WORKING PAPER

Uploaded to bioRxiv

Reliable population utility signal from diverse economic value coding in orbitofrontal neurons

Simone Ferrari-Toniolo* and Wolfram Schultz

Department of Physiology, Development and Neuroscience

University of Cambridge, Cambridge, UK

*** Corresponding author**

Email addresses of all authors:
Simone Ferrari-Toniolo: Simone.FerrariToniolo@gmail.com
Wolfram Schultz: Wolfram.Schultz@protonmail.com

Abbreviated title: Neuronal economic utility and probability weighting

Conflict of interest: The authors declare no conflicts of interest.

Acknowledgements: We thank Aled David and Christina Thompson for animal and technical support, Polly Taylor for expert anesthesia, and Federico M. Echenique and David M. Grether (Caltech) for helpful comments. This work was supported by Wellcome Grants WT 095495 and WT 204811 and European Research Council Advanced Grant 293549

Abstract

Economic value encapsulates the subjective combination of reward magnitude and probability. We investigated the mechanism for subjective value computation in single neurons using an economic axiomatic approach. We found that single neurons in the macaque orbitofrontal cortex, known to be sensitive to reward magnitude and probability, encode the economic value functions (utility and probability weighting) in a heterogeneous manner, such that the activity of individual neurons did not match the animal's choices. However, the utility and probability weighting code from a population of these varied neurons reliably matched the animals' choices and risk attitudes. Thus, the neuronal population code for economic value amounted to a distributional representation of the formal economic functions. With a diverse single-unit economic value code converging into a reliable population-level utility code, this scheme suggests a brain mechanism for the flexible accommodation of multiple choice patterns and risk attitudes.

Main text

Utility is a fundamental economic concept which defines subjective values. Expected Utility Theory (EUT) describes choices as a utility-maximization process: we assign a numerical utility to each option and choose the option with the highest utility (von Neumann and Morgenstern, 1944). The abstract utility concept, which can only be inferred from observed choices, is not necessary for explaining behavior (we make choices "as if" maximizing utility) but represents a testable brain mechanism for choice (Camerer, 2008; O'Doherty and Camerer, 2015).

Neurons sensitive to reward magnitude and probability have been identified in midbrain dopamine neurons and in the orbitofrontal cortex (OFC) (Tremblay and Schultz, 1999; Fiorillo, C. D., Tobler, P. N., Schultz, 2003; Padoa-Schioppa and Assad, 2006; Padoa-Schioppa and Cai, 2011; Lak et al., 2014; Stauffer et al., 2014). Dopamine neurons show nonlinear responses to reward magnitudes, compatible with utility coding (Stauffer et al., 2014). These data suggest that single neurons drive a reward-maximization process by encoding the subjective values associated with the different choice options, although direct evidence for such a mechanism is missing.

We investigated whether single neurons and neuronal populations in the macaque OFC encode a utility signal compatible with the axiomatic economic theory, implying a specific neuronal mechanism for value computation. We identified a quantitative utility measure based on the continuity axiom of EUT, a test for robustness of magnitude-probability integration into a scalar utility value, across multiple settings. This showed that utility maximization reliably described risky choices in macaques (Ferrari-Toniolo et al., 2021). Using the same paradigm, we recorded the activity of single OFC neurons, to infer a corresponding neuronal utility measure. Both behavioral and neural procedures considered the non-linear contribution of magnitudes and probabilities to the formation of utility values, in line with modern economic theories (e.g. Prospect Theory).

We found neurons that were nonlinearly sensitive to reward magnitude and probability. Utility was heterogeneously encoded, such that only a few individual neurons reflected the animal's choices. However, the population utility code computed from these heterogeneous neurons reliably matched the animals' choices and risk attitudes.

Task design and behavioral results. Monkeys were presented with visual cues indicating different combinations of reward magnitude and probability (Fig. 1a) making binary decisions through joystick movements (Fig. 1b). Options with different reward magnitude (m, 0 to 0.5 ml) and probability (p, 0.1 to 0.9) were pseudorandomly intermixed (Fig. 1a). Based on the continuity axiom of EUT, we defined a behavioral utility measure as the reward probability for which the monkey was indifferent between a gamble (AC) and a safe option (B) (Fig. 1c). The AC gamble was defined as a combination of two safe options (A, C) with probabilities p and $(1 - p)$, respectively. According to the continuity axiom, the resulting behavioral indifference point (IPb) represents a utility measure, being a numerical quantity associated with the subjective evaluation of B in relation to A and C. Monkeys showed consistent choice behavior, allowing for the definition of the IPb, in a single option-set {A,B,C} (Fig. 1d) as well as in a range of different option-sets with varying A and B values (A: 0.20–0.50 ml, B: 0.15–0.45 ml, C: 0 ml). We represented the results of these multiple-option-sets measurements in a two-dimensional magnitude-probability space, where each point corresponded to a gamble (Fig. 1e). The full set of IPbs identified a series of indifference curves (ICs), points with equal utility in the magnitude-probability space (Fig. 1f). The ICs revealed a choice behavior compatible with a utility maximization process: by fitting an economic model in which each gamble's value was computed as the subjective probability-weighted expected utility, we recovered the behavioral ICs with good approximation (Fig. 1f, colored lines) (see Ferrari-Toniolo et al., 2021 for more details). These modeled values originated from the combination of a utility function (u) and a probability weighting function (w), describing the nonlinear subjective weighting of reward magnitude (m) and probability (p), respectively.

Utility measure in single neurons. We turned to the neural foundation of the observed utility maximization process. Via the same axiomatic principle used to characterize behavior, we defined a

neuronal utility measure as the indifference point computed from the neural activity. A neuron coding utility would respond equally to the two choice options at the IPb. The same neuron, when presented with an AC option preferred (not preferred) to the B option, would produce a stronger (weaker) activation. This scheme was used to identify a neuronal indifference point (IPn), and was repeated for different A and B magnitude levels (Fig. 2a,b), as in the behavioral measures. Each option set {A,B} defined one test for computing the IPn. We focused on single neurons that were sensitive to both reward probability (Fig. 2c) and magnitude. Their IPn values (Fig. 2d, black line) varied across neurons, generally differing from the corresponding behavioral IPb values (Fig. 2d, green line). In some units (example unit 1 in Fig. 2), the IPn was close to the IPb in most tests; in other units the IPn was larger (example unit 2) or smaller (example unit 3) than the IPb across all tests. For each test, the IPn varied continuously across the population of recorded neurons (Fig. 2e), showing a general non-conformity of the neuronal utility measure with the behaviorally defined one. In economic terms, neurons with IPn > IPb were more "risk averse" than the measured behavior, while neurons with IPn < IPb were more "risk seeking" than the behavior.

Reward magnitude and probability coding in single neurons. To investigate the source of the observed heterogeneity in utility coding (IPn) across cells, we looked at the neuronal responses to a set of reward magnitude and probability levels. Two crucial features emerged when comparing the responses to safe rewards with the responses to gambles with equal expected value (EV) (Fig. 3a): (1) the responses were a nonlinear function of the EV, and (2) the relative sensitivity to reward magnitude (red) and probability (blue) varied across neurons. These two characteristics defined the "risk attitude" of each single neuron, producing a relatively stronger activity for the safe option (red) or for the gamble option (blue), at each EV.

We linked this idea with the economic constructs of utility and probability weighting functions, which correspond to subjective nonlinearities in the evaluation of reward magnitudes and probabilities, respectively. To this aim, we directly fitted a nonlinear economic model to the neuronal activity. In brief, we defined the economic value of a gamble as $v(m,p) = u(m) \cdot w(p)$ (multiplicative value model), with u and w as 2-parameter functions (Prelec, 1998). We then fitted the neural activity rate N to the value function through nonlinear least squares ($N(m,p) = a \cdot v(m,p) + b + \varepsilon$, with free parameters a, b and error term ε), using data from all presented gambles. As a result, we characterized the activity of each neuron across all tests with a set of 6 parameters, fully identifying the neuronal counterparts of the utility and probability weighting functions (Fig. 3a). By comparing these functions with the behavioral ones we confirmed the units' heterogeneity in relation to the behavior (Fig. 3b), with the neural u and w functions being generally different from the behavioral ones. Furthermore, using the neuronal economic functions we reconstructed the ICs coded by single cells: the neuronal ICs (Fig. 3c). One example neuron (unit 1) produced ICs very similar to the behavioral ones, although its u and w functions differed from the behavioral ones; other units (unit 2, unit 3) produced ICs that were clearly incompatible with the behavior, being strongly skewed towards the right (less risk seeking than the behavior) or to the left (more risk seeking than the behavior).

A reliable population utility code. Our economic-based characterization of single neurons' activity allowed us to identify unit-specific utility and probability weighting functions, which nevertheless did not generally explain or correspond to the monkey choice behavior. This puzzling result opens two questions about the role of OFC neurons: are behavior-compatible economic values represented at a population level? What are the roles of the different units in this varied population?

To answer the first question, we computed the median recovered economic functions' parameters across a subpopulation of units (N=40) and reconstructed the corresponding population utility and probability weighting functions. Results showed that, although single units in the selected subpopulation had highly varying economic u, w functions (Fig. 4a, left), their population-computed counterparts were surprisingly close to the behavioral ones (Fig. 4a, right). This was even more evident in the pattern of population-recovered ICs, which approximated the behavioral ones in great detail (Fig. 4b).

Overall, these results represent an insight into the brain mechanism for subjective value computation, proposing the existence of a neuronal counterpart of economic utility. Crucially, this neuronal signal is only evident at the neuronal population level, with single units' representation of reward variables largely incompatible with the choice behavior.

The heterogeneity in the coding of reward magnitude and probability across single units suggests possible neural implementations of the value computation mechanism. We found a few single neurons directly able to describe monkey choices, usually for a limited number of choice sets; the activity of most units corresponded to potentially different choice behaviors (e.g. risk seeking or risk averse). A downstream mechanism might modulate the contribution of single neurons, with distinct modulation patterns producing specific choice attitudes. This would allow for the flexible accommodation of different choice patterns, as required by the decision context. The current study defines a framework for the experimental validation of this interpretation.

In conclusion, we showed that a population mixture of heterogeneously coded economic constructs closely matched behavior, suggesting a neuronal mechanism for the computation of economic utility. The existence of such neuronal utility signal demonstrates a possible neural implementation of the utility-maximization principle and provides a basis for understanding the origin of subjective risk attitudes and choice biases.

Figures

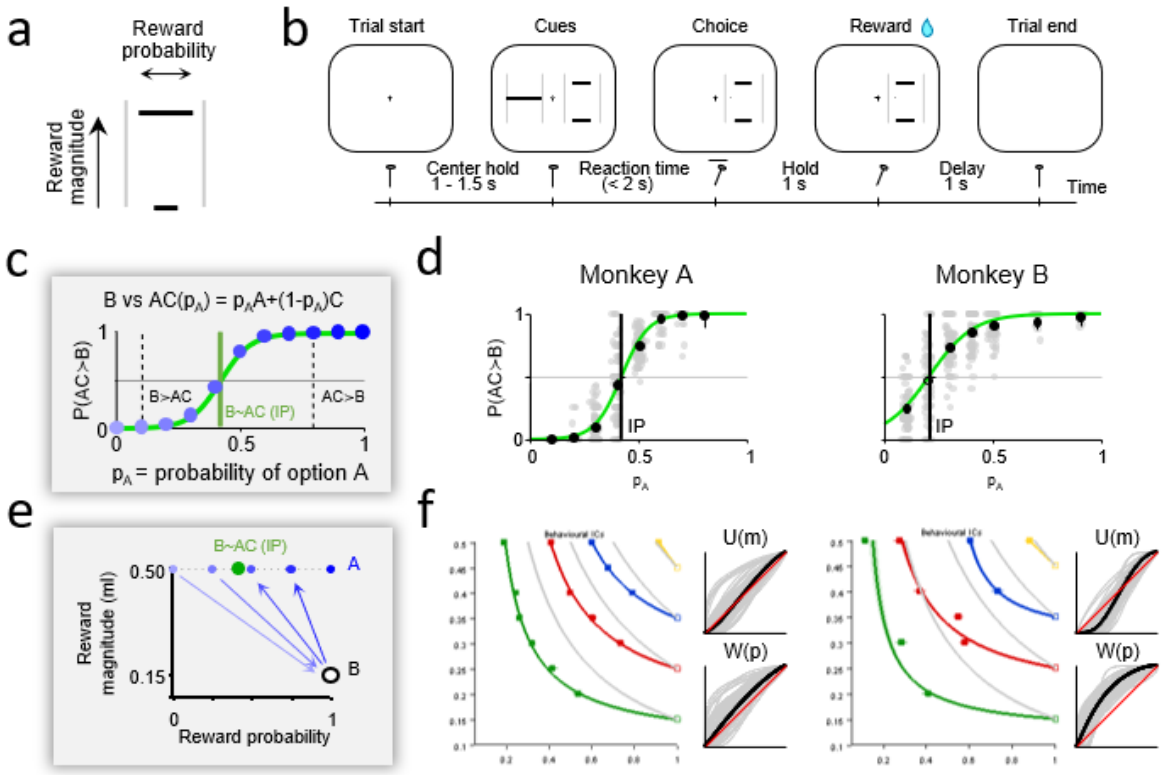

Figure 1. Utility and probability weighting functions describe risky choices.

a) Visual cue. Each horizontal bar represents the magnitude (vertical position) and probability (horizontal width) of a possible outcome. Rewards varied between 0 and 0.5 ml.

b) Trial sequence. A monkey was presented with two choice options (Cues) and revealed its preference by moving a joystick towards one side. After 1 sec the reward was delivered, conditional on the reward distribution of the selected cue.

c) Scheme of utility measurement. In choices between a gamble AC (magnitude A with probability pA and magnitude C with probability 1-pA) and safe option B, an indifference point (IP) was identified. The IP was defined as the probability (pA) for which the two options were chosen with the same choice probability, and corresponded to a utility measure for B in relation to the specific A and C, according to the continuity axiom of EUT.

d) Experimental utility measure. Utility measure for one set of magnitudes {B = 0.25 ml, A = 0.50 ml, C = 0 ml} in two monkeys. Grey dots: probability of choosing option AC (pAC) in single sessions; black dots: averages across all sessions with the corresponding 95% confidence intervals from a binomial distribution (vertical bars). The green curve is a softmax function that was fitted to the pAC to identify the behavioral indifference point (IPb) as the pA for which the softmax had value 0.5.

e) Scheme of extended utility measurements. The utility measure procedure described in (c) can be represented in a two-dimensional magnitude-probability space, where multiple utility measures can be shown. Arrows identify a choice pair, pointing to the preferred option.

f) Indifference curves. ICs in the magnitude-probability space were elicited by first measuring IPs for several combinations of starting magnitudes {A,B,C} (14 IPbs in monkey A, 10 IPbs in monkey B); for a fixed B, all IPbs had the same utility, belonging to the same IC. Choice-elicited utility (u) and probability weighting (w) functions (inset) predicted ICs from the economic model. The difference between ICs and EV-lines (points with the same expected value) quantifies the subjective valuation of choice options, indicating risk seeking (ICs shifted to the left) or risk averse (ICs shifted to the right) behavior.

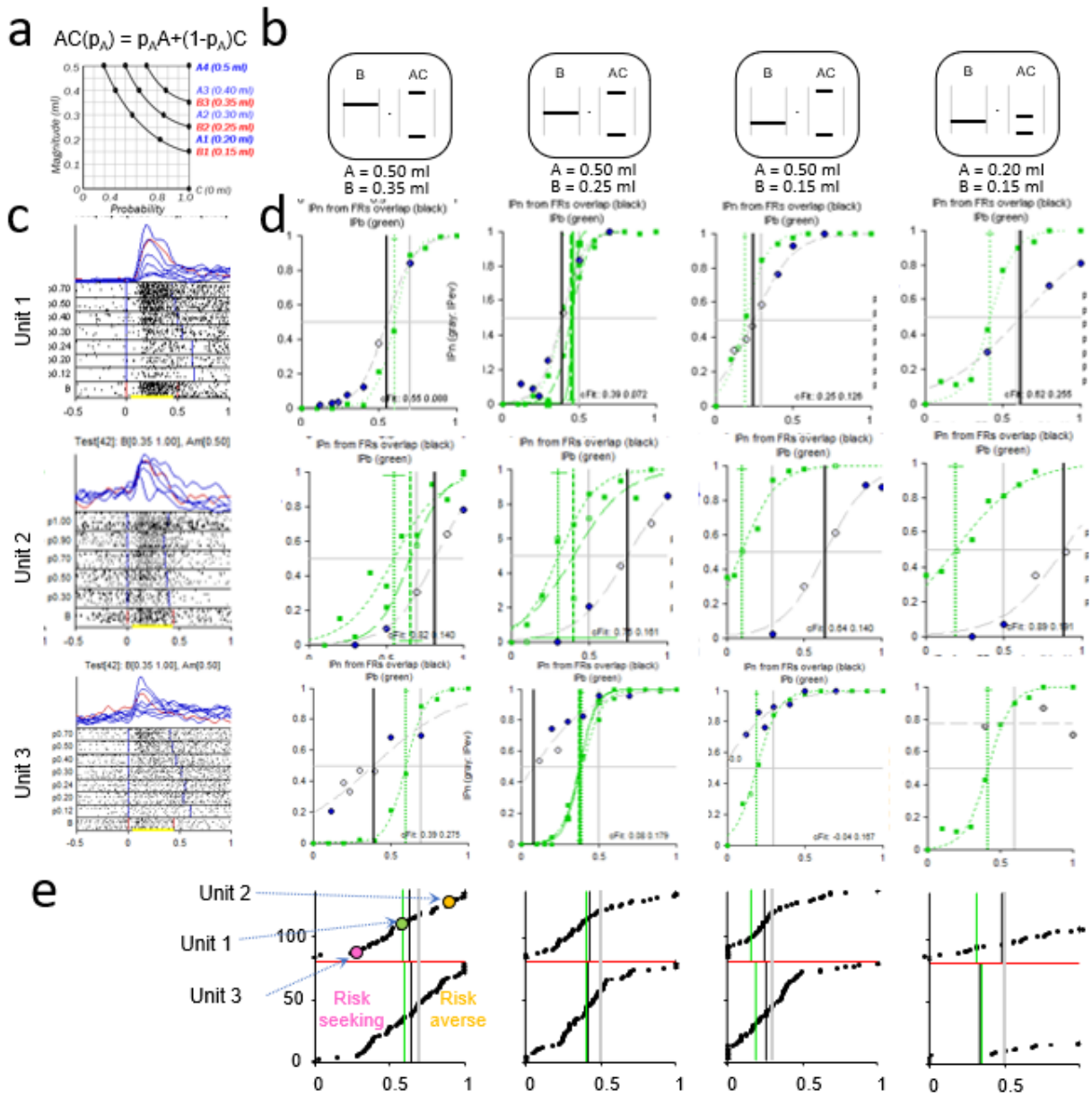

Figure 2. Heterogeneous utility code across single neurons.

a) Example set of options (magnitudes and probabilities) tested in neurons. From each options set {A,B,C} we define a utility measure test.

b) Example utility tests. Gamble cues for four sample neural and behavioral utility measures. Each test is identified by the indicated combination of A and B magnitudes (with fixed C = 0 ml).

c) Raster plots for three example neurons. Response to the specific option indicated on the y-axis (safe B, or gamble AC with probability pA). Spike density functions (gaussian smoothed activity rates, kernel 30 ms) for the safe (red) or gamble (blue) option. Yellow bar represents the time window used for computing each trial's activity rate. The x-axis represents time from Cue onset (in sec).

d) Neuronal utility measure. Neuronal indifference points (IPn, black vertical line) are compared to behavioral ones (IPb, green vertical line). Each dot represents a probability of choosing the AC option (pAC) for a specific probability of A (x-axis). The neuronal pAC (black) were decoded from the neuronal response distributions to the two choice options (see Methods). The behavioral pAC (green) were calculated as proportion of trials for which the AC option was chosen. The IPn varied across neurons and the difference between IPn and IPb varied across tests, indicating that single units were not homogeneously encoding the options' utility.

e) Heterogeneity of utility coding across units. Utility measure IPn across all recorded neurons in each of the four example tests. Each dot corresponds to IPn coded by one neuron in one specific option-set, highlighting the broad and continuous variation of prospective utility measures in single units. This variation corresponded to the coding of different risk attitudes, e.g. neurons on the left-end of each graph encoded a more risk-seeking behavior than the monkey's.

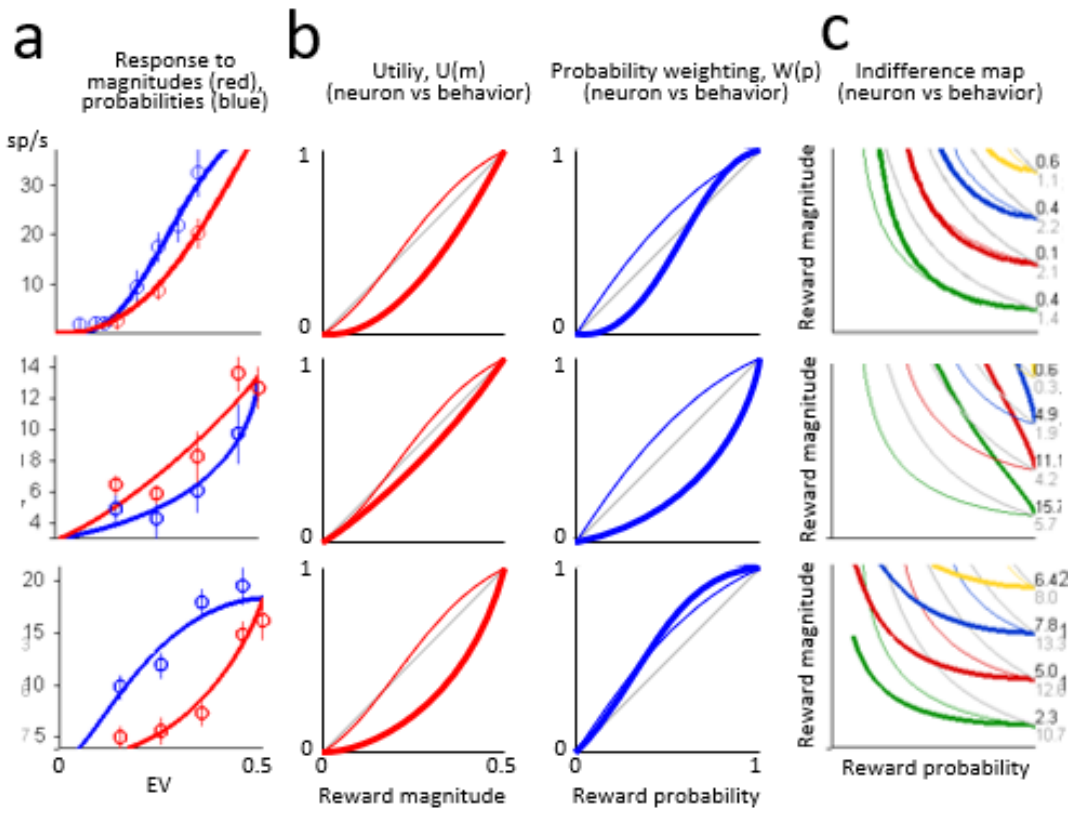

Figure 3. Heterogeneous neuronal economic functions across single neurons.

a) Single neurons' magnitude and probability response. Activity (mean ± SEM) of three example neurons to different reward magnitudes (red) and reward probabilities (blue). The curves represent the economic value functions directly fitted to each neuron's activity, and correspond to the neuronal utility and probability weighting functions, in line with the economic concepts.

b) Neuronal and behavioral utility. The utility function computed from each neuron (thick line) varied across neurons and differed from the behavioral one (thin line).

c) Neuronal and behavioral probability weighting. The probability weighting function computed from each neuron (thick) varied across neurons and differed from the behavioral one (thin).

d) Neuronal and behavioral indifference map. Comparison between neuronal (thick) and behavioral

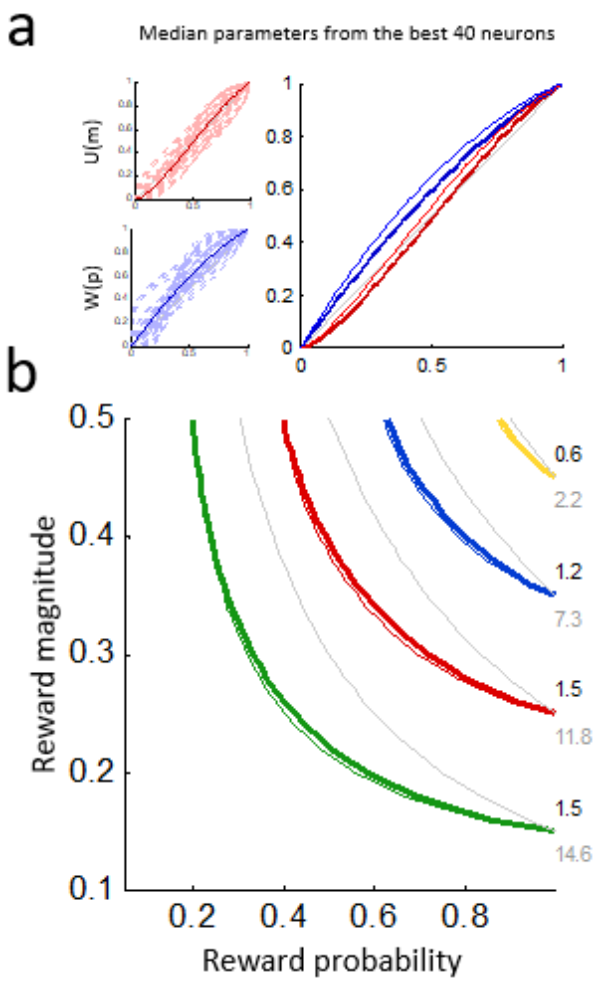

Figure 4. Reliable population utility code.

a) Population economic functions. Left: utility and probability weighting functions elicited in a population of neurons (N=40). Right: median economic functions across the neuronal population (thick line) superimposed with the corresponding behavioral estimates (thin line). The heterogeneity of single units' economic functions (left) and their dissimilarity with the behavioral ones is resolved when averaging across the neuronal population (right).

b) Population indifference map. Indifference curves (ICs) computed from the neural population economic functions (thick lines) compared with the behavioral ICs (thin lines). The precise overlap between neuronal and behavioral ICs indicates a reliable reconstruction of the behavioral choice pattern by a population of diverse neurons.

Methods

Animals

Two adult male macaque monkeys (*Macaca mulatta*; Monkey A, Monkey B), weighing 12.6 kg and 13.8 kg, respectively, were used in the experiments. Animal had not been used in other studies; behavioral results have been published in detail (Ferrari-Toniolo et al., 2021).

Ethical approval

This research has been ethically reviewed, approved, regulated, and supervised by the following UK and University of Cambridge (UCam) institutions and individuals:

UK Home Office, implementing the Animals (Scientific Procedures) Act 1986, Amendment Regulations 2012, and represented by the local UK Home Office Inspector

UK Animals in Science Committee

UK National Centre for Replacement, Refinement and Reduction of Animal Experiments (NC3Rs)

UCam Animal Welfare and Ethical Review Body (AWERB)

UCam Biomedical Service (UBS) Certificate Holder

UCam Welfare Officer

UCam Governance and Strategy Committee

UCam Named Veterinary Surgeon (NVS)

UCam Named Animal Care and Welfare Officer (NACWO).

Experimental design

Each monkey was presented with choices between a gamble and a safe option. Gambles had one non-zero outcome, which could vary in magnitude (m, from 0 ml to 0.5 ml in 0.05 ml steps) and probability (p, from 0 to 1, in 0.02 steps), and were represented respectively by the vertical position and width of a horizonal bar (Fig. 1a). The safe option represented a degenerate gamble ($p = 1$). A trial started when a cross appeared at the center of the computer monitor (eye-screen distance 50 cm); a cursor, controlled by the monkey through left-right joystick hand movements, was required to be screen-centered, and after a control period two visual cues appeared, representing the two choice options. The monkey revealed its preference by moving the cursor towards the chosen option; after holding it there for 1 sec, a liquid reward (water) was delivered through a spout directly in front of the animal's mouth. The reward was contingent on the selected option's amount and probability (Fig. 1b).

During neural recording sessions, intermixed with binary choices between a gamble and a safe option, choices between one option (gamble or safe) and no reward were presented. These *zero-alternative* trials represented a fully controlled situation for studying the neuronal responses to one specific option, and thus to investigate the neural coding of value without confounds introduced by varying, simultaneously presented alternative options. We first present neural data from the *zero-alternative* trial types, in which information was only relative to the non-zero option. We then use the same analysis for gamble-safe choice trials, in which we characterized the neural activity as a function of the reward variables (magnitude and probability) relative to the chosen option.

Behavioral data analysis

<u>Elicitation of utility measure</u>. Based on the continuity axiom of Expected Utility Theory (EUT), we defined a behavioral utility measure as the reward probability for which the monkey was indifferent between a gamble (AC) and a safe option (B) (Fig. 1c). The AC gamble was defined as a combination of two safe options (A, C) with probabilities p and $(1 - p)$, respectively. According to the continuity axiom, the resulting behavioral indifference point (IPb) represents a utility measure, being a

numerical quantity associated with the subjective evaluation of B in relation to A and C. The IPb was computed by fitting a softmax function to the probability of choosing the AC option, and identifying the point for which the softmax' value was 0.5, which corresponded to a 0.5 probability of choosing both options, i.e. to choice indifference.

Indifference map. For each fixed B, we varied the reward magnitude associated to A, obtaining a series of indifference points which identified an indifference curve (IC). For a different B magnitude, we obtained a different corresponding IC. By construction, all points on one IC had the same utility, i.e. monkey was indifferent in choices between any two points on the IC; and points on an IC corresponding to a B with increased magnitude had a higher utility.

Economic functions' modelling. We fitted an economic multiplicative value model to the choice data, through a maximum likelihood estimation procedure. The multiplicative model represented the subjective value of a gamble (reward magnitude m and probability p) as the product of a utility function $u(m)$ and a probability weighting function $w(p)$. Formally, $v(m,p) = u(m) \cdot w(p)$, where u and w were represented by 2-parameter functions (Prelec, 1998). Each *Prelec* function $y = exp(- a \cdot (- ln(x)^b))$ had two free parameters (a,b) that were estimated from the data. Using the economic u and w functions we calculated the ICs predicted by the model.

Neurophysiological recordings

Recording area. Single neurons were recorded from the orbitofrontal cortex, putatively from area 13/11 (central orbital gyrus). – setereotaxic coordinates: AP 27 to 37 mm, ML 3 to 12 mm.

Electrodes. Custom glass-coated tungsten single electrodes (up to two per session) were lowered using an electronic micromanipulator (NAN instruments).

Recoding tasks. When single units were identified and isolated, a short task (*zero-alternative* task, presenting either a gamble (magnitude 0.5 ml, probability 0.3 or 0.7) or a safe option (0.15 ml or 0.35 ml), with 8 repetitions each) was run to assess any reward related response. For units with a significant response at the Cue period (a 400 ms time-window starting at Cue onset), the main task was then run. The task included 8 to 24 *zero-alternative* trial classes (at least 4 safe options and 4 gambles), and 6 to 18 binary choice trial classes, all intermingled and equally presented for at least 8 repetitions (12 replications on average). We interrupted the session if a neuron's waveform became too small or was lost during the session. We included units that were well-isolated for at least 4 repetitions. The single-unit activity in one trial was calculated as the rate of action potentials within the Cue period, measured in spikes per second (sp/s).

Neural utility measure

Elicitation of neuronal indifference point. Mirroring the behavioral utility measure, we defined a neuronal utility measure for a particular set of reward magnitudes {A, B, C}. The neuronal indifference point (IPn) corresponded to the probability of A for which the neuron had the same response to the gamble (AC) and to the safe option (B). The IPn was computed by first decoding the probability of choosing the gamble AC from the neural activity (see below). Then, a softmax function was fitted to the decoded choice probability, identifying the point at which the softmax had a value of 0.5 (i.e. decoded choice indifference).

Decoding of choice probability. We decoded the probability of choosing option X over option Y from the neural activity, through a Receiver Operating Characteristic (ROC) analysis. Assuming that the activity of a neuron was proportional to the instantaneous value of an option, including any noise in the system, a higher activity would identify the preferred option. We thus decoded the probability of choosing option X as the probability that the activity for X was higher than the activity for Y.

Economic model fit

We fitted an economic multiplicative value model directly to the neural activity recorded in all trials. The multiplicative model represented the subjective value of a gamble (reward magnitude m and probability p) as the product of a utility function and a probability weighting function, as in the

behavioral model. This corresponded to the equation $v(m,p) = u(m) \cdot w(p)$, where u and w were 2-parameter functions (Prelec, 1998). Each Prelec function $y = exp(- a \cdot (- ln(x)^b))$ had two free parameters (a,b) to be estimated from the data. The two dimensional value function was fitted to the neural activity N through nonlinear least squares ($N = c \cdot v(m,p) + d + \varepsilon$, with additional free parameters c, d, and error term ε). The constraints imposed to the procedure (using matlab function *fminsearchbnd*) imposed $N > 0$, ensuring that meaningful activities were produced by the model. Additional constraints were imposed for biological plausibility and fit stability, by limiting the modeled activity to not exceed two standard deviations outside the recorded activity range. This was applied separately to each unit.

Neural population code

Average across neurons. To compute the coding of economic functions by a neuronal population, we computed the median value of the elicited parameters. Using the median parameters, we then calculated the population economic functions. This procedure was equivalent to averaging the fitted activity across all neurons after normalizing it, thus representing the simplest form of combining activity from multiple neurons.

References

Camerer CF (2008) Neuroeconomics: Opening the Gray Box. Neuron 60:416–419 Available at: http://dx.doi.org/10.1016/j.neuron.2008.10.027.

Ferrari-Toniolo S, Bujold PM, Grabenhorst F, Báez-Mendoza R, Schultz W (2021) Nonhuman Primates Satisfy Utility Maximization in Compliance with the Continuity Axiom of Expected Utility Theory. J Neurosci 41:2964–2979 Available at: https://www.jneurosci.org/content/41/13/2964 [Accessed October 2, 2021].

Fiorillo, C. D., Tobler, P. N., Schultz W (2003) Discrete Coding of Reward Dopamine Neurons. Science (80-) 299:1898–1902 Available at: http://www.sciencemag.org/content/299/5614/1898.short.

Lak A, Stauffer WR, Schultz W (2014) Dopamine prediction error responses integrate subjective value from different reward dimensions. Proc Natl Acad Sci U S A 111:2343–2348.

O'Doherty JP, Camerer CC (2015) Editorial overview: Neuroeconomics. Curr Opin Behav Sci 5:v–viii Available at: http://dx.doi.org/10.1016/j.cobeha.2015.10.004.

Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. Nature 441:223–226 Available at: https://www.nature.com/articles/nature04676.

Padoa-Schioppa C, Cai X (2011) The orbitofrontal cortex and the computation of subjective value: consolidated concepts and new perspectives. Ann N Y Acad Sci 1239:130–137 Available at: http://doi.wiley.com/10.1111/j.1749-6632.2011.06262.x.

Prelec D (1998) The Probability Weighting Function. Econometrica 66:497 Available at: https://www.jstor.org/stable/2998573?origin=crossref.

Stauffer WR, Lak A, Schultz W (2014) Dopamine reward prediction error responses reflect marginal utility. Curr Biol 24:2491–2500.

Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. Nature 398:704–708 Available at: www.nature.com.

von Neumann J, Morgenstern O (1944) Theory of Games and Economic Behavior. Princet Univ Press:625 Available at: http://www.archive.org/details/theoryofgamesand030098mbp.