope and curvature (characteristics 2 and 3) that had not been investigated previously. We established 2D plots whose dots indicated the relative contribution of each of the two bundle rewards to the neuronal response. We then compared vectors of behavioral choices with vectors of averaged neuronal population responses before and during satiety.

For behavioral choices, we plotted vectors (with 95% CIs) from averaged dot positions defined by reward amount (distance from center: sqrt $(\beta_1^2 + \beta_2^2)$) and relative weight (elevation angle: arctangent (β_1 / β_2)); coefficient β_1 refers to Reward A (blackcurrant, y-axis), coefficient β_2 refers to any of the other rewards (x-axis) (Eq. 1a). The angle of the vector reflects the relative contribution the two bundle rewards to the choice, as estimated by the a and b coefficients (Eq. 1). A deviation of the alignment angle from the diagonal line indicates an unequal contribution weight to bundle choice, and thus a non-1:1 reward ratio.

For neuronal responses, each dot on the 2D plot was defined by the two β regression coefficients for neuronal responses (Eq. 3; $P < 0.01$, t-test) for each of the two rewards in any of the four task epochs. The distance from center indicates the z-scored response magnitude (sqrt $(\beta_1^2 + \beta_2^2)$), coding sign (positive or negative), and relative weight (elevation angle; arctangent (β_1 / β_2)) of the two β coefficients. Coefficient β_1 refers to Reward A (blackcurrant, y-axis), coefficient β_2 refers to any of the other rewards (x-axis). Responses with negative (inverse) coding were rectified. Further IC characteristics such as systematic trade-off across multiple IPs and IC curvature played no role in these graphs. The alignment of the dots along the diagonal axis shows the relative coding strength for the two bundle rewards, as estimated by the β regression coefficients; a deviation from the diagonal line indicates an unequal influence of the two bundle rewards on the neuronal responses, reflecting a neuronal correlate of reward ratio.

Neuronal decoders

We used linear support vector machine (SVM) algorithms to decode neuronal activity according to bundles presented at different behavioral ICs during choice over zero-reward bundle (bundle distinction) and, separately, according to the behavioral choice between two non-zero bundles located on different ICs (choice prediction). As in our main study on revealed preferences (Pastor-Bernier et al., 2019), we implemented both decoders as custom-written software in Matlab R2015b (Mathworks). The SVM decoder with linear kernel was accomplished with svmtrain and svmclassify procedures (our previous work had shown that use of nonlinear SVM kernels does not improve decoding Tsutsui et al., 2016). The SVM decoder was trained to find the optimal linear hyperplane for the best separation between two neuronal populations relative to lower vs. higher ICs.

All analyses employed single-neuron data, consisting of single-trial impulse counts that had been z-normalised to the activity during the Pretrial epoch in all trials recorded with the neuron under study. The analysis included activity from all neurons whose responses followed the IC scheme of revealed preferences during any of the four task epochs, as identified by our three-test statistics, except where noted. The neurons were recorded one at a time; therefore, the analysis concerned aggregated pseudo-populations of neuronal responses.

The decoding analysis used 10 trials per neuron for each of two ICs (total of 20 trials). Extensive analysis suggested that higher inclusion of 15-20 trials per group did not provide significantly better decoding rates (while reducing the number of included neurons). For neurons that had been recorded with > 10 trials per IC, we selected randomly 10 trials from each neuron for each of the two ICs. We used a leave-one-out cross-validation method in which we removed one of the 20 trials and trained the SVM decoder on the remaining 19 trials. We then used the SVM decoder to assess whether it accurately detected the IC of the left-out trial. We repeated this procedure 20 times, every time leaving out another one of the 20 trials. These 20 repetitions resulted in a percentage of accurate decoding (% out of $n = 20$). The final percentage estimate of accurate decoding resulted from averaging the results from 150 iterations of this 20-trial random selection procedure. To distinguish from chance decoding, we randomly shuffled the assignment of neuronal responses to the tested ICs, which should result in chance decoding (accuracy of 50% correct). A significant decoding with the real, non-shuffled data would be expressed as statistically significant difference against the shuffled data ($P < 0.01$; Wilcoxon rank-sum test).

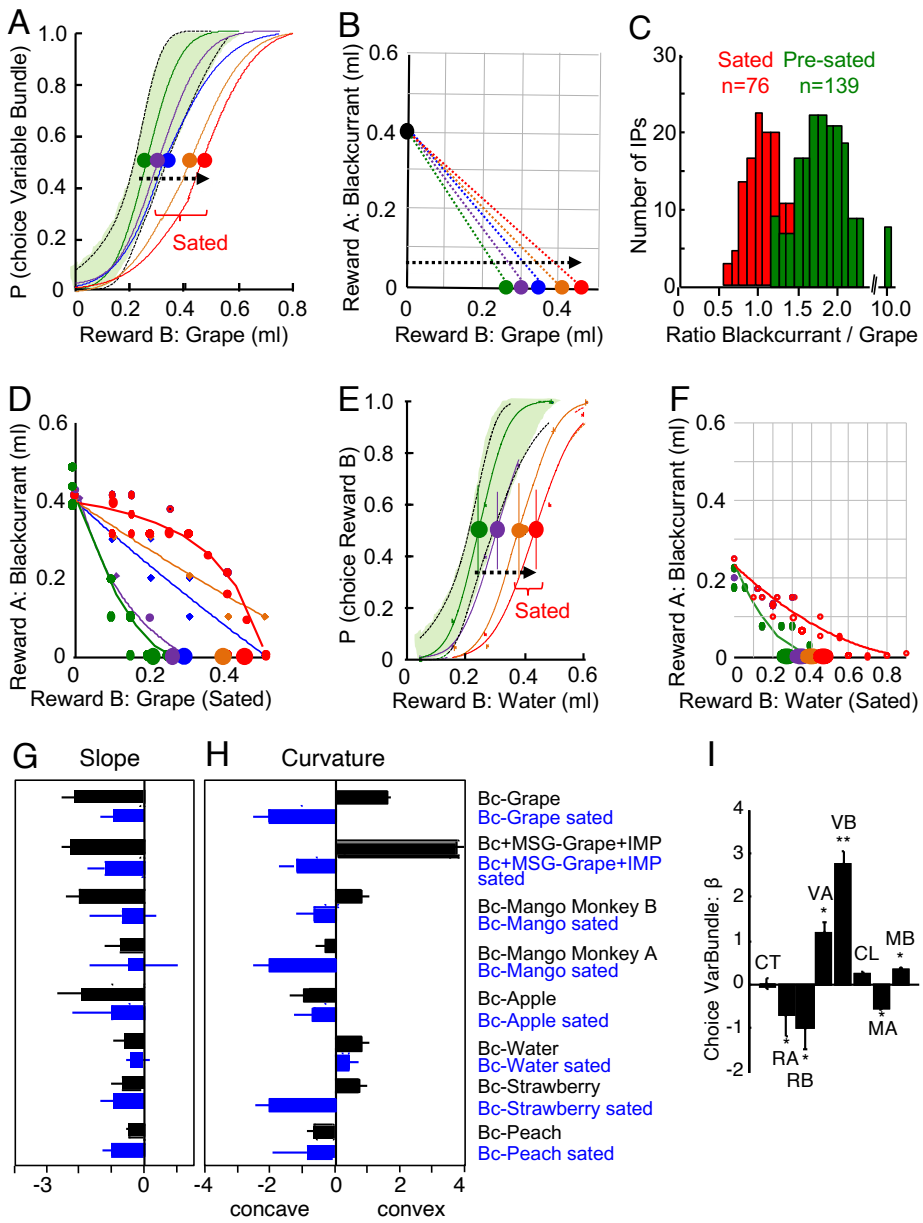

Figure S1. Additional behavioral measures

(A) Psychophysical assessment of choice between single-reward bundles with grape juice variation (constant Reference Bundle: 0.6 ml blackcurrant juice, 0.0 ml grape juice; Variable bundle: 0.0 ml blackcurrant juice, varying grape juice). Green and violet curves inside green ±95% confidence intervals: initial choices; blue, orange and red curves: on-going consumption. The decrease in ratio blackcurrant/grape juice amounts at IP was significant between the confidence interval of the first IP and all IPs exceeding it (ratios of 1.9857 ± 0.0173, N = 139, green, vs. 1.0077 ± 0.02, orange and red; mean ± standard error of the mean, SEM; individual trial blocks: $p = 9.6943 \times 10^{7}$, Kolmogorov-Smirnov test; $p = 2.336 \times 10^{-32}$, Wilcoxon rank-sum test; $p = 3.1712 \times 10^{-46}$, t-test; Monkey A). Each curve and indifference point (IP) were estimated from 80 trials in a single block (Weibull fits).

(B) Gradually developing relative satiety for grape juice indicated by increasing choice indifference points (IP; same bundles and animal as in A): with on-going consumption of both juices, the animal gave up progressively more grape juice for obtaining the same 0.4 ml of blackcurrant juice (from green to red). The ratio blackcurrant/grape juice amounts at IP decreased from approximately 2:1 (0.4 ml of blackcurrant juice for 0.25 ml of grape juice, black vs. green dots) to about 1:1 (0.4 ml

blackcurrant for 0.45 ml grape juice, black vs. red), suggesting subjective value loss of grape juice relative to blackcurrant juice.
(C) Significant decrease of ratio blackcurrant/grape juice amounts at IP with on-going consumption (same bundles as in A; Wilcoxon test). N = 139 and 76 IPs estimated in 43 trial blocks (Monkey A).
(D) Gradual changes with grape juice variation in slope and curvature of choice indifference curves (IC) between pre-satiety (green, violet) and during increasing satiety (blue, orange, red) (single session; 2,960 trials; 80 trials/IP; Monkey A).
(E), (F) Psychophysical tests and consumption-dependent change of ICs in Monkey B during choice between single-component bundles (constant Reference Bundle: 0.25 ml blackcurrant juice, 0.0 ml water; Variable bundle: 0.0 ml blackcurrant juice, varying water). With on-going consumption of both liquids, the animal gave up progressively more water for obtaining the same 0.25 ml of blackcurrant juice (from green to red), suggesting subjective value loss of water relative to blackcurrant juice. Same conventions as in A and D (2,400 trials; 80 trials/IP), Monkey B.
(G), (H) Significant IC slope and curvature changes from pre-sated to sated states with on-going consumption with individual bundles (Bc, blackcurrant juice; MSG, monosodium glutamate; IMP, inosine monophosphate; $p = 0.0156$ and $p = 0.0313$, respectively; Wilcoxon test). The slope parameter reflects the amount ratio blackcurrant/other liquids at IP.
(I) Value control by logistic regression for choice of Variable Bundle over non-zero Reference Bundle during satiety (Eq. 2). According to significance of β regression coefficients, choice of the Variable Bundle (Choice VarBundle) correlated significantly with amount of rewards A and B in the Variable Bundle (VA, VB) and the Reference Bundle (RA, RB) and the consumed amount of bundle rewards A (blackcurrant; MA) and B (various other liquids; MA). Choice varied insignificantly with consecutive trial number within blocks (CT) and left-right choice (CL). N = 7,243 trials pooled from several sessions; * $P < 0.05$; ** $P < 0.01$; t-test on βs.

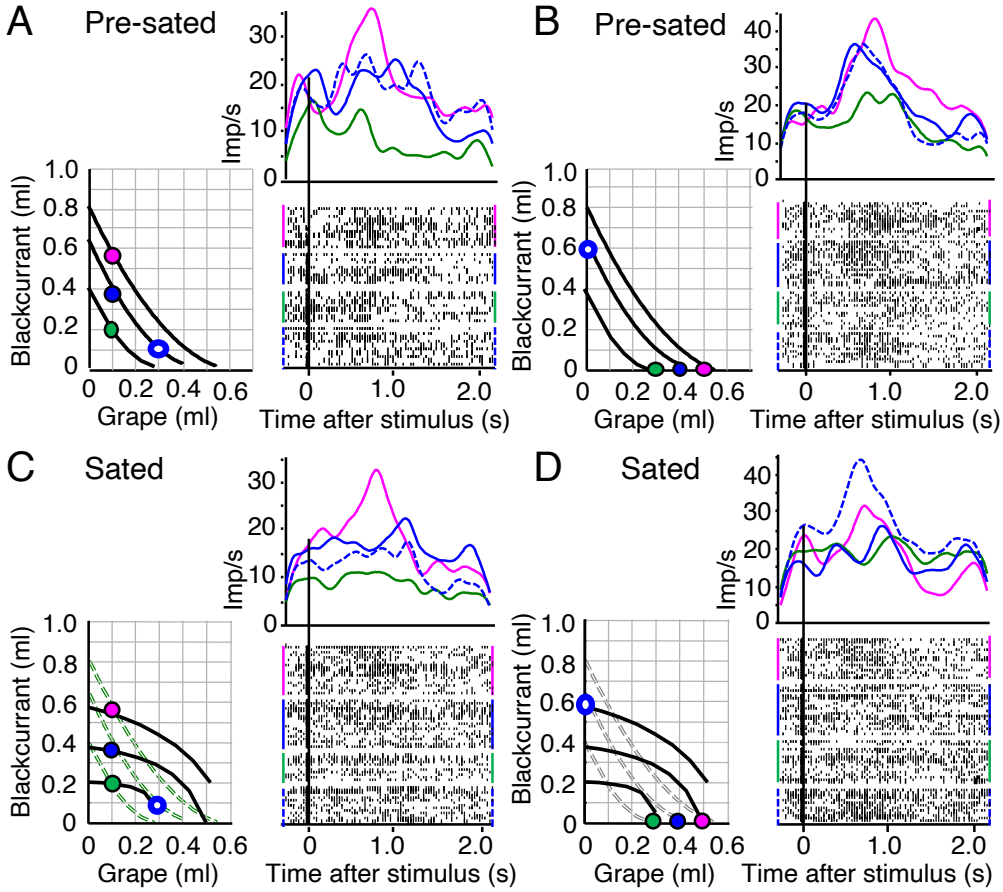

Figure S2. Satiety-related neuronal response change during choice between two non-zero bundles

(A) Significant monotonic neuronal response increase with value of chosen bundle across indifference curves (IC) before satiety (from green via blue to red) (P = 0.0055, F = 10.49, 17 trials; 1-way Anova). The animal chose between the Reference Bundle (hollow blue dot) and one of the Variable Bundles (solid colored dots). The responses to the two blue bundles on the same IC (indicating equal preference) varied insignificantly despite different juice composition (P = 0.5488, F = 0.38, 18 trials). Response to Reference Bundle (hollow blue dot) is indicated by dotted line.
(B) As (A) but for grape juice variation. Responses varied significantly across ICs with grape juice (P = 0.0046, F = 9.7, 27 trials). The responses to the two blue bundles on the same IC differed insignificantly (P = 0.2622, F = 1.31, 29 trials). Same color labels as in (A).
(C) Despite IC change indicating satiety, the neuronal response increase across ICs remained significant (P = 0.0014, F = 10.87, 17 trials). However, the two unchanged blue bundles were now on different ICs, and their responses varied significantly (P = 0.0028, F = 5.46, 40 trials).
(D) With slope and curvature change indicating satiety, the three bundles with grape juice variation were now located within only two ICs. Although the neuronal response increase across ICs remained significant (P = 0.0144, F = 6.02, 35 trials), the peak response was reduced by 25% (from 40 to 30 imp/s, red) and the three responses were closer to each other. Further, the two unchanged blue bundles were now on different ICs, and their responses now differed significantly (P = 0.0201, F = 9.27, 52 trials). Thus, the changes of neuronal responses were consistent with the IC change indicating satiety.

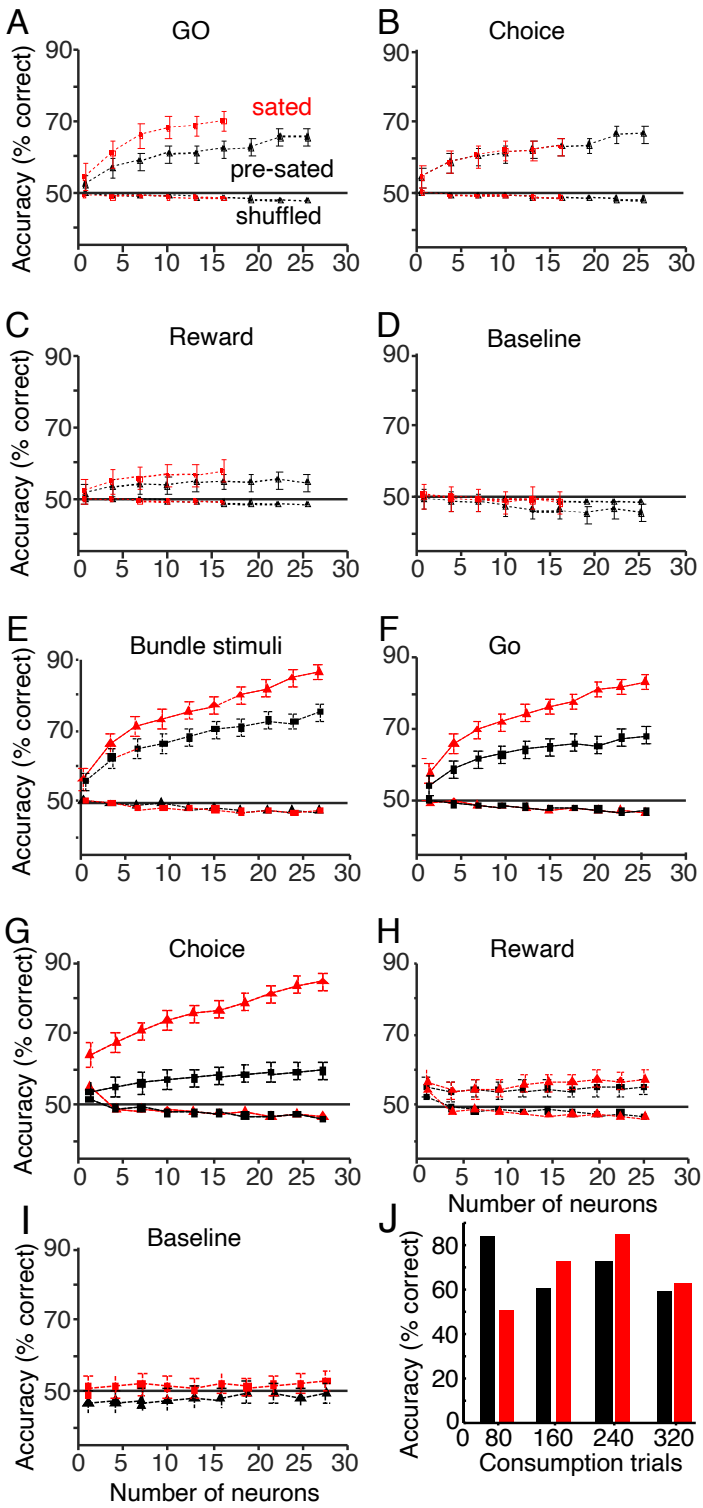

Figure S3. Satiety-induced changes in bundle classification during different task epochs
(A) - (D) Choice over zero-bundle. Baseline refers to 1 s Pretrial control epoch before Bundle stimuli. For details, see Figure 6.
(E) - (I) As (A-D) but for choice prediction by neuronal responses during choice over non-zero bundle.
(J) Classification accuracy of neuronal responses across on-going liquid consumption. Same data selection as for (A-D) and collapsed across all task epochs. Black: before satiety, red: during satiety.