Smarter than humans: rationality reflected in primate neuronal reward signals

Wolfram Schultz[1], Wiliam R Stauffer[2], Armin Lak[3] and
Alexandre Pastor-Bernier[1]

Rational choice, in all its definitions by various disciplines, allows agents to maximize utility. Formal axioms and simple choice designs are suitable for assessing rationality in monkeys. Their economic preferences are complete and transitive. In this paper I will describe how neuronal reward signals demonstrate a propensity for rational choice. Dopamine signals follow transitivity and satisfy first-order stochastic dominance that defines the better option. Neurons in orbitofrontal cortex reflect unchanged preferences when a dominated option is removed from the option set, thus satisfying Independence of Irrelevant Alternatives (IIA). While monkeys, with their reward neurons, may not be more rational than humans, the constraints of controlled experiments seem to allow them to behave rationally within their informational, cognitive and temporal bounds.

Addresses
[1] Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3DY, UK

[2] Systems Neuroscience Institute, University of Pittsburgh, Pittsburgh, PA 15261, USA

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

Corresponding author:
Schultz, Wolfram (Wolfram.Schultz@protonmail.com)

Introduction

At the 2017 Nobel dinner in Stockholm, my head of department at the time, Bill Harris, told his dining neighbor Richard Thaler, who had just won the Nobel Prize in economics for his work on irrational human choice, that a guy in his department had shown that monkeys can make rational economic choices. Bill had seen a talk I had given a few months earlier in a series of departmental seminars in which we catch up on each other's work. Thaler contradicted him by citing the famous example of a monkey throwing a piece of nutritious cabbage after an experimenter who had just given another monkey a much nicer food. Nevertheless, I am often asked by economists after a talk why monkeys can make rational choices when humans have such difficulties. I will use this short article to make my point, which is of course not as simple as its title, but monkeys do possess the neuronal hardware that allows them to make rational choices.

Rationality has many connotations in psychology, economics, biology and evolution [1]. Whatever the perspective, the consequences from its violation, irrationality, are quite evident; after 'poor' choices we end up worse than we would have ended otherwise. By contrast, true rational choice results in the best possible outcome, irrespective of being based on instinctive, automatic behavior without awareness of consequences or on predictions, reasoning, deliberation, emotional control and consistency. Thus, rational choice leads to maximization of economic utility that is specific for the individual decision-maker. We can measure choice and infer preference (revealed by the choice) and utility, both of which are not directly observable; we consider revealed preferences and utility as equivalent: if I prefer option x to option y, it can be said that option x has higher utility u for me than option y (formally: $x \succ y$ is equivalent to $u(x) \succ u(y)$, with $\succ$ as preference operator). By restricting rationality to the single definition of utility, we can search for its neuronal basis by studying reward utility signals in the brain.

The observation of a decision-maker trying to optimize utility has a major problem: we don't know what is best for the agent, even if s/he says 'this is best' or at least thinks so, both of which require insight and truthfulness and thus are problematic. We may extrapolate from our own utility functions, thinking what is best for me should also be best for you, but that may ultimately lead to putting value and norms onto others' behavior. Or we infer others' utility functions from their choices, but then nothing would be irrational. One source of ignorance may derive from a potential but yet unknown longer-term or even evolutionary gain, like the cabbage-throwing monkey trying to discourage a potential competitor getting advantageous food. Such behavior would reflect emotions, and indeed emotions may provide a decent definition of irrationality for psychologists. More objective, basic definitions

assume that more is better, as represented by a positive monotonic value function, but more is not always better. For an obese agent, more food is obviously not better, unless we think of famines that were frequent until a hundred years ago and which an obese agent would survive better than an individual with normal weight. Thus, rational choice is not a ready-made, unequivocal concept, and this article will approach it with minimal and empirically testable assumptions.

## Basic rationality requirements: completeness and transitivity

A basic definition of rational choice employs two axioms that need to be satisfied: completeness of preference and transitivity [2•]. Completeness postulates that agents have well-defined preferences: in a set of two options, either option x is preferred to option y, or option y is preferred to option x, or there is choice indifference between the two options. Satisfaction of the completeness axiom requires that choices are deliberate and distinguishes them from unreflected, automatic behavioral reactions, as seen with simple Pavlovian and habitual operant conditioning. Completeness can be achieved within a well-defined, restricted and fully known option set containing a finite number of alternatives. The options are mutually exclusive (choose one option or its alternative but not both) and collectively exhaustive (the set includes all available options). Frequently tested option sets contain two simultaneously presented options, allowing binary choice. In this situation, agents are induced to express complete preferences, although they could choose not to select any option, avoiding to express a preference, which would violate the completeness axiom; this is usually not observed with well-motivated agents. Thus, the completeness axiom defines choice scenarios in which agents can make rational choices. However, it is difficult to postulate specific neuronal signals representing completeness of preference, beyond general signals that occur only during choice and not with Pavlovian or habitual operant reactions.

Satisfaction of the second rationality axiom, transitivity, can be directly tested on neurons. Transitivity assumes that an agent who prefers option x to option y and option y to option z should also prefer option x to option z. If reward probabilities increase very mildly and gradually from option x to y to z, together with an under-compensating decrease of reward amounts, the agent may prefer option z to option x in the transitivity test and thus violate transitivity [3]. One reason may be that the probability increases from x to y to z are almost imperceptible and only get noticed with the bigger difference between options x and z; the now-perceived probability increase outcompetes the minor amount decrease, and option z is preferred to option x. Thus, limits of sensory and

cognitive discrimination ('just noticeable difference') may be a factor in irrational choice.

Monkeys' choices typically satisfy transitivity. Neurophysiological data usually require multiple trials for statistical analysis. Choices vary slightly between trials, which is attributed to stochastic processes intervening between option presentation and overt choice [4•]. In one such experiment [5], a monkey chooses between various amounts of fruit juice and banana (Figure 1a,b). The animal prefers large to small juice amounts (Figure 1b, black) and small juice to small banana amounts (grey), as inferred from the probability of choosing one option over the other option with $P > 0.5$ (the not-all-or-none,

Figure 1

Current Opinion in Behavioral Sciences
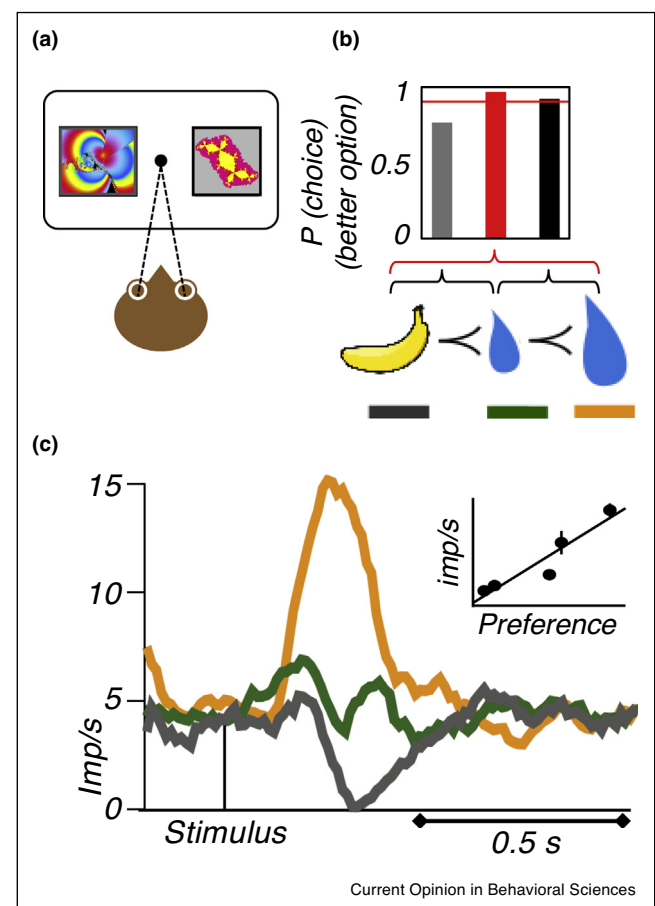
Satisfaction of transitivity.
(a) Binary choice. The animal fixates the central spot, chooses one of the two fractal stimuli by eye movement, and receives the associated reward.
(b) Large juice drop is preferred to small juice drop that itself is preferred to small banana morsel (black and grey bars). Preference of large juice drop to small banana morsel closes strong stochastic transitivity (red).
(c) Average responses from 20 dopamine neurons follow rank order of transitive preference. (a)–(c) reprinted with permission from Ref. [5].

$P < 1.0$ choice of the preferred option reflects the stochastic process that is captured by an S-shaped, rather than rectangular, choice function). Transitivity is satisfied when the animal prefers the large juice drop to the small banana morsel (red); the higher probability for the transitive choice compared to the two initial choices indicates strong stochastic transitivity and strengthens the suggestion of rational choice [6]. Starlings satisfy also stochastic transitivity [1].

Midbrain dopamine neurons in monkeys code economic utility relative to prediction (reward prediction error) [7] (for an update on dopamine responses and their multi-component nature and movement signals, see Ref. [8]). The stimulus predicting the most preferred option (large juice amount) elicits a positive dopamine prediction error response relative to the average reward predicted from past trials (Figure 1c, orange), whereas the stimulus predicting the least preferred option (small banana morsel) elicits a negative prediction error response (grey); the stimulus predicting the intermediate small juice amount elicits no response, reflecting little deviation from the predicted reward average (green). Thus, dopamine responses follow the preference ranks confirmed by compliance with transitivity and constitute a neuronal correlate for rational choice.

More is better: first-order stochastic dominance

In addition to formal axioms, rational choice can be tested with a few well-conceptualized designs. A basic test involves first-order stochastic dominance, that defines unequivocally the better option in a given choice set and tests whether more is preferred to less. Here, 'stochastic' refers to the natural tendency of rewards to be uncertain ('gambles'); this use of 'stochastic' differs from the stochasticity of choice processes that underlie preference variations in repeated trials; both meanings may apply to behavioral neurophysiology studies.

In a first-order stochastically dominant option, the probability of receiving each outcome is at least as high as in the alternative option and higher for at least one outcome. Statewise dominance is a reduced, more intuitive version in which every outcome of the dominant option is at least as good as the corresponding outcome in the alternative option and strictly better in at least one instance. Thus, would you prefer one pound for sure or, alternatively, one pound to which occasionally a second pound is added? Some people may be so horrified by not knowing whether they will ultimately receive one or two pounds that they might settle for the one pound for sure, in which case they violate first-order stochastic dominance and make an irrational choice that gives them less than the best possible outcome. There is no need to estimate a utility function to indicate how much more two pounds are worth than one pound, as long as two pounds are preferred to one pound (positive value function). By contrast, on-average better outcomes alone do not amount to first-order stochastic dominance (as choices would be prone to subjective weighting of reward amounts and probabilities).

In an empirical test for first-order stochastic dominance, a monkey chooses between a safe reward (Figure 2a, blue) and a risky gamble with two equiprobable rewards ($P = 0.5$ each) (red). Choice of the blue option delivers a small reward on every trial, whereas choice of the red option delivers either the same small reward or a larger reward. Thus, the red option is equal or better compared to its alternative in every trial and thus first-order stochastically dominates the blue option (statewise dominance). Conversely, in Figure 2b, the safe black option is better than the green gamble that delivers a smaller reward than the black option on half the trials. Cumulative plots in Figure 2c show that reward probabilities for the dominant options (red, black) are always the same or higher than for the dominated option (blue, green). Monkeys choose the better option on most trials (Figure 2d); occasional violations reflect the stochastic choice process (blue, green). Preference for the red option is not explained by risk seeking, as preference for the same risky gamble is lost to a higher safe reward (black in Figure 2b-d). Thus, the animal's choices satisfy first-order stochastic dominance with two-outcome gambles [7], and also with three-outcome gambles [9]. Monkeys also obey second-order and third-order stochastic dominance that requires intuitive, but obviously not conceptual, understanding of the reward distributions and demonstrate maximization of utility rather than physical value [7,9].

The responses of midbrain dopamine neurons to individual option stimuli are stronger for the dominant than the dominated gamble in no-choice trials (Figure 2e,f; red versus blue; first-order stochastic dominance is defined by the higher low-outcome in the red compared to blue gamble) [7]. The result is confirmed when dominance is solely defined by different outcome probabilities; the dominant option delivers the lower amount on fewer trials (both lower and upper outcomes are identical in both options) (Figure 2g). Monkeys' stochastic preferences satisfy this more general test of first-order stochastic dominance (Figure 2h). Dopamine responses in true choice trials reflect the option the animal is going to choose (chosen value response; Figure 2i): when choosing the dominant option (red), dopamine responses are higher compared to choosing the dominated option (blue). Thus, dopamine responses follow the behavioral satisfaction of first-order stochastic dominance as a neuronal correlate for rational choice.

Independence of irrelevant alternatives (IIA)

Each choice option, irrespective of being a biological reward or an economic good, is composed of multiple components (also called attributes, dimensions or aspects)

Figure 2
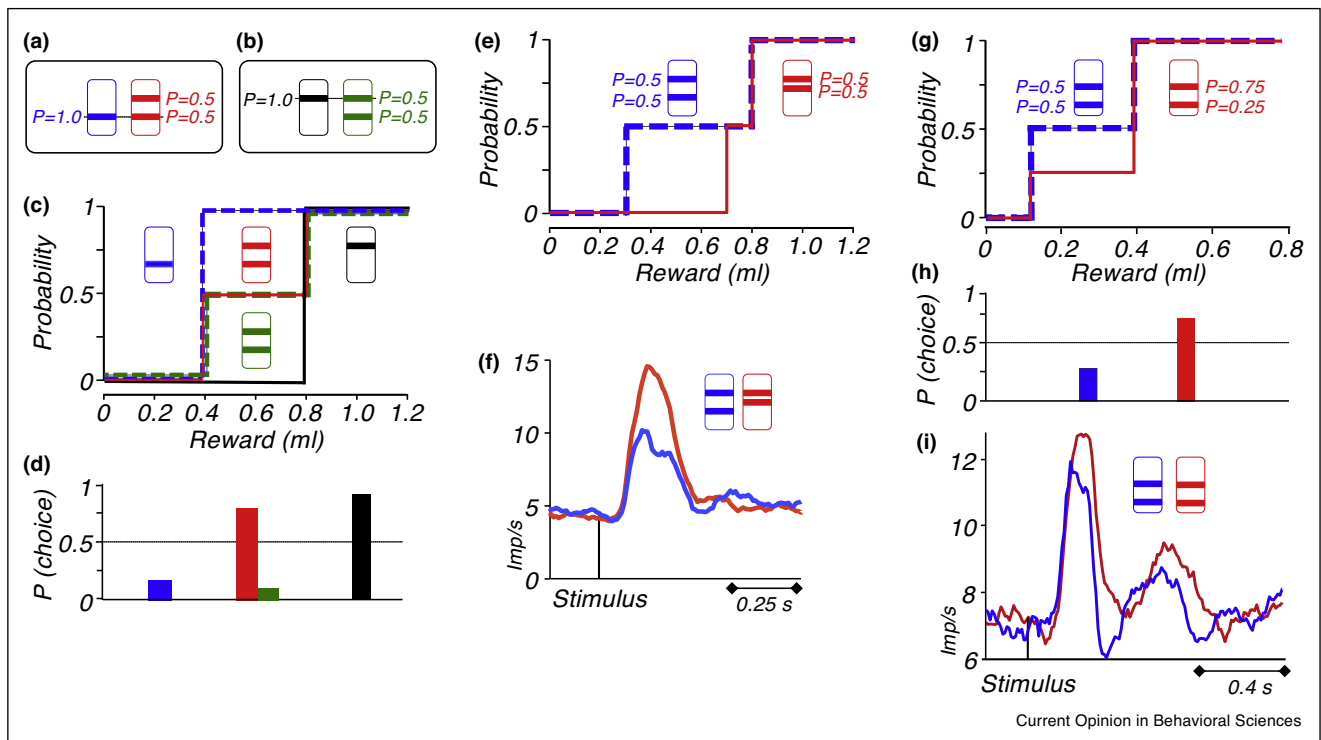
Satisfaction of first-order stochastic dominance.

(a) Two options for choice by eye movement. Red option statewise dominates blue option. Bar height indicates juice amount (higher is more, an inter-species valid metaphor [30], which is delivered with the indicated probability. The thin line connecting the two options indicates same reward amount for the connected bars.

(b) as (a) but black option dominates green option.

(c) Cumulative probability distributions for options shown in (a) and (b). Higher probability for larger rewards indicates dominance (from right to left).

(d) Choice of red and black options stochastically satisfies statewise dominance.

(e) Two-bar choice options with appropriate symmetry for controlled neuronal recordings, and their probability distributions.

(f) Stronger average responses from 52 dopamine neurons to dominating option. (a)–(f) reprinted with permission from Ref. [7].

(g) Choice options and their probability distributions for testing first-order stochastic dominance based only on reward probability differences. Red option first-order stochastically dominates blue option.

**(h)** Choice of red option stochastically satisfies first-order stochastic dominance.

(i) Stronger average responses from 99 dopamine neurons to dominating option. Re-analyzed data from Ref. [31].

and thus constitutes a bundle. The bundle components may be integral parts of a reward or good, like quantity and probability of goods [3], or consist of separable items, like meat and vegetable of a meal. Preferences for bundles are revealed by measurable choice of bundles within a given option set that contains two or more bundles. Thus, the revealed preference for a given option can be defined as the probability of choosing that option over all other options within the option set. The preferences reflect the integrated utility of their components (rather than the utility of a single component alone, called lexicographic preferences). The integration requires additional computation and thus adds unreliability that renders bundle choice more vulnerable [10] and may lead to preference reversal and thus irrational choice. The rationality is captured by the Weak Axiom of Revealed

Preference (WARP); when option x is preferred to option y in a given option set, there cannot be any option set containing both x and y in which y is preferred to x [2•]. Popular tests of WARP concern the Independence of Irrelevant Alternatives (IIA) that is violated when reduction or expansion of option sets by removing or adding one or more dominated (non-preferred) options leads to preference reversal, as seen with the similarity, compromise, asymmetric dominance and attraction effects [11,12•,13••,14••].

An empirical assessment of IIA may employ Arrow's version of WARP, whose satisfaction requires the dominant bundle to remain preferred to all other bundles within the bundle set when a dominated (irrelevant) bundle is removed from the set [15•]. Figure 3a shows

Figure 3
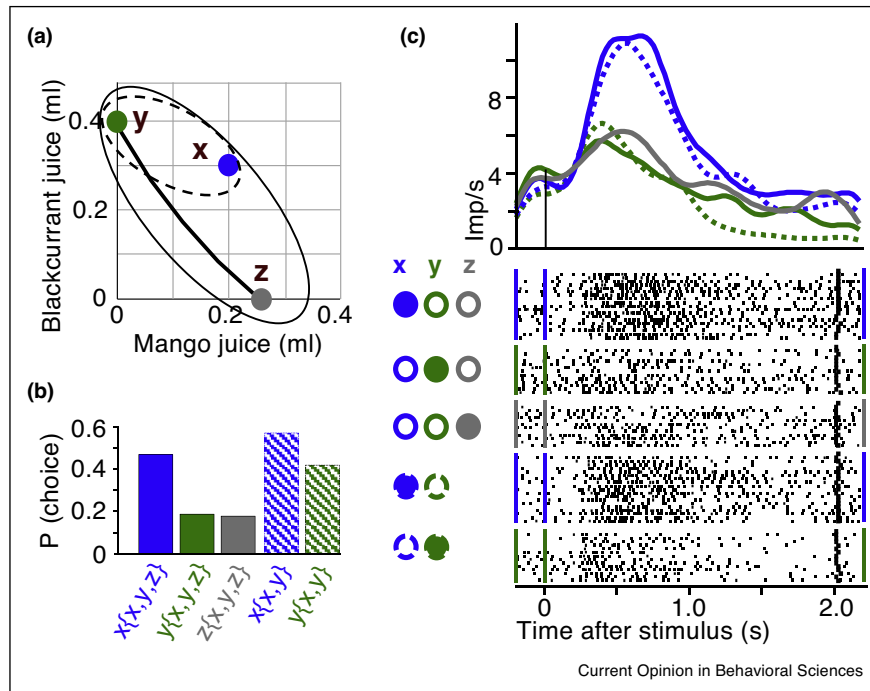
Satisfaction of Arrow's Weak Axiom of Revealed Preference (WARP) addressing Independence of Irrelevant Alternatives (IIA).
(a) Positions of bundles on behavioral choice indifference curve (y, z) and above (x), suggesting preference of bundle x to bundles y and z. Elipsoids refer to three-bundle set ({x, y, z}; solid) and two-bundle set ({x, y}; dotted).
(b) Behavioral satisfaction of WARP: bundle x remains preferred (x{x,y,z}, blue) when restricting the three-bundle set {x,y,z} to two-bundle set {x,y} (bundle x was chosen with highest probability in both sets). Bars show choice probability while recording the neuron shown in c (N = 83 choices).
(c) Neuronal satisfaction of WARP in orbitofrontal cortex. Stronger response with choice of preferred bundle x (blue) compared to choice of alternative bundles y (green) or z (grey) ($P < 0.002$ for bundle x versus y and z; two-factor Anova). Responses to preferred bundle x remain strongest with restriction from three-bundle set {x,y,z} (solid blue) to two-bundle set {x,y} (dotted blue) (N = 16 neurons; $P > 0.05$; response to bundle x in three-bundle set versus two-bundle set; two-factor Anova). Filled dots at left indicate chosen bundles whose responses are shown in rasters to the right. (a)–(c) reprinted with permission from Ref. [16•].

such a test [16•]: within the three-bundle set {x,y,z} (solid ellipsoid), bundles y and z are positioned on the same choice indifference curve, indicating equal preference between them, whereas bundle x is positioned above the indifference curve, suggesting preference to bundles y and z. The reduced bundle set {x,y} (dotted ellipsoid) contains the same two bundles x and y but not bundle z. Within the three-bundle set {x,y,z}, a monkey prefers bundle x to bundles y and z (Figure 3b; filled bars). Importantly, when the option set is reduced to two options {x,y} by removing bundle z, the animal keeps preferring bundle x to bundle y (striped bars). These choices satisfy Arrow's WARP. In a different species and with a different design, the choices of starlings comply with IIA between two-option and three-option sets [1].

Reward neurons in the orbitofrontal cortex (OFC) of monkeys, when tested with the three-bundle option set {x,y,z}, show the strongest response to the dominant bundle x and weaker responses to the dominated bundles y and z (Figure 3c, solid lines and filled circles); bundle z,

which is as much preferred by the animal as bundle y, elicits a similar response as bundle y, confirming the neuronal relationship to bundle dominance. Importantly, when tested with the reduced option set {x,y}, the still-preferred bundle x elicits very similar responses in the same OFC neurons (blue dotted line and filled circle) as in the three-bundle set {x,y,z} (and bundle y elicits a similar response as in the {x,y,z} set, green dotted). These OFC responses reflect the maintained bundle dominance despite change in option set size and thus provide a neuronal correlate for satisfaction of Arrow's WARP for IIA.

Not smarter, just more constrained

Why do monkeys and their reward signals seem rational, as opposed to humans? Contrary to the monkeys' satisfaction of first-order stochastic dominance and Arrow's WARP, humans often choose first-order dominated options and typically violate IIA [12•]. The reason are unlikely to be superior monkey intelligence. Even plants are making rational 'choices' without conceivable access

to insight, awareness or other cognitive processes [17[••]]: with low, close-to-survival fertilizer concentrations, plant roots prefer 'risky' soil whose fertilizer varies around the survival level, rather than soil with constant, below survival-level fertilizer concentration of same mean; only with more fertilizer do the plants prefer 'safe' soil. This behavior follows risk-sensitive foraging theory: with mean food amount too low to survive the night, birds prefer risky options with food peaks sufficient for survival, rather than fixed food with same low mean insufficient for survival [18[•]]. Thus, plants and animals rationally satisfy first-order stochastic dominance as shown in Figure 2a: when bar height scales with survival, only the top red bar allows survival. So, how come that plants, birds and monkeys can make rational choices when humans so often fail?

Monkeys perform tens of thousands of trials over several weeks and months in constant, well-constrained laboratory environments, with full knowledge and daily experience of constant reward distributions. Such stable situations conceivably reduce reference-dependency [19] and adaptation to reward probability distributions [20–25,26[•]], which may underlie irrational choices [14[••],27,28,29[•]]. By contrast, humans are tested in much fewer trials and more complex tasks with fluent or unknown reward distributions. The primate tests may be artificial and unrepresentative of daily life, but they clearly demonstrate the existence of a basic propensity for rational choice and the necessary neuronal hardware.

The fact that monkeys make rational choices in well defined, experienced and understood situations but humans perform with many fewer constraints relates to the issue of 'Bounded Rationality': agents are more likely to make rational choices within bounds defined by available information, cognitive abilities and time to make a decision [29[•]]. The well-constrained laboratory situation may not require the animals to exceed their informational, cognitive and temporal limits, thus reducing uninformed decisions, poor understanding and time pressure. Working within such bounds, monkeys may perfectly well choose rationally, as would humans, birds and plants. Of course, their rational choices is likely restricted to the laboratory and similarly well defined and constrained environments. As the saying goes 'Good fences make good neighbors', but agents should beware of acting outside their bounds where they have insufficient understanding, information and time and make disadvantageous decisions.

While irrational choice is usually perceived as being bad and preventing utility maximization, it has survived evolutionary pressure and is present in modern homo sapiens post-Neanderthalensis. Like exploration, irrational choice may inadvertently lead to sampling of options that are believed to be suboptimal from past experience but which may have unknowingly improved in the meantime and are now advantageous for the agent. Maybe a linear combination of bounded rationality and irrational choice is better in the long run than either the comfortable rigidity from bounded rationality or the disruptive chaos from irrational choice alone?

Conflict of interest statement
Nothing declared.

Acknowledgements

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
•• of outstanding interest

1. Monteiro T, Vasconcelos M, Kacelnik A: **Starlings uphold principles of economic rationality for delay and probability of reward**. *Proc R Soc B* 2013, **280**:20122386.

2. Mas-Colell A, Whinston M, Green J: *Microeconomic Theory*. New
• York: Oxford University Press; 1995.
A very helpful and broad textbook of economic theory at graduate student level.

3. Tversky A: **Intransitivity of preferences**. *Psychol Rev* 1969, **76**:31-48.

4. Woodford M: **Modeling imprecision in perception, valuation,**
• **and choice**. *Ann Rev Econ* 2020, **12**:579-601.
The development of a model of stochastic choice that should be useful for understanding neural mechanisms.

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

6. Luce RD, Suppes P: **Preference, utility, and subjective probability**. In *Handbook of Mathematical Psychology*, , vol 3. 252-410. Edited by Luce RD, Bush EE, Galanter E. New York: Wiley; 1965:

7. Stauffer WR, Lak A, Schultz W: **Dopamine reward prediction error responses reflect marginal utility**. *Curr Biol* 2014, **24**:2491-2500.

8. Schultz W: **Recent advances in understanding the role of phasic dopamine activity**. *F1000Research* 2019, **8**:1680.

9. Genest W, Stauffer WR, Schultz W: **Utility functions predict variance and skewness risk preferences in monkeys**. *Proc Natl Acad Sci U S A* 2016, **113**:8402-8407.

10. Pelletier G, Fellows LK: **A critical role for human ventromedial frontal lobe in value comparison of complex objects based on attribute configuration**. *J Neurosci* 2019, **39**:4124-4132.

11. Simonson I: **Choice based on reasons: the case of attraction and compromise effects**. *J Consum Res* 1989, **16**:158-174.

12. Rieskamp J, Busemeyer JR, Mellers BA: **Extending the bounds**
• **of rationality: evidence and theories of preferential choice**. *J Econ Lit* 2006, **44**:631-661.
A very helpful survey of IIA violations.

13. Gluth S, Hotaling JM, Rieskamp J: **The attraction effect**
•• **modulates reward prediction errors and intertemporal choices**. *J Neurosci* 2017, **37**:371-382.
Important demonstration of neural reward prediction error responses following IIA violations, thus combining concepts of Reinforcement Learning Theory and Revealed Preference Theory.

14. Dumbalska T, Li V, Tsetsos K, Summerfield C: **A map of decoy**
•• **influence in human multialternative choice**. *Proc Natl Acad Sci U S A* 2020, **117**:25169-25178.

A comprehensive model for defining different IIA problems on a numeric scale.

15. Arrow KJ: **Rational choice functions and orderings**. *Economica*
   • 1959, **26**:121-127.
The classic, and one of the first, formalisations of the Independence of Irrelevant Alternatives (IIA), which our experiments used for testing rational choice by reducing option sets.

16. Pastor-Bernier A, Stasiak A, Schultz W: **Orbitofrontal signals for**
   • **two-component choice options comply with indifference
   curves of Revealed Preference Theory**. *Nat Commun* 2019,
   10:4885.
The first neurophysiological study on distinct multi-component choice options, employing basic economic formalisms.

17. Dener E, Kacelnik A, Shemesh H: **Pea plants show risk**
   •• **sensitivity**. *Curr Biol* 2016, **26**:1763-1767.
The stunning description of rational 'choice' in plants.

18. Caraco T, Martindale S, Whitham TS: **An empirical**
   • **demonstration of risk-sensitive foraging preferences**. *Anim Behav* 1980, **28**:820-830.
The foundation of risk-sensitive foraging theory.

19. Köszegi B, Rabin M: **A model of reference-dependent
   preferences**. *Q J Econ* 2006, **121**:1133-1165.

20. Tremblay L, Schultz W: **Relative reward preference in primate
   orbitofrontal cortex**. *Nature* 1999, **398**:704-708.

21. Tobler PN, Fiorillo CD, Schultz W: **Adaptive coding of reward
   value by dopamine neurons**. *Science* 2005, **307**:1642-1645.

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

23. Kobayashi S, Pinto de Carvalho O, Schultz W: **Adaptation of
   reward sensitivity in orbitofrontal neurons**. *J Neurosci* 2010,
   30:534-544.

24. Cai X, Padoa-Schioppa C: **Neuronal encoding of subjective
   value in dorsal and ventral anterior cingulate cortex**. *J Neurosci*
   2012, **32**:3791-3808.

25. Louie K, Grattan LE, Glimcher PW: **Reward value-based gain
   control: divisive normalization in parietal cortex**. *J Neurosci*
   2011, **31**:10627-10639.

26. Padoa-Schioppa C, Rustichini A: **Rational attention and**
   • **adaptive coding: a puzzle and a solution**. *Am Econ Rev* 2014,
   104:507-513.
A welcome formalization of adaptive neuronal coding.

27. Roe RM, Busemeyer JR, Townsend JT: **Multialternative decision
   field theory: a dynamic connectionist model of decision
   making**. *Psychol Rev* 2001, **108**:370-392.

28. Soltani A, De Martino B, Camerer C: **A range-normalization
   model of contextdependent choice: a new model and
   evidence**. *PLoS Comput Biol* 2012, **8**:e1002607.

29. Simon HA: **A behavioral model of rational choice**. *Q J Econ*
   • 1955, **69**:99-118.
The foundation of the concept of Bounded Rationality.

30. Lakoff G, Johnson M: *Metaphors We Live By*. University of Chicago
   Press; 1980.

31. Lak A, Stauffer WR, Schultz W: **Dopamine neurons learn relative
   chosen value from probabilistic rewards**. *eLife* 2016, **5**:e18044.