David V. Smith · Patricia L. Lockwood ·
Dominic S. Fareri *Editors*

Neuroeconomics: Core Topics and Current Directions

Springer

Neuroeconomics: Core Topics and Current Directions

David V. Smith • Patricia L. Lockwood •
Dominic S. Fareri
Editors

Neuroeconomics: Core Topics and Current Directions

Springer

Editors

David V. Smith
Department of Psychology and Neuroscience
Temple University
Philadelphia, PA, USA

Patricia L. Lockwood
Centre for Human Brain Health
University of Birmingham
Birmingham, UK

Dominic S. Fareri
Gordon F. Derner School of Psychology
Adelphi University
Garden City, NY, USA

ISBN 978-3-032-02924-9 ISBN 978-3-032-02925-6 (eBook)
https://doi.org/10.1007/978-3-032-02925-6

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Neuroeconomics has matured into a dynamic and methodologically sophisticated discipline dedicated to understanding the neural, cognitive, and computational mechanisms underlying decision-making. What began as an interdisciplinary effort integrating concepts from neuroscience, psychology, and economics has evolved into a cohesive research enterprise that investigates how organisms assign value, form preferences, and choose among competing options. At its core, neuroeconomics provides a robust framework for exploring how decisions are made under uncertainty, emotional influences, social contexts, or constraints of time and effort. This volume, *Neuroeconomics: Core Topics and Current Directions*, offers a structured overview of recent advances, reflecting both the conceptual maturity and ongoing evolution of the field.

This book is designed to serve as an essential resource for researchers, clinicians, practitioners, and graduate students. Rather than providing an exhaustive historical survey, the chapters focus explicitly on current directions in neuroeconomics, synthesizing recent empirical findings in concise and accessible reviews. A key aim is to foster dialogue not only among specialists within neuroeconomics but also with researchers from adjacent disciplines who wish to engage with contemporary work in the field. By highlighting current trends and new insights, we hope this volume can stimulate novel connections and interdisciplinary collaborations.

The initial chapters establish historical and methodological foundations, tracing neuroeconomics from its origins in reward processing and reinforcement learning. These contributions emphasize how computational models and neuroscientific techniques have reshaped traditional assumptions about value and choice. Interdisciplinary methods—including lesion studies, neuroimaging, and behavioral paradigms—are discussed as central tools, providing a common framework that unites research across the diverse topics covered in subsequent chapters.

Building upon these foundations, later chapters investigate how value is represented and updated within dynamic and uncertain environments. Research on learning and valuation elucidates the processes by which individuals acquire and revise beliefs through feedback, integrate probabilistic information, and encode risk and ambiguity across neural systems. Rather than assuming stable preferences,

these chapters emphasize the dynamic nature of valuation and its responsiveness to contextual influences. Closely related are discussions of temporal discounting, effort-based choice, and self-control. These chapters clarify how people evaluate trade-offs between immediate and future outcomes, linking behavioral observations to underlying neural and cognitive processes of goal-directed behavior and impulse regulation.

Subsequent chapters explore the modulatory roles of affect and context on decision-making behavior. Here, authors examine how emotional states, environmental framing, and cognitive interpretations shape valuation and choice architecture. Whether through amplifying affective responses or subtly altering evaluations via contextual cues, these modulatory influences demonstrate the inherent flexibility of preferences. Social context emerges as particularly influential, profoundly shaping how decisions are evaluated and enacted. Chapters on social preferences and social decision-making highlight the role of neural circuits involved in interpersonal evaluations of fairness, cooperation, and reciprocity. These findings underscore the notion that decision-making rarely occurs in isolation; instead, it is deeply embedded within networks of relationships and shared experiences.

The book also addresses significant variability in decision-making processes across individuals. Chapters on individual differences explore how factors such as developmental stage, psychiatric conditions, hormonal influences, and sociocultural identities affect valuation processes and behavioral outcomes. Rather than treating such variability as extraneous noise, these contributions position it as essential for understanding how decision-making systems function across diverse populations. This emphasis challenges normative assumptions and advocates for more inclusive, representative models of neuroeconomic behavior.

Finally, chapters in the concluding section extend insights from neuroeconomics into applied and translational domains, demonstrating the field's relevance to real-world decision making. Topics include consumer behavior, marketing influence, prosocial giving, and public policy design. These chapters illustrate how neural and computational methods can forecast aggregate behaviors, reveal latent preferences, and guide interventions aimed at enhancing decision outcomes. Translational applications serve as practical tests for theoretical models, allowing researchers to evaluate the validity of neuroeconomic insights beyond controlled laboratory settings.

Throughout this volume, contributors were encouraged to integrate foundational knowledge with recent empirical developments, identify important open questions, and situate their research within broader theoretical and societal contexts. Although each chapter stands independently as a concise and accessible review, the overall structure emphasizes conceptual continuity across diverse research domains. Readers will encounter recurring themes—such as interactions between valuation and self-control systems, the role of uncertainty, and the pervasive influence of social cognition—that connect insights from individual choices to collective outcomes.

In summary, this volume aims to provide students, researchers, clinicians, and practitioners with a broad yet focused introduction to contemporary neuroeconomics. By highlighting current directions, synthesizing recent findings, and

presenting a wide range of topics in an accessible form, we hope this book will serve as both a valuable reference and a catalyst for future interdisciplinary inquiry and collaboration, deepening engagement with the fundamental questions that define the field today.

Philadelphia, PA, USA — David V. Smith
Birmingham, UK — Patricia L. Lockwood
Garden City, NY, USA — Dominic S. Fareri

Contents

About the Editors

David V. Smith is an Associate Professor of Psychology and Neuroscience at Temple University, where he directs the Neuroeconomics Laboratory and serves as Associate Director of Temple's Brain Research and Imaging Center. After earning his PhD in Cognitive Neuroscience from Duke University in 2012, he completed a postdoctoral fellowship at Rutgers University–Newark and launched his lab at Temple in 2017. His research examines the neural mechanisms that support social and economic decision-making across the lifespan, with particular interest in how people evaluate rewards, respond to incentives, and cooperate with others. Combining neuroimaging, computational modeling, and transcranial electrical stimulation, his work aims to advance basic understanding and inform translational applications in clinical and policy contexts. His research has been supported by the National Institute of Mental Health, the National Institute on Drug Abuse, and the National Institute on Aging. At Temple, he teaches courses on sensation and perception, decision neuroscience, behavioral economics, statistics, and functional neuroimaging. He has been named a Rising Star and a Fellow by the Association for Psychological Science. In addition, he has received Temple's Psychology Honors Excellence in Mentoring Award as well as the College of Liberal Arts' Presidential Faculty Award for excellence in teaching and mentoring. Finally, he also serves as an Associate Editor at *Collabra: Psychology* and is an active peer reviewer across leading journals in psychology, neuroscience, and decision science.

Patricia L. Lockwood is a Professor of Decision Neuroscience and Wellcome Trust/Royal Society Sir Henry Dale Fellow at the University of Birmingham, where she directs the Social Decision Neuroscience Lab (www.sdn-lab.org). She was previously a Junior Research Fellow at Christ Church, University of Oxford, and a Medical Research Council Fellow at the University of Birmingham, University of Oxford, and the University of Zurich. She holds a PhD in Psychology from University College London. Her lab investigates social learning and decision-making across the lifespan and in neurological and psychiatric disorders using a mixture of computational modeling, behavioral measures, self report, patient

studies, and neuroimaging. She combines frameworks from psychology, economics, ecology, and decision neuroscience to capture how, when, and why people learn and make choices that have consequences for themselves and other people. She holds multiple nationalities (British, Portuguese, Brazilian), and she is a Fellow of the Association for Psychological Science, and a member of the Women of the Future Network. She has received the Janet Taylor Spence Award for Transformative Early Career Contributions from the Association for Psychological Science and the Society for Neuroeconomics Early Career Award.

Dominic S. Fareri is an Associate Professor of Psychology at Adelphi University where he directs the Fareri Lab for Social Learning and Decision-Making. He is also the Director of Undergraduate Psychology and Neuroscience at Adelphi University. He received his PhD in Psychology at Rutgers University-Newark in 2013, and completed his postdoctoral training at the University of California, Los Angeles and Columbia University in 2015 before joining the faculty of Adelphi University in 2015. His lab investigates how the social environment becomes integrated into value-based computations supporting decision-making. Implementing functional neuroimaging alongside behavioral and computational approaches, his research draws on frameworks from psychology, cognitive neuroscience, decision neuroscience to characterize how day-to-day social contexts change our decision-making and our subjective experiences of rewards. His work has been supported by the National Institute of Mental Health and the National Science Foundation. He serves on the editorial board for *Social Cognitive and Affective Neuroscience, Imaging Neuroscience and Scientific Reports.* He was named a Rising Star by the Association for Psychological Science, has received the Faculty Excellence Award for Scholarship from Adelphi University, and is serving as the President of the Social and Affective Neuroscience Society (SANS) for the 2025–2026 year.

Part I
Foundations of Neuroeconomics: History and Methods

David V. Smith, Patricia L. Lockwood, and Dominic S. Fareri

Neuroeconomics is an interdisciplinary field integrating neuroscience, psychology, and economics to understand how people make decisions. Departing from traditional economic models, which infer preferences solely from observed choices, neuroeconomics seeks to elucidate how the brain computes value, updates beliefs, manages uncertainty, and incorporates social information (Glimcher, 2010). This section introduces foundational studies that defined the early trajectory of the field, establishing core concepts, questions, and methodologies. It also presents key empirical and computational methods—such as neuroimaging, reinforcement learning models, lesion studies, and behavioral economic tasks—that have become standard approaches in neuroeconomics. Collectively, these contributions provide an essential starting point for understanding both the historical context and future directions of neuroeconomic research.

One of neuroeconomics' earliest and most influential findings is the discovery that midbrain dopamine neurons encode reward prediction errors—the discrepancy between expected and actual outcomes. Schultz (this volume) reviews foundational research demonstrating how these neural signals instantiate learning mechanisms described by temporal difference models, facilitating adaptive value updating over time (Schultz et al., 1997; Schultz, 2016). His chapter further details how the brain differentiates reward components, such as magnitude, probability, and delay, across distinct yet interacting valuation systems. This biologically grounded understanding of utility laid the groundwork for computational modeling in human decision-making research.

D. V. Smith (✉)
Temple University, Philadelphia, PA, USA
e-mail: david.v.smith@temple.edu

P. L. Lockwood
University of Birmingham, Birmingham, UK

D. S. Fareri
Adelphi University, Garden City, NY, USA

Building upon this neural understanding of value, Terry and Huettel (this volume) explore foundational experiments addressing uncertainty, specifically distinguishing between risk and ambiguity. They highlight findings that brain regions such as the anterior insula and dorsolateral prefrontal cortex exhibit differential sensitivity to uncertainty. Such sensitivity challenges economic perspectives that treat risk merely as a function of expected value, instead revealing that uncertainty engages distinct affective and cognitive neural systems (Loewenstein et al., 2008). This research advanced a richer conceptualization of decision making under uncertainty, recognizing it as a process shaped by emotional experiences and cognitive evaluations rather than purely statistical calculations.

A complementary methodological perspective emerges in Fellows (this volume), who leverages lesion studies to uncover the causal roles of prefrontal subregions in value-based decision making. Her chapter illustrates that orbitofrontal cortex damage specifically impairs option evaluation and comparison without broadly affecting executive functions, emphasizing the distinct contributions of prefrontal regions to economic behaviors. Lesion-based dissociations thus provide crucial insights, highlighting how economic choice involves distributed yet distinct components within the frontal lobe—insights not achievable through neuroimaging alone.

Cooper and McClure (this volume) further expand methodological approaches by synthesizing two decades of research on delay discounting. Their chapter traces how intertemporal choice paradigms have revealed distinct neural signatures differentiating immediate from delayed rewards. From a computational perspective, they argue discounting behaviors arise not from a singular self-control mechanism but from dynamic interactions among memory, valuation, and prospection systems. They identify new frontiers, including episodic memory's role and direct intracranial recording's potential to clarify dopamine's role in temporal decision-making processes.

Another foundational strand of neuroeconomic research examines how social contexts influence valuation and learning processes. Wu and O'Doherty (this volume) describe how individuals evaluate outcomes affecting others through neural circuits similar to those involved in self-relevant reward processing. Extending this theme, Ratala and Sanfey (this volume) analyze interactive decisions related to trust, fairness, and cooperation through structured tasks like the Trust and Ultimatum Games. They demonstrate that valuation systems dynamically interact with social cognition networks, notably the temporoparietal junction and medial prefrontal cortex, integrating economic and interpersonal motivations (Chang et al., 2011; Rilling & Sanfey, 2011). These insights establish social decision-making as a central and integral area within neuroeconomic inquiry, rather than a peripheral topic.

Finally, these core neuroeconomic mechanisms extend to practical, everyday decisions in consumer contexts. Reeck (this volume) applies neuroeconomic principles to consumer behavior, illustrating how neural responses to prices, branding, and product features encode subjective value. Her chapter highlights neural data's potential to predict preferences more accurately than traditional behavioral or self-report methods (Plassmann et al., 2007; Karmarkar & Plassmann, 2019), demonstrating neuroeconomics' relevance to real-world applications.

While this section emphasizes foundational contributions that shaped neuroeconomics, it also identifies emerging directions for future research. Key open questions include understanding how value representations evolve over developmental and contextual timescales, elucidating neural mechanisms underlying social learning in complex multi-agent settings, and leveraging neural models to inform economic policy and ecological decision-making contexts. These early studies laid a robust groundwork, continuing to inspire ongoing efforts to deepen our understanding of decision-making.

References

Chang, L. J., Smith, A., Dufwenberg, M., & Sanfey, A. G. (2011). Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron, 70*(3), 560–572. https://doi.org/10.1016/j.neuron.2011.02.056

Cooper, A., & McClure, S. M. (this volume). Delay discounting: Empirical foundations and future directions. In D. V. Smith, P. Lockwood, & D. S. Fareri (Eds.), *Neuroeconomics: Core topics and current directions*. Springer. Preprint: https://doi.org/10.31234/osf.io/twn3f_v1

Fellows, L. K. (this volume). What lesion studies have taught us about frontal lobe contributions to decision-making. In D. V. Smith, P. Lockwood, & D. S. Fareri (Eds.), *Neuroeconomics: Core topics and current directions*. Springer. Preprint: https://doi.org/10.31234/osf.io/geqnc_v1

Glimcher, P. W. (2010). *Foundations of neuroeconomic analysis*. Oxford University Press.

Karmarkar, U. R., & Plassmann, H. (2019). Consumer neuroscience: Past, present, and future. *Organizational Research Methods, 22*(1), 174–194. https://doi.org/10.1177/1094428117730598

Loewenstein, G., Rick, S., & Cohen, J. D. (2008). Neuroeconomics. *Annual Review of Psychology, 59*, 647–672. https://doi.org/10.1146/annurev.psych.59.103006.093710

Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2007). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences, 105*(3), 1050–1054. https://doi.org/10.1073/pnas.0706929105

Ratala, A., & Sanfey, A. G. (this volume). Foundations of social decision-making. In D. V. Smith, P. Lockwood, & D. S. Fareri (Eds.), *Neuroeconomics: Core topics and current directions*. Springer.

Reeck, C. (this volume). Foundations of consumer neuroscience. In D. V. Smith, P. Lockwood, & D. S. Fareri (Eds.), *Neuroeconomics: Core topics and current directions*. Springer. Preprint: https://doi.org/10.31234/osf.io/aeghp

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology, 62*, 23–48. https://doi.org/10.1146/annurev.psych.121208.131647

Schultz, W. (2016). Dopamine reward prediction-error signalling: A two-component response. *Nature Reviews Neuroscience, 17*(3), 183–195. https://doi.org/10.1038/nrn.2015.26

Schultz, W. (this volume). Reward learning and choice: From theory to neuronal signals. In D. V. Smith, P. Lockwood, & D. S. Fareri (Eds.), *Neuroeconomics: Core topics and current directions*. Springer. Preprint: https://ssrn.com/abstract=4890719

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Terry, J. M., & Huettel, S. A. (this volume). Resolving uncertainty: Foundational neuroeconomic studies of risk and ambiguity. In D. V. Smith, P. Lockwood, & D. S. Fareri (Eds.), *Neuroeconomics: Core topics and current directions*. Springer.

Wu, Q., & O'Doherty, J. P. (this volume). Neural mechanisms of reward, decision-making and learning in a social context: Insights from computational neuroimaging. In D. V. Smith, P. Lockwood, & D. S. Fareri (Eds.), *Neuroeconomics: Core topics and current directions*. Springer. Preprint: https://doi.org/10.31219/osf.io/ncz4j_v2

Reward Learning and Choice: From Theory to Neuronal Signals

Wolfram Schultz

Abstract Rewards satisfy the needs of biological agents for survival and gene propagation that result in evolutionary fitness. To do so, rewards elicit learning (positive reinforcement), approach behaviour, economic choice and emotions like pleasure, happiness and desire. Dopamine neurons signal reward prediction errors (RPE) that are suitable for learning according to reinforcement learning theory, including Pavlovian learning and temporal difference learning. As biological needs vary between individual agents, the value of reward is subjective, which can be assessed using principles of economic choice theory. Theory-driven tests ensure that animals make meaningful choices for inferring subjective reward value. Dopamine neurons code RPEs in the subjective value formalized as utility. Concepts of Revealed Preference Theory allow to extend these tests to rewards that typically contain multiple components. Neuronal responses in monkey orbitofrontal cortex follow preferences for multi-component rewards, even when preferred rewards have smaller individual components. These responses reflect the transformation of vectorial multi-component reward into scalar subjective value. In sum, concepts from animal learning theory and economic choice theory inform experimental designs of neurophysiological reward experiments that transcend ad-hoc situations and validate, extend and ultimately reduce theories of reinforcement learning and economic choice.

1 Introduction

The investigation of neuronal reward signals during the past 30 years has substantially benefitted from concepts of animal learning and economic choice theories. Reward is a theoretical concept that explains learning, approach and choice. However, there are no specific sensory receptors for rewards that provide a hardware

W. Schultz (✉)
Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK
e-mail: ws42@pm.me

D. V. Smith et al. (eds.), *Neuroeconomics: Core Topics and Current Directions*,
https://doi.org/10.1007/978-3-032-02925-6_1

definition of their function. Rather, the function of rewards is defined by their value for the survival of humans and animals (collectively called agents henceforth) and the propagation of their genes. Thus, rewards serve evolutionary fitness by ensuring that agents survive and successfully propagate their genes into the next generation until that generation can propagate their genes again.

As the needs defining reward function differ between agents, reward value is subjective. Take the example of satiety; a pineapple loses value for an agent who has just eaten a pineapple, even though the pineapples may be physically exactly the same. And we have no way to compare subjective values between agents as there exists no inter-individual metric for subjective reward value. Thus, we need to address subjective reward value when we like to understand how the brain gets us the rewards we need.

While reinforcement learning serves reward acquisition by establishing predictions, economic choice theory provides estimation methods for subjective reward value. Already in 1738, Bernoulli postulated that the subjective gain from the same additional unit of money ('marginal utility') decreases with increasing wealth of the agent (Bernoulli, 1738). Since then, economic choice theory has become suitable for designing well-controlled behavioural animal studies for exploring neuronal signals for reward value and thus establish neuronal foundations for economic choice. This short review will present a few such examples from several laboratories, including the author's own laboratory, that have helped in our understanding of economic choice behaviour.

2 The Dopamine Reward Signal

Several mammalian brain structures contain neurons that process reward information distinct from sensory and movement processes. The structures include the dopamine neurons, whose large majority carries a rather specific signal reflecting reward prediction errors (RPE), and the orbitofrontal cortex, striatum and amygdala containing subsets of reward-sensitive neurons. Additional brain structures contain neurons that report specific aspects of rewards, including hypothalamic neurons signalling nutrients.

Reinforcement learning (RL) theory defines RPE as key variable for learning. Pavlovian learning of reward-predicting stimuli advances with RPEs: only rewards that are better or worse than predicted contribute to learning (Kamin, 1969; Rescorla & Wagner, 1972). A positive RPE, elicited by a reward that is better than predicted, increases the prediction of that reward, and a negative RPE, elicited by a reward that is worse than predicted, decreases the reward prediction (Fig. 1a, red vs. green). The process constitutes error-driven learning: adjusting the prediction via an error between the actually experienced reward and the prediction. However, RL is not an error-correcting mechanism, as it does not involve a fixed setpoint to which the system returns after an error has been corrected. Indeed, predictions in RL are basically unlimited: predictions can grow as long as rewards can increase enough to

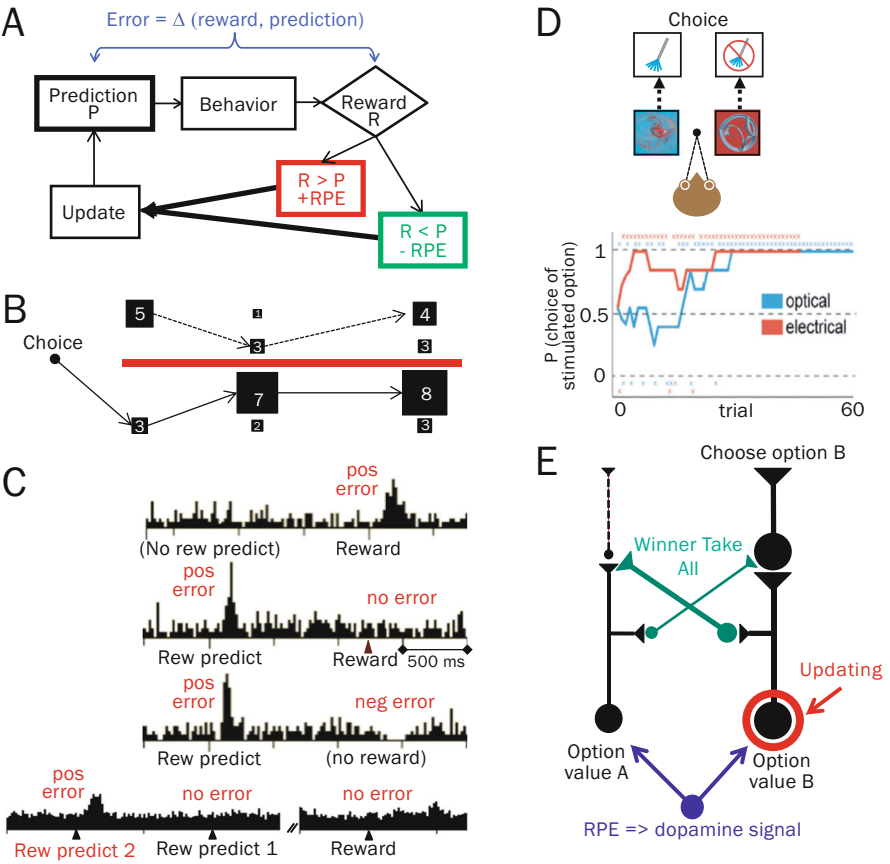

Fig. 1 Dopamine reward signal and reinforcement learning (RL) (**a**) RL block scheme. Positive reward prediction errors (+RPE: reward R exceeding prediction P) increase reward prediction, whereas negative RPEs decrease the prediction. Both predictions are unbounded, which distinguishes RL from error correction around a pre-determined setpoint (**b**) Maximization of state value rather than individual rewards using temporal difference (TD) learning of reward sequence. The agent learns the value of each primary or higher-order reward backwards after initial encounter of an ultimate primary reward. Overall state value is higher by foregoing an earlier larger reward for entering a sequence with later larger rewards (red line: sequence choice cannot be revisited). Numbers in black squares indicate values of primary and higher-order rewards (**c**) Dopamine RPE signal with reward and single or sequential conditioned stimuli (top three graphs, bottom graph). (Reproduced from Schultz et al., 1997 and Schultz, 1998) (**d**) Optogenetic stimulation of dopamine neurons in rhesus monkey. Repeated optical midbrain stimulation after choice of one option elicits increasing preference (P(choice)) of that option (blue fractal) to its alternative (red fractal). Midbrain electrical stimulation, directly opening voltage-gated sodium channels, has similar but stronger effects (bottom red). (Reproduced from Stauffer et al., 2016) (**e**) Winner-Take-All scheme of competitive choice between two options. See text for description

elicit positive RPEs; predictions can decrease as long as rewards can shrink enough to elicit negative RPEs.

Pavlovian learning provides the simplest and minimally confounded model of learning, but it is admittedly an abstraction of real-world learning processes. An RL model that conceptualizes real-world situations better is temporal difference (TD) learning. Agents are often several steps away from the ultimate reward but can be guided to that reward by a sequence of predictive stimuli that elicit specific choices. For example, a restaurant sign predicts a menu that predicts a specific meal that predicts specific foods that predict essential nutrients. The nutrients are the ultimate rewards, and the preceding stimuli are predictors of these rewards; by eliciting learning themselves, the predictors are higher-order rewards. Agents usually can choose at some or all steps between different predictive stimuli. At some steps, the choice may be between a better immediate option and a smaller option that predicts a later but much better option. In Fig. 1b, the chooser foregoes the best immediate option to obtain much larger rewards in the next steps (the red line blocks subsequent sequence change). Thus, TD learning conceptualizes how agents benefit from maximizing the overall reward value of a whole sequence, called state value, rather than maximizing only the immediate reward. Each reward-predicting stimulus in a TD sequence is learned using a basic Pavlovian process, starting with the last prediction before the ultimate reward and then advancing backwards in steps to the earliest stimulus; each predictive stimulus serves as higher-order reward that elicits learning of the preceding stimulus.

The dopamine response to rewards and reward-predicting stimuli in monkeys, rodents and fruit flies codes a bidirectional RPE (Fig. 1c) (Schultz et al., 1997; Cohen et al., 2012). An unpredicted reward at a given moment in time constitutes a better-than-predicted reward and elicits a positive temporal RPE and a dopamine excitation (top). A reward that occurs exactly as predicted and at the predicted time elicits no RPE and no dopamine response at all (middle high). A reward that fails to occur at its predicted time constitutes a worse-than-predicted reward and elicits a negative temporal RPE and a dopamine inhibition at the time at which the predicted reward would have occurred (middle low). Thus, dopamine neurons code a bidirectional temporal RPE. The bidirectional nature distinguishes the RPE from attention elicited by both positive and negative RPEs that would result in a common unidirectional neuronal response, usually an excitation. Further tests reveal that the dopamine response to reward-predicting stimuli itself constitutes also an RPE signal. Dopamine neurons are not only excited by a temporally unpredicted reward-predicting stimulus (left in both middle traces) but lose the response when that reward-predicting stimulus is predicted in time by an earlier predictive stimulus (bottom) (Schultz et al., 1993). Thus, the dopamine RPE response occurs with temporally unpredicted RPEs elicited by both ultimate rewards and reward-predicting stimuli. This feature resembles key characteristics of the TD RL model in which learning is mediated by RPEs at each rewarding event (Montague et al., 1996). Thus, the phasic dopamine response signals reward, and the form of the signal is not reward itself but the difference between reward and prediction across their respective time steps (Δr-p or dr/dt).

Besides the RPE signal, dopamine neurons also show a brief earlier unselective response to any event that reflects the novelty, sensory and motivational salience of the event (Schultz, 2016). Dopamine neurons are also activated with behavioural activation from movements of sufficient amplitude, speed or acceleration (Schultz, 2007; Cox & Witten, 2019). These multiple functions of phasic dopamine changes seem confusing when assuming that one brain system has only one function. However, pluripotent functions of individual neurons may be advantageous for an evolutionary ancient brain system.

The dopamine RPE signal registers rewards not only passively but also drives behaviour actively towards reward. Work in the laboratory of Roy Wise (Corbett & Wise, 1980) suggested that dopamine neurons are a major substrate for the intracranial electrical self-stimulation discovered earlier (Olds & Milner, 1954). Recent optogenetic experiments have confirmed that excitation of dopamine neurons drives monkeys and rodents towards behaviours that elicit dopamine excitation again (Tsai et al., 2009; Stauffer et al., 2016), for example, during choices (Fig. 1d). Correspondingly, inhibition of dopamine neurons drives animals away from behaviours eliciting dopamine inhibition (Tan et al., 2012). The artificially elicited dopamine excitations and inhibitions mimic positive and negative dopamine RPE signals. The observations suggest that dopamine RPE signals are causally sufficient for driving RL. It appears that animals seek better-than-predicted and avoid worse-than-predicted rewards in order to experience positive dopamine RPE signals and avoid negative RPE signals, although only work on humans may tell us about explicit intentional stance.

While a function in RL may seem abstract, the neuronal reinforcing effects of dopamine signals may consist of updating reward value in downstream neurons involved in economic choice. In a typical competitive Winner-Take-All model of binary choice, the stronger of two value inputs exerts stronger excitation on downstream neurons and stronger inhibition on neurons for the alternative option, as compared to the option with the weaker value input (Fig. 1e). A subsequent threshold mechanism assures that only the stronger of the two inputs reaches the output, and thus determines the choice, whereas the weaker input of the alternative option dissipates without consequence. Circuits implementing lateral inhibition are found in the striatum (caudate nucleus, putamen, nucleus accumbens) and orbitofrontal cortex (OFC). Dopamine release elicited by the RPE from the received reward is indiscriminately directed at basically all postsynaptic striatal neurons but only updates the option value at synapses whose previous activity has left an eligibility trace for modification (Fig. 1e red). Thus, according to this decision model, the widespread dopamine RPE signal updates the value in those neurons that contributed to the preceding rewarded choice, constituting a kind of neuronal RL mechanism.

3 Dopamine Neuroeconomics

Satiety makes a good case for the notion that rewards have value for satisfying the needs of individual agents; therefore, reward value is subjective. A candidate brain system for updating economic decision variables should code reward value on a subjective basis without requiring additional downstream mechanisms that translate objective into subjective value. Another case for subjective value is temporal discounting; the value of a later reward is lower than the value of an earlier but physically identical reward. While satiety and temporal distance are useful ad-hoc situations for identifying subjective value and its neuronal correlates, mathematically defined utility (Bernoulli, 1738) provides a metric for subjective value, defined as a typically nonlinear function of reward amount (Fig. 2a). A further basic determinant of subjective value is weighted reward probability (Kahneman & Tversky, 1979) that has recently been found to affect neurons in orbitofrontal cortex (Ferrari-Toniolo & Schultz, 2023) but will not be further elaborated here. While subjective reward value is determined by many other factors, utility and weighted probability are simple enough for testing in animals suitable for neurophysiological studies on reward.

An experimentally viable way to estimate utility functions employs choice under risk. Risk in economics concerns both gains and losses and is defined by the higher moments of reward distributions, namely variance, skewness and kurtosis. By contrast, the often-used definition of risk as probability of damage or loss fails to include gains that are usually the reason for engaging in risky choice. Intuition often seems to combine these two definitions.

Risk is conveniently investigated during choices involving two options. To avoid confounds and be well understood by animals and interpretable by experimenters, tasks should be simple and have straightforward parameters (notwithstanding the higher complexity of real-world choices):

- Options are distinct and well separated.
- Options are collectively exhaustive: all options are included in the presented option set.
- Options appear simultaneously: there are no intervening stimuli that may distract, require additional memory, or cause adaptation.
- Options are mutually exclusive: only one option can be chosen.
- Individual options may have two mutually exclusive reward amounts; only one amount is paid out with a specific probability (all reward probabilities within an option add to $P = 1.0$).
- The cost for obtaining an option remains constant throughout the test (but cost can differ between options).
- The reward in any option does not depend on other options.
- The reward of the chosen option is paid out immediately.

An important issue with animals is to assure they make meaningful choices. Three simple conditions can test this requirement. The first test for meaningful

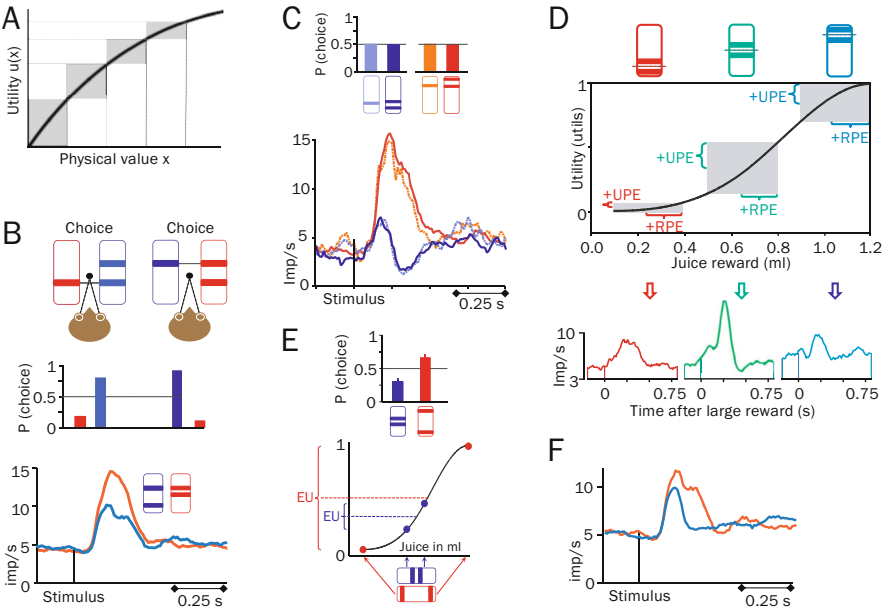

Fig. 2 Neuroeconomics of the dopamine reward signal (**a**) Bernoulli utility function. Decreasing height of grey squares indicates gradually declining marginal utility (du/dx) as a hallmark of subjective reward valuation (**b**) Assessment of meaningful choice with first-order stochastic dominance (FOSD). Top: choice between safe and risky rewards. Line height within each vertical rectangle indicates reward amount, higher is more; two lines inside the same rectangle indicate mutually exclusive reward amounts, each delivered with probability of $P = 0.5$. Blue options are better. Middle: monkey shows stochastic preference for better options, thus satisfying FOSD. Bottom: dopamine response to stimuli predicting the options is stronger for better reward (red) (**c**) 'Utility maximization' during choice between risky and safe rewards. At choice indifference (choice $P = 0.5$ of each of two options), subjective value is by definition the same in both options. Relative to the mean amount of a given gamble option, the safe reward is higher with small rewards (top left) but lower with large rewards (top right). Thus, the animal chooses the smaller amount in half the trials, indicating maximization of utility rather than amount. Bottom: the reward-sensitive dopamine neurons (red vs. blue) show the same response to each option of a given option set (blue and red), suggesting coding of subjective value rather than physical amount (**d**) Dopamine RPE signal codes utility. Top: three equiprobable gambles with same variance projected onto nonlinear utility function. The positive RPE of each gamble is identical (x-axis) but changes into varying positive utility prediction errors (UPE; y-axis) due to utility nonlinearity. Bottom: dopamine RPE signal varies with UPE, not physical RPE (**e**) Testing second-order stochastic dominance (SOSD) with mean-preserving spread. Despite same mean physical amounts, projection of the higher-variance gamble from the x-axis onto the y-axis results in higher utility (red) for the animal and thus predicts preference for that gamble, which is indeed the case (top) (**f**) Dopamine signal with SOSD. Despite same mean amounts, the higher-variance gamble elicits a stronger dopamine response compared to the lower-variance gamble shown in panel **e**, confirming neuronal coding of subjective value rather than physical amount. (Panels **b**, **d**, **e**, and **f** reproduced from Stauffer et al., 2014)

choice involves first-order stochastic dominance, which defines the better option: every option is at least as good as its alternative and better in at least one instance. Figure 2b (top) shows how reward amount in each option is predicted by the vertical position of a bar within a rectangle (higher is more; two bars indicate split amounts with $p = 0.5$ each); blue options are better and thus 'dominate' the option set. Monkeys choose the better options stochastically, with a few exploratory choices of the worse, 'dominated' option (Fig. 2b middle). Dopamine neurons respond strongly to the stimuli predicting the better option (Fig. 2b bottom). Thus, monkeys recognize what is better for them, and dopamine responses reflect these values (the test does not distinguish objective from subjective value).

The second test for meaningful choice involves maximization of utility. In psychophysically controlled choices between safe and risky rewards, the animal chooses both options equally frequently at some point (choice indifference) (Fig. 2c). As we infer subjective value from observable choice, both options have the same subjective value for the animal at choice indifference. However, with lower amounts, the safe reward amount exceeds the mean reward of the gamble at choice indifference (blue), whereas the converse is true with higher amounts (red). Thus, by choosing each option equally frequently at choice indifference, the animal chooses in half the trials the smaller reward amount (the gamble at lower amounts, and the safe reward at higher amounts; blue and red). It foregoes the larger amount on these trials as the smaller amount has the same subjective value. Thus, monkeys maximize subjective value (informal utility) rather than objective amount in a meaningful way. Dopamine responses show very similar responses to both options at choice indifference (while their increase with higher amounts attests to their sensitivity for value) (Fig. 2c bottom). Thus, dopamine neurons seem to code subjective value in this specific test.

An empirically estimated utility function would allow to test neuronal coding of subjective value on a more general basis. Utility functions can be estimated from risky choice using the so-called fractile procedure that employs a chain of indifference points with precisely defined utility, starting at 0.5 utils and working towards increasingly finer resolution. Utility functions are fitted from choices using splines or nonlinear functions. The obtained utility functions in monkeys are convex for small reward amounts and with increasing amounts become gradually more linear and finally concave (Fig. 2d, black curve) (Stauffer et al., 2014). Choices of monkeys between safe rewards analysed with the random utility model reveal similarly S-shaped utility functions with somewhat less nonlinearity (Bujold et al., 2022). For humans, the S-shaped pattern is referred to as gambling for peanuts (Weber & Chapman, 2005). When projecting three gambles with identical variance but different means onto different parts of this utility function, the positive RPEs (top amount minus mean; x-axis) are identical, whereas the positive utility prediction errors (UPE; y-axis) vary with the steepness of the utility function (Fig. 2d, grey rectangles). The varying dopamine responses to the larger reward follow the varying positive UPEs rather than the constant positive physical RPEs (Fig. 2d, bottom) (Stauffer et al., 2014), thus demonstrating RPE coding in the metric of subjective utility rather than objective physical reward amount.

The demonstration of dopamine utility coding allows to use second-order stochastic dominance as a third test for meaningful choice. The test uses the so-called mean-preserving spread that provides a minimal-confound definition of economic risk and tests risk sensitivity: two gambles with equal mean objective value differ only in the variance of their amounts (Rothschild & Stiglitz, 1970). A risk avoider would prefer the less risky gamble (smaller variance), whereas a risk taker would prefer the riskier gamble (larger variance). Monkeys prefer the riskier option (Fig. 2e top), thus confirming risk-taking with low amounts (Fig. 2c blue at top left). When projecting the two gambles onto the empirically estimated utility function for that monkey, the mean subjective value (expected utility, EU) of the riskier gamble exceeds the mean subjective value of the less risky gamble despite the same physical mean amount (Fig. 2e bottom), confirming dissociation between subjective and objective value. Thus, true to the predictive nature of mathematical functions, the utility function empirically estimated from measured choices predicts risk-taking choice (note that the test gambles differed from the gambles used for estimating the utility function and thus served also as out-of-sample test for the validity of the utility function). The dopamine responses to the two gambles followed the subjective value rather than the objective value, thus confirming nonlinear utility coding in this test (Fig. 2f). Although the dopamine response increases with risk, it is not a risk signal but a subjective value signal boosted by risk. Analogous utility prediction error signals are found in the human dopamine-receiving ventral striatum using functional magnetic resonance imaging (Konova et al., 2023). True risk signals, distinct from value (and not coding prediction errors), are found in monkey orbitofrontal cortex (O'Neill & Schultz, 2010).

Taken together, the meaningful choices under risk in rhesus monkeys demonstrate that subjective reward value, expressed as formal utility, is a key decision variable. It is coded by the RPE signal of dopamine neurons in the metric of formal utility, strictly spoken as UPE. The identification of subjective value coding would allow dopamine signals to participate in the updating of value signals for economic choice as proposed in Fig. 1e.

4 Neuronal Value Signals for Multi-component Rewards

Rewards are typically composed of several components. An apple has taste and size, a sausage contains calories and fat, and a coffee provides caffeine and water (and milk and sugar and maybe a few other ingredients). Each of these components has its own value that varies along a single dimension. A single-dimensional value is either low or high or somewhere in between and is therefore called a scalar. Rewards that have several components also have several values, and the value of each component is a distinct scalar. Thus, multi-component rewards are vectors composed of several scalars.

Even when rewards have multiple components, agents choose them only more or less along a single dimension, as if the value of even these rewards is a scalar.

Reward signals of individual neurons also vary along a single dimension; they can only be larger or smaller and thus are scalars just like reward value. The common scalar nature of reward value and neuronal reward signals encourages to ask the question how the single-dimensional scalar neuronal responses can reflect multi-component, vectorial rewards. Help comes from Revealed Preference Theory in economics (Samuelson, 1937) that derives scalar value from vectorial rewards. Once we understand that step, we can explore how such scalar value is encoded as scalar neuronal signals.

A convenient graphic demonstrates a solution for obtaining scalar subjective value from vectorial rewards. The scalar amounts of individual bundle components are plotted onto a two-dimensional x-y graph, and the positions of equally preferred bundles are connected as curves (Fig. 3a red, orange, green, purple). Due to their equal preference, all bundles on each curve have the same subjective value, and bundles on higher curves (farther away from the origin) are preferred to, and thus have higher subjective value than, bundles on lower curves. Note that the reward values are inferred from choice and therefore are subjective by definition, despite being set in the metric of physical amounts.

In the experiment, a monkey chooses between two visually similar two-component reward options ('bundles') (Fig. 3b). Each of the two bundles contains the same two reward types, like grape juice and blackcurrant juice, with individually set amounts (Fig. 3b) (Pastor-Bernier et al., 2017). In this stochastic design, individual bundle amounts are varied psychophysically until each bundle is chosen equally frequently (choice indifference). Systematic variation of initial bundle settings allows to estimate several indifference points that can be fitted by an indifference curve. Further variation of bundle settings results in different such indifference curves that can then be combined into an indifference map. The map demonstrates the vectorial-to-scalar transformation of value: same scalar value along curves, different scalar value across curves (Fig. 3c).

Now that bundles have integrated scalar value from both reward components, we can ask whether scalar neuronal reward responses can code such scalar values. This is indeed the case for a substantial population of neurons in the monkey orbitofrontal cortex (OFC). As Fig. 3d shows, the bundle with the highest value, on the higher indifference curve, elicits the strongest neuronal response (orange), whereas the two lower-valued bundles, on the lower indifference curve, elicit smaller responses (red, green). Note that the two lower-valued bundles, being positioned on the same curve, have similar value among them; accordingly, their neuronal responses fail to differ much from each other. Interestingly, when comparing across curves, the higher-valued bundle on the higher curve (Fig. 3c orange) elicits the stronger neuronal response, even though it contains less grape juice and less blackcurrant juice than the lower-valued bundles (red, green) on the lower indifference curve; it is the combined juice amounts that determine the integrated scalar value of the bundles and their neuronal responses (Pastor-Bernier et al., 2019). Thus, it seems that specific OFC reward neurons code the integrated scalar value of vectorial, multi-component rewards.

Index

D. V. Smith et al. (eds.), *Neuroeconomics: Core Topics and Current Directions*,
https://doi.org/10.1007/978-3-032-02925-6