The phasic dopamine signal maturing: from reward via behavioural activation to formal economic utility

Wolfram Schultz[1], Wiliam R. Stauffer[2], Armin Lak[3]

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

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

[3]Institute of Ophthalmology, University College London, 11-43 Bath Street, London EC1V 9EL, UK

Corresponding author:
Wolfram Schultz
Department of Physiology, Development and Neuroscience
University of Cambridge
Cambridge CB2 3DY
United Kingdom
Email: ws234@cam.ac.uk
Phone: +44-1223-333 779
Fax: +44-1223-333 840

Highlights:
• Two-component dopamine response structure explains lack of activation by aversiveness.
• Other phasic dopamine changes reflect behavioural activation.
• Dopamine reward prediction error signalling codes formal economic utility.
• Optogenetic dopamine activation affects learning and economic choices of primates.

Short title:
Phasic dopamine signals

Abstract
The phasic dopamine reward prediction error response is a major brain signal underlying learning, approach and decision making. This dopamine response consists of two components that reflect, initially, stimulus detection from physical impact and, subsequenttly, reward valuation; dopamine activations by punishers reflect physical impact rather than aversiveness. The dopamine reward signal is distinct from earlier reported and recently confirmed phasic changes with behavioural activation. Optogenetic activation of dopamine neurones in monkeys causes value learning and biases economic choices. The dopamine reward signal conforms to formal economic utility and thus constitutes a utility prediction error signal. In these combined ways, the dopamine reward prediction error signal constitutes a potential neuronal substrate for the crucial economic decision variable of utility.

Introduction

Neuronal signals for reward prediction errors (RPE) are reaching adulthood. For more than 21 years now, empirical studies demonstrate activations (increase in activity) in midbrain dopamine neurones with unpredicted reward and depressions with reward omission that conform to the formalism of reward prediction error known from animal learning theory [1-6]. (The term 'error' is meant to indicate a discrepancy between actual and predicted outcome, not wrong behaviour; even 'unpredicted' rewards occur on the basis of some, often poorly defined, prediction from the animal's past experience, and thus can be accomodated into the error concept). During these juvenile years, their robustness has been tested and necessary amendments were made [7-12]. Now the dopamine RPE signal should show signs of sophistication and perspective. This text describes four recent developments, namely a two-component signal structure that also addresses relationships to salience and aversiveness, a distinction from recently confirmed phasic changes reflecting general motor and sensory processes leading to behavioural activation, optogenetic dopamine stimulation in primates, and formal utility coding that builds the bridge to economic goods. This review is admittedly a short, personal selection, and we apologise to the many authors of recent dopamine studies whose work we will not have the space to discuss.

Two-component phasic dopamine responses

Similar to neuronal responses in other brain systems, the midbrain dopamine RPE signal consists of several components. A short-latency, initial activation (increase in neuronal activity) is unselective and detects any event of sufficient intensity; a subsequent activation or depression codes a positive or negative error, respectively, in the prediction of reward or reward-predicting stimulus as soon as the animal has evaluated the event (as evidenced by behavioural reactions) [13•]. The transition from unselective event detection to value processing is likely to emerge gradually via intermediate steps, including stimulus identification and comparison; the transition appears to be more distinct with more demanding stimuli taking longer to identify and evaluate [14]. The initial component is enhanced by physical stimulus impact, reward context, similarity to reward (generalisation), and novelty [15] (Figure 1). In their capacity to generate stimulus-driven attention, these four factors confer physical, motivational and novelty/surprise salience, respectively. However, salience concerns only the first component and transitions rapidly to ultimate reward coding in the second component. The initial component follows also the formality of temporal prediction error coding [14] and thus forms an integral part of the dopamine RPE signal compatible with temporal difference reinforcement learning [16]. Despite its unselective nature, the initial component may provide advantages; its short latency allows early preparatory neuronal processing for faster reward acquisition while allowing cancellation of behaviour if the object turns out not to be a reward; the salience would enhance subsequent neuronal processing [17] and facilitate behavioural learning according to attentional learning theories [18]. The observation that the ultimate reward elicits a graded RPE suggests that the prediction from the second, value component lasts well beyond the behavioural action [14]. Due to the early onset of the lasting value component, the animal would not be confused and be able to perform the action according to the signalled value.

The issue of aversive dopamine activation

The two-component structure of the dopamine RPE response may explain why dopamine research for > 30 years had consistently reported dopamine activations to aversive stimuli that now turn out to result probably from the physical impact of the punisher. Starting with Chiodo et al. [19], followed by ourselves [20,21] and confirmed by others [22-25], subsets of dopamine neurones are activated by airpuffs and loud tones. A recent study used several punishers, various physical intensities, psychophysically assessment of aversive values and multiple linear regression analysis to demonstrate that the ‘aversive’ dopamine activation seems to concern the initial dopamine response component and

reflect the physical intensity of punishers rather than their negative aversive value [26•,27]. None of the investigated dopamine neurones so far show an unanimous true aversive activation in such well controlled conditions.

The physical impact attribution to 'aversive' dopamine activations may affect the interpretation of aversiveness in dopamine responses. First, dopamine activations by reward are reduced by an added aversive liquid [27], which follows an economic model of net benefit being the sum of reward utility and punisher disutility. Second, dopamine activation with aversive stimuli may signal punishment relief, as voltammetric striatal dopamine decreases with experienced aversive footshock but increases with successfully avoided footshock [28•]. Third, habenula stimulation induces place avoidance (staying away due to signalled aversiveness) or place dispreference (reduced preference due to depressed reward signalling) [29,30,31], either by disynaptically inhibiting reward-activated dopamine neurones and or by monosynaptically exciting supposedly punisher-activated dopamine neurones. An interpretation towards aversive dopamine activation seen with amazing molecular methods [31] may be constrained by difficulties in empirically distinguishing place avoidance from place dispreference, by dopamine activation shown by Fos possibly arising from rebound excitation of dopamine neurones inhibited by habenula stimulation [23,30,32-34], by excitatory postsynaptic currents possibly not leading to propagated action potentials, and by blunting from dopamine antagonists not distinguishing between dopamine excitation and inhibition (although regrettably misstated [13,34], the dopaminergic nature of the neurones tested with Fos expression and whole-cell recording were identified by tyrosine hydroxylase immunostaining [31]).

Taken together, the typical dopamine response to punishers seems to be a depression of activity that reflects the negative value RPE. Of course, despite all arguments and reservations, the existence of some truly aversively activated dopamine neurones can not be completely excluded, but the potential confound of physical impact should be addressed experimentally.

Homogeneous RPE responses of heterogeneous dopamine neurones

Truly aversive activations in subgroups of reward-processing dopamine neurones would indicate categorical heterogeneity of dopamine RPE signalling. However, the attribution of 'aversive' activations to the initial, physical response component [27,35] rather suggests graded sensitivities to physical impact, conforming to one single distribution rather than to two statistically distinct populations. This more nuanced view contrasts with the notion of categorically different dopamine populations suggested before the two-component nature became recognised [23]. The notion of rather homogeneous RPE signalling is supported by the stereotyped computational subtraction of received minus predicted reward with correspondingly scaled sensitivities to positive and negative RPEs [35,36,37••], and the high noise correlation and synchrony between dopamine responses [37••,38]. As a functional consequence, if dopamine responses vary along a single continuum, each dopamine neurone contains the full RPE information, and single or small groups of dopamine neurones effectively convey a full RPE signal to postsynaptic neurones, without requiring summed population activity through overlapping projections [37••].

The issue is not whether dopamine neurones are heterogeneous or not; like all neurones, they differ in many aspects including cytology, input source, co-neurotransmitter, presynaptic interaction, reuptake transporter, projection territory, and pre- and postsynaptic receptor location on heterogeneous postsynaptic neurones. Additional, non-RPE activity changes reflecting more global phenomena such as behavioural activation (see below) are also heterogeneous when considering their time course, polarity, event relationship and proportion of neurones engaged. By contrast, what seems to be rather homogeneous and stereotyped is the RPE response, which varies in a graded rather than categorically

distinct manner across the otherwise heterogeneous dopamine neurones. The similarity in dopamine RPE responses contrasts with the classically well separated distributions of heterogeneous neuronal responses to distinct task events in prefrontal cortex, striatum, amygdala and most other brain structures [39-41].

Phasic non-RPE dopamine changes

Given the notion that dopamine neurotransmission acts at different time courses [42,43] and the renewed interest in non-RPE signals [44-49], it might be worth considering the initial studies of non-RPE dopamine signals. During large arm reaching movements toward food boxes, which engage a large proportion of arm and eye muscles in monkeys, select subpopulations of dopamine neurones are activated (31-44%) or depressed (15-17%) [50,51] (Figure 2a, b), sometimes distinguishing between contra- and ipsilateral reaching [50]. Similar activations occur when these movements occur spontaneously (12%) in a design that addressed the most impaired movement type in Parkinsonism [52]. Mouth movements are also accompanied by activations (9%) and depressions (1%) [51] (Figure 2c). Further, some dopamine neurones show slow activations (5%) or depressions (1%) lasting over whole trial durations of several seconds [51] (Figure 2d). Besides the later identified RPE responses, this variety gives rise to impressive schemes of heterogeneous activities that likely reflect motor activation, sensory stimulation, and general behavioural reactivity [50-52] (Figure 2e). However, these activities lack consistency, reproducibility and coherent functional interpretation. Clear movement relationships are absent in dopamine neurones of monkeys performing in Pavlovian or simple operant tasks that separate well sensory from motor events and engage only limited numbers of muscles; thus, well controlled reaching to levers, precise elbow flexion-extensions and ocular saccades consistently fail to activate dopamine neurones in monkeys [53-55] (Figure 2f-h), as Pavlovian licking fails to do in mice [12]. Specific analysis identifies RPE coding as what initially appears to reflect movement during spatial delayed-alternation and -response [1,55], cognition during delayed matching-to-sample [56], subjective perception during signal detection [11], and complex behaviour during multistep choices [10]. So far, the only known slower activation in simple tasks reflects reward risk [57]. Thus, non-RPE dopamine changes occur with general behavioural activation processes involving many muscles and sensory receptors, but less so with well-controlled movements involving fewer muscles.

Optogenetics is shifting behavioural neurophysiology towards rodents. Many operant tasks used on rodents engage large fractions of the body's musculature and sensory receptors. Neurophysiological, voltammetric and optical imaging studies reveal fast and slow changes in rodent midbrain dopamine neurones and striatal dopamine axons. Global motor activity, stimuli and reward delivery during T-maze navigation, whole-body turns, locomotion, wheel running and nose pokes are associated with dopamine activations [44•,45,47,48], depressions [49] or both [46] (Figure 2i-k). Compatible with these tasks engaging major body musculature and somatosensory receptors, some of these activities vary parametrically with locomotion velocity and acceleration [46,47] and differ between ipsi- and contralateral movements in some striatal regions [48]. These studies confirm the dopamine changes with behavioural activation in monkeys and, in addition, describe optogenetic stimulation effects on locomotion speed [46] and willingness for operant performance [58] reflecting reward value. Most of these rodent studies demonstrate also RPE coding, although the many overlapping sensory and motor components in these complex behaviours render its precise identification difficult (as previously in monkeys [50,53]). More simple Pavlovian or operant tasks comprise well-separated stimuli and well-controlled movements engaging limited numbers of muscles; accompanied by sophisticated data analysis, such tasks facilitate the assessment of reward prediction and identification of RPEs [36,37••] and, as in monkeys [53-55], do not seem to yield much non-RPE activities.

The question arises whether an absence of dopamine changes reflecting behavioural activation might provide an explanation for Parkinsonian akinesia. The notion would less easily apply to movement-related depression [50] that is associated with striatal dopamine reduction in normal rats [49], as Parkinsonian akinesia is already associated with reduced striatal dopamine; also, electrical or optogenetic dopamine activation elicits behavioural learning and approach [59,60,61•], whereas optogenetically induced depression induces behavioural dispreference [48,62]. But even for activations, dopamine receptor agonists alleviating Parkinsonian akinesia cannot reinstate phasic neuronal changes, thus making absent dopamine activations an unlikely explanation for the akinesia. Perhaps dopamine agonists would act by boosting the influence of dopamine released by the few remaining dopamine neurones with behavioural activation?

From biological rewards to economic goods

Biological rewards derive their function from the body's need to acquire substances for correcting and preventing homeostatic challenges. This notion applies to basic liquid and food rewards that contain such substances and to other, non-nutrient stimuli, objects and events facilitating the acquisition of basic rewards. As not all biological agents have access to the same, optimal rewards, they benefit from exchanging plentiful objects against objects they may be lacking. The need for exchange has expanded the survival function of biological reward to the notion of economic good enhancing the chance for survival and ultimately evolutionary fitness. Consequently, a more thorough and representative investigation of neuronal reward function should incorporate concepts of economic decision theory.

The survival value of rewards is determined by the subjective needs of the agent rather than by physical and chemical factors alone. The value of a third Porterhouse steak is lower than the value of the identical first one; this is why all-you-can-eat restaurants can stay in business. Populations of midbrain dopamine neurones, and voltammetrically assessed striatal dopamine concentrations, code reward value on a subjective rather than a physical basis, as seen with delays (temporal discounting) [63,64], different reward types [65], risk [65], salt depletion [66••] and effort [64,67,68]. Human striatal voltammetric dopamine changes reflect RPEs combined from obtained and foregone (counterfactual) reward, rather than from obtained reward alone [69]. Thus, dopamine RPE signals reflect subjective value in a variety of scenarios.

Although temporal delay, satiety, risk and effort contribute to subjective reward value, only formal economic utility is normative for the choices individuals undertake for maximising utility. Utility represents, on an internal scale (utils), the subjective reward value as a mathematical function of objective value. Once estimated from specific choices, a utility function predicts reward-maximising choices. As neuronal activity is quasi-continuous and numeric, a utility function processed by neurones should have similar properties. Such 'cardinal' utility functions can be estimated under risk [70••], using the fractile, chaining procedure with a specifically structured series of psychophysically controlled choices between an adjustable safe reward and a fixed binary, equiprobable gamble defined by variance-risk [71]. The certainty equivalent (CE) is defined by the safe reward amount at choice indifference (50% choice of each option) (Figure 3a, top). A CE above the Expected Value of the gamble (EV; summed product of amount and probability) indicates a higher subjective value of the gamble compared to the safe reward of equal amount, and thus variance-risk seeking; by contrast, a CE below EV suggests variance-risk avoidance. Our rhesus monkeys are variance-risk seeking with small rewards but variance-risk avoiding with larger rewards (Figure 3a, bottom) [72•]. Correspondingly, their utility function is initially convex and then linear and concave with larger rewards (Figure 3b, red) [72•]. Thus, the same, single utility function of a monkey independently predicts variance-risk seeking and avoidance with different reward amounts, positive skewness-risk seeking and negative skewness-risk avoidance [73•]. Importantly, dopamine responses to unpredicted rewards generating positive

RPEs show a similar, nonlinear, convex-concave amount-response function (Figure 3b, black bars), rather than a linear increase in physical amount (dotted line). RPE responses to the outcomes of well defined, binary, equiprobable gambles follow the utility function in a similar way, being stronger in the steeper part of the function and weaker in its flatter parts despite identical variance [72•]. Thus, the dopamine RPE response constitutes a utility prediction error signal and as such represents a physical implementation of the theoretical construct of formal economic utility.

Optogenetics: of mice and monkeys

The optogenetic excitation of midbrain dopamine neurones or striatal dopamine axons induces learning of behavioural tasks in rodents, such as place preference, nose pokes, lever pressing and unblocking of learning a stimulus not associated with a RPE [58,61•,74-78], compatible with simple behavioural learning deficits with reduced dopamine burst firing in NMDA receptor knock-out mice [79]. Optogenetic dopamine activation also leads to immediate behavioural approach, higher motivation (more responses, shorter latencies), whole-body rotation, locomotion, treadmill movement and reward choice [47,58,77,80]. To the contrary, optogenetic inhibition of dopamine neurones, or excitation of synaptically upstream GABA neurones, induces learning of place and choice dispreference (reduced preference for place or choice option) and disrupts reward consumption [48,62,78,81]. Optogenetic dopamine stimulation may exert its behavioural effects by mimicking dopamine RPE signals, particularly when using similar stimulation parameters as natural dopamine RPE responses (number of impulses, instantaneous frequency, duration) [75,77]. Optogenetic dopamine stimulation might also mimick non-RPE activations, as phasic stimulations resembling natural dopamine activations during bouts of treadmill locomotion induce such movements [47]; however, the distinction from mimicking RPEs is difficult given the complexity of rodent behaviour. Much longer (1 sec) stimulations not mimicking phasic RPE signals fail to induce place preference learning [60]. Together, these studies demonstrate a function of phasic dopamine activity in inducing behavioural outputs.

Compared to rodents, monkeys have higher perceptual capabilities, can more easily perform well-controlled and well-separated individual movements involving interpretable numbers of muscles, allow better distinction between movement and sensory processes, and have higher cognitive abilities that facilitate the expression of distinct behaviours, such as specific attitudes to different, mathematically characterised forms of economic risk [73]; work on these closer relatives of humans also increases the relevance for human brain disorders. Initial monkey studies infected cortical and subcortical neurones with light-sensitive excitatory and inhibitory ion channels, distinguishing neurones from glia [82-90]. Current work aims for specific neuronal types. A two-vector approach labels midbrain dopamine neurones in wild-type monkeys with Cre recombinase coupled to a tyrosine hydroxylase promotor and inserts channelrhodopsin 2 dependent on the expressed Cre-recombinase [91]. With >35% infection efficacy and >95% specificity, laser stimulation induces impulses in dopamine neurones, but not in other neurones in the area, usually in a < 1:1 ratio. Laser stimulation simultaneously with unpredicted juice reward enhances the dopamine responses to the reward itself, makes an arbitrary visual stimulus acquire reward-predicting properties and dopamine responses, and gradually biases the animal's oculomotor choices towards the reward-cum-laser option and away from the reward-only option (Figure 3c, d), thus validating the notion of dopamine-mediated neuronal and behavioural reinforcement in monkeys. These optogenetic behavioural effects are consistent with those induced by electrical midbrain stimulation in monkeys [91-92]. Thus, a first step has been done to transfer neurone-type specific optogenetic techniques to monkeys.

The question arises by which output route dopamine stimulation may exert its long recognised behavioural functions [59]. Among the many projection territories, the dorsal and ventral striatum receive the densest dopamine innervation. Schematically, the dopamine RPE signal (Figure 4a) acts on

different medium spiny neurone classes in dorsal but not ventral striatum [93], via D1-type receptors (motor-excitatory 'direct' pathway to output) or via D2-type receptors (motor-inhibitory 'indirect' pathway). Optogenetic excitation of D1- and D2-receptor-expressing striatal neurones induces generally opposite learning and immediate behavioural effects. The opposing learning effects consist of increases and decreases of self-stimulation [94] (Figure 4b) and of changes of bimanual joystick movement speed [95]. The opposing immediate effects consist of biasing nose poke target choices for reward towards contralateral and ipsilateral body sides, respectively [96] (Figure 4c), and of exciting and inhibiting the basal ganglia, thalamus and cortex [97]. It is interesting to find such straightforward behavioural effects when stimulating striatal neurone populations with stereotyped protocols that cannot incorporate the heterogenous patterns of striatal neurones activitated with various sensory and motor task components.

Conclusions

Recent advances in the behavioural neurophysiology of the phasic dopamine signal seem to follow both a forward path and a full circle. The forward path concerns the development from reward response via neuronal reward prediction error signal to the characterisation of neuronal reward coding of subjective value and formal economic utility. Subjective, rather than objective, neuronal value coding matches intuitively the subjective aspects of reward. Maybe more importantly, the neuronal coding of utility signifies the conceptual step from biological reward, required for survival, to economic good, allowing welfare enhancement by exchange; that exchange value is based on utility.

The full circle concerns changes of dopamine activity with large reaching movements of the arm that are, embarrasingly, not replicated with fine, precise and well-controlled movements in monkeys. The result of the well-controlled studies is that the notion of a clear movement relationship of phasic dopamine activity cannot be maintained. However, in an apparent full circle swing, recent work on rodents replicates the dopamine changes during movements of the whole body or of large parts of the body; these tasks engage considerable proportions of the body's muscles and sensory receptors. As a result, it appears that the dopamine changes during large movements probably concern general motor activation, behavioural activation, or sensory stimulation derived from moving, all of which is difficult to distinguish from specific motor control. There is still no clear evidence of phasic dopamine changes with precise, well-controlled movements of concise numbers of muscles and devoid of reasonable confounds.

Acknowledgements

Our work has been supported by the Wellcome Trust (095495, 106101), European Research Council (ERC, 293549), and NIH Conte Center at Caltech (P50MH094258).

Legends

Figure 1. Two response components of dopamine neurones. During the learning of visual stimuli predicting probabilistic rewards, the initial component reflects stimulus novelty, as it decreases with repetition of the identical stimulus but does not distinguish between different reward probabilities (top). The second, subsequent component codes the reward value of the stimulus (bottom); across successive learning trials, this component increasingly discriminates between the three stimuli predicting reward with different probability (p=0.25, p=0.5, p=0.75). Not shown is the conclusion from other experiments, suggesting that the initial activation reflects also physical impact, similarity to rewarded stimuli, and reward context [13]. Reprinted with permission from [15].

Figure 2. Fast and slow dopamine changes not coding reward prediction error.
(a, b) Activation and depression, respectively, of impulse activity with large arm reaching movements of monkey engaging > 40 muslces. Reprinted with permission from [50].
(c) Slow activation with reward consumption and associated mouth movements in monkey. Reprinted with permission from [51].
(d) Slow activation during whole behavioural trial comprising stimuli, reaching movement and reward delivery and consumption in monkey. Reprinted with permission from [51].
(e) Scheme of dopamine activation with behavioural activation shown in (a-d). Reprinted with permission from [51].
(f) No activation with well-controlled flexion-extension movements around the ellbow in monkey. Reprinted with permission from [53].
(g) No activation with arm movement despite prediction error response time-locked to conditioned stimulus (CS). Reprinted with permission from [55].
(h) No activation with spontaneous saccadic eye movement. This neurone is activated in response to a CS (not shown). Reprinted with permission from [54].
(i) Gradual increase of voltammetrically measured dopamine concentration in rat striatum over whole trial of movement in T-maze. Reprinted with permission from [44•].
(j) Depression with onset of spontaneous treadmill movement in mouse. Reprinted with permission from [49].
(k) Increased calcium signalling in striatal dopamine axons during treadmill movements. Reprinted with permission from [47].

Figure 3. Dopamine utility prediction error signal and the effects of optogenetic stimulation.
(a) Top: Choices between an adjustable safe reward and a fixed binary, equiprobable gamble, using eye or arm movements. Bottom: Psychophysical assessment of Certainty Equivalent (CE) at low and high reward amounts, respectively. The CE indicates the subjective value of the gamble in ml of the safe juice reward (choice indifference). The CE>EV indicates risk seeking, CE<EV indicates risk avoidance. EV, Expected Value of gamble.
(b) Dopamine responses to increasing reward amount follow nonlinear utility function. Red: utility function estimated from CEs in specifically structured choices under risk. Black: average responses of 16 dopamine neurones to unpredicted reward outside of any behavioural task. **(a)** and **(b)** reprinted with permission from [72•].
(c) Optogenetic stimulation increasing dopamine response to a conditioned, reward-predicting, visual stimulus (CS). The animal received laser stimulation of dopamine neurons expressing channelrhodopsin 2 (ChR2) together with unpredicted juice reward (blue stimulus, left), or it receives the reward alone (red stimulus, right).
(d) Developing oculomotor choice bias toward the blue, optogenetically stimulated CS and away from the red, unstimulated CS (for training history, see (**a**)). Traces and dots, blue: choices (overall % and

individual selections) of blue vs red CS with laser stimulation in the ChR2-injected midbrain (blue); red: laser stimulation in the contralateral, non-ChR2-injected midbrain. **(a)** and **(b)** reprinted with permission from [91].

Figure 4. Effects of dopamine reward prediction error signalling on learning and choices.
(a) Block diagramme of error-driven learning. The equation reflects the Rescorla-Wagner learning rule (V, associative strength or prediction; λ, reward (reinforcer); t, trial). The red arrows denote the positive and negative dopamine prediction error signal and its action on striatal neurones involved in learning and choices. The green-red triangles represent the two-component dopamine signal (green, initial component; red, value response).
(b) Learning of self-stimulation via optogenetic activation of mouse striatal neurones that differentially express dopamine D1 receptors (blue, increased frequency) or D2 receptors (red, decreased frequency). Reprinted with permission from [93].
(c) Biasing of nose poke target choices toward contra- and ipsilateral sides by optogenetic activation of D1-expressing and D2-expressing striatal neurones, respectively. Reprinted with permission from [93].

References

- of special interest
•• of outstanding interest

[1] Ljungberg T, Apicella P, Schultz W: **Responses of monkey midbrain dopamine neurons during delayed alternation performance.** *Brain Res* 1991, **586**:337-341.

[2] Schultz W, Apicella P, Ljungberg T: **Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task**. *J Neurosci* 1993, **13**:900-913.

[3] Mirenowicz J, Schultz W: **Importance of unpredictability for reward responses in primate dopamine neurons**. *J Neurophysiol* 1994, **72**:1024-1027.

[4] Montague PR, Dayan P, Sejnowski TJ: **A framework for mesencephalic dopamine systems based on predictive Hebbian learning**. *J Neurosci* 1996, **16**:1936-1947.

[5] Schultz W, Dayan P, Montague RR: **A neural substrate of prediction and reward**. *Science* 1997, **275**:1593-1599.

[6] Schultz W: **Predictive reward signal of dopamine neurons**. *J Neurophysiol* 1998, **80**:1-27.

[7] Satoh T, Nakai S, Sato T, Kimura, M: **Correlated coding of motivation and outcome of decision by dopamine neurons**. *J Neurosci* 2003, **23**:9913-9923.

[8] Bayer HM, Glimcher PW: **Midbrain dopamine neurons encode a quantitative reward prediction error signal**. *Neuron* 2005, **47**:129-141.

[9] Day JJ, Roitman MF, Wightman RM, Carelli RM: **Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens**. *Nat Neurosci* 2007, **10**:1020-1028.

[10] Enomoto K, Matsumoto N, Nakai S, Satoh T, Sato TK, Ueda Y, Inokawa H, Haruno M, Kimura M: **Dopamine neurons learn to encode the long-term value of multiple future rewards**. *Proc Natl Acad Sci USA* 2011, **108**:15462-15467.

[11] De Lafuente O, Romo R: **Dopamine neurons code subjective sensory experience and uncertainty of perceptual decisions**. *Proc Natl Acad Sci USA* 2011, **108**:19767-19771.

[12] Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N: **Neuron-type-specific signals for reward and punishment in the ventral tegmental area**. *Nature* 2012, **482**:85-88.

• [13] Schultz W: **Dopamine reward prediction error signalling: a two-component response**. *Nat Rev Neurosci* 2016, **17**:183-195. • This review presents the most comprehensive account of the two-component dopamine response.

[14] Nomoto K, Schultz W, Watanabe T, Sakagami M: **Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli**. *J Neurosci* 2010, **30**:10692–10702.

[15] Lak A, Stauffer WR, Schultz W: **Dopamine neurons learn relative chosen value from probabilistic rewards**. *eLife* 2016, **5**: e18044.[16] Sutton RS, Barto AG: *Reinforcement Learning*. MIT Press; 1998.

[17] Bushnell MC, Goldberg ME, Robinson DL: **Behavioral enhancement of visual responses in monkey cerebral cortex. I. Modulation in posterior parietal cortex related to selective visual attention**. *J Neurophysiol* 1981, **46**:755-772.

[18] Pearce JM, Hall G: **A model for Pavlovian conditioning: variations in the effectiveness of conditioned but not of unconditioned stimuli**. *Psychol Rev* 1980, 87:532-552.

[19] Chiodo LA, Antelman SM, Caggiula AR, Lineberry CG: **Sensory stimuli alter the discharge rate of dopamine (DA) neurons: Evidence for two functional types of DA cells in the substantia nigra**. *Brain Res* 1980, **189**:544-549.

[20] Schultz W, Romo R: **Responses of nigrostriatal dopamine neurons to high intensity somatosensory stimulation in the anesthetized monkey**. *J Neurophysiol* 1987, **57**:201-217.

[21] Mirenowicz J, Schultz W: **Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli**. *Nature* 1996, **379**:449-451.
[22] Joshua M, Adler A, Mitelman R, Vaadia E, Bergman H: **Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials**. *J Neurosci* 2008, **28**:11673–11684.
[23] Matsumoto M, Hikosaka O: **Two types of dopamine neuron distinctively convey positive and negative motivational signals.** *Nature* 2009, **459**:837-841.
[24] Brischoux F, Chakraborty S, Brierley DI, Ungless MA: **Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli**. *Proc Natl Acad Sci USA* 2009, **106**:4894–4899.
[25] Lerner T, N, Shilyansky C, Davidson TJ, Luo L, Tomer R, Deisseroth K: **Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits**. *Cell* 2015, **162**:635–647.
• [26] Fiorillo CD, Song MR, Yun SR: **Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli**. *J Neurosci* 2013, **33**:4710–4725. • This study was the first to distinguish 'aversive' dopamine activations from the physical impact of punishers.
[27] Fiorillo CD: **Two dimensions of value: Dopamine neurons represent reward but not aversiveness**. *Science* 2013, **341**:546-549.
• [28] Oleson EB, Gentry RN, Chioma VC, Cheer JF: **Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance**. *J Neurosci* 2012, **32**:14804-14808. • This study was the first to show that dopamine increases may derive from (rewarding) punishment relief.
[29] Stamatakis AM, Stuber GD: **Activation of lateral habenula inputs to the ventral midbrain promotes behavioural avoidance**. *Nat Neurosci* 2012, **15**:1105–1107.
[30] Stopper CM, Tse MTL, Montes DR, Wiedman CR, Floresco SB: **Overriding phasic dopamine signals redirects action selection during risk/reward decision making**. *Neuron* 2014, **84**:177-189.
[31] Lammel S, Lim BK, Ran C, Huang KW, Betley MJ, Tye KM, Deisseroth K, Malenka RC: **Input-specific control of reward and aversion in the ventral tegmental area**. *Nature* 2012, **491**:212-217.
[32] Christoph GR, Leonzio RJ, Wilcox KS: **Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat**. *J Neurosci* 1986, **6**:613-619.
[33] Ji H, Shepard PD. **Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a GABAA receptor-mediated mechanism**. *J Neurosci* 2007, **27**:6923-6930.
[34] Schultz W. **Neuronal reward and decision signals: from theories to data**. *Physiol Rev* 2015, **95**:853-951.
[35] Fiorillo CD, Yun SR, Song MR: **Diversity and homogeneity in responses of midbrain dopamine neurons**. *J Neurosci* 2013, **33**:4693–4709.
[36] Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Naoshige Uchida N: **Arithmetic and local circuitry underlying dopamine prediction errors**. *Nature* 2015, **525**: 243–246.
•• [37] Eshel N, Tian J, Bukwich M, Naoshige Uchida N: **Dopamine neurons share common response function for reward prediction error**. *Nat Neurosci* 2016, **19**:479-486. •• Together with [36], this study quantified the mechanisms of RPE computation common to dopamine neurones.
[38] Joshua M, Adler A, Prut Y, Vaadia E, Wickens JR, Hagai Bergman H: **Synchronization of midbrain dopaminergic neurons is enhanced by rewarding events**. *Neuron* 2009, **62**:695-704.
[39] Fuster JM: **Unit activity of prefrontal cortex during delayed-response performance: Neuronal correlates of transient memory**. *J Neurophysiol* 1973, **36**:61-78.

[40] Hollerman JR, Tremblay L, Schultz W: **Influence of reward expectation on behavior-related neuronal activity in primate striatum**. *J Neurophysiol* 1998, **80**:947-963.
[41] Grabenhorst F, Hernadi I, Schultz W: **Prediction of economic choice by primate amygdala neurons**. Proc Natl Acad Sci USA 2012, **109**:18950-18955.
[42] Schultz W: **Multiple dopamine functions at different time courses**. *Ann Rev Neurosci* 2007, **30**:259-288.
[43] Schultz W: **Behavioral dopamine signals**. *Trends Neurosci* 2007, **30**:203-210.
• [44] Howe MW, Tierney PL, Sandberg SG, Phillips, PEM, Graybiel AM: **Prolonged dopamine signalling in striatum signals proximity and value of distant rewards**. *Nature* 2013, **500**:575-579. • This study was the first to take up again presumably non-RPE-related dopamine activity.
[45] Totah NKB, Yunbok Kim Y, Moghaddam B: **Distinct prestimulus and poststimulus activation of VTA neurons correlates with stimulus detection**. *J Neurophysiol* 2013, **110**:75-85.
[46] Barter JW, Li S, Lu D, Bartholomew RA, Rossi MA, Shoemaker CT, Salas-Meza D, Gaidis E, Yin HH: **Beyond reward prediction errors: the role of dopamine in movement kinematics**. *Front Neurosci* 2015, **9**:39.
[47] Howe MW, Dombeck DA: **Rapid signalling in distinct dopaminergic axons during locomotion and reward**. *Nature* 2016, **535**:505-510.
[48] Parker NF, Cameron CM, Taliaferro JP, Lee J, Choi JY, Davidson TJ, Daw ND, Witten IB: **Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target**. *Nat Neurosci* 2016, **19**:845-854.
[49] Dodson, PD, Jakob K. Dreyer JK, Jennings KA, Syeda ECJ, Wade-Martins R, Cragg SJ, Bolam JP, Magill PJ: **Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism**. *Proc Natl Acad Sci USA* 2016, **113**:E2180-E2188.
[50] Schultz W, Ruffieux A, Aebischer P: **The activity of pars compacta neurons of the monkey substantia nigra in relation to motor activation**. *Exp Brain Res* 1983, **51**:377-387.
[51] Schultz W: **Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey**. *J Neurophysiol* 1986, **56**:1439-1462.
[52] Romo R, Schultz W: **Dopamine neurons of the monkey midbrain: Contingencies of responses to active touch during self-initiated arm movements**. *J Neurophysiol 1990*, **63**:592-606.
[53] DeLong MR, Crutcher MD, Georgopoulos AP: **Relations between movement and single cell discharge in the substantia nigra of the behaving monkey**. *J Neurosci* 1983, **3**:1599-1606.
[54] Schultz W, Romo R: **Dopamine neurons of the monkey midbrain: Contingencies of responses to stimuli eliciting immediate behavioral reactions**. *J Neurophysiol* 1990, **63**:607-624.
[55] Ljungberg T, Apicella P, Schultz W: **Responses of monkey dopamine neurons during learning of behavioral reactions**. *J Neurophysiol* 1992, **67**:145-163.
[56] Matsumoto M, Takada M: **Distinct representations of cognitive and motivational signals in midbrain dopamine neurons**. *Neuron* 2013, **79**:1011-1024.
[57] Fiorillo CD, Tobler PN, Schultz W: **Discrete coding of reward probability and uncertainty by dopamine neurons**. *Science* 2003, 299:1898-1902.
[58] Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ, Berke JD: **Mesolimbic dopamine signals the value of work**. Nat Neurosci 2016, **19**:117-126.
[59] Corbett D, Wise RA: **Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: A moveable microelectrode study**. *Brain Res* 1980, 185:1-15.
[60] Tsai H-C, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, Deisseroth K: **Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning**. *Science* 2009, 324:1080-1084.
• [61] Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH: **A causal link between prediction errors, dopamine neurons and learning**. *Nat Neurosci* 2013, **16**:966-973. • The

classic and elegant study incorporating the previously described learning effects of electrical [59] and optogenetic [60,62] dopamine stimulation into a formal animal learning theory approach (blocking and unblocking).

[62] Tan KR, Yvon C, Turiault M, Mirzabekov JJ, Doehner J, Labouèbe G, Deisseroth K, Tye KM, Lüscher C: **GABA neurons of the VTA drive conditioned place aversion**. *Neuron* 2012, **73**:1173-1183.

[63] Kobayashi S, Schultz W: **Influence of reward delays on responses of dopamine neurons**. *J Neurosci* 2008, **28**:7837-7846.

[64] Day JJ, Jones JL, Wightman RM, Carelli RM: **Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs**. *Biol Psychiat* 2010, **68**:306-309.

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

•• [66] Cone JJ, Fortin SM, McHenry JA, Stuber GD, McCutcheon JE, Roitman MF: **Physiological state gates acquisition and expression of mesolimbic reward prediction signals**. *Proc Natl Acad Sci USA* 2016, 113:1943-1948. •• A particularly elegant study testing the appearance of reward value due to salt depletion.

[67] Pasquereau B, Turner RS: **Limited encoding of effort by dopamine neurons in a cost–benefit trade-off task**. *J Neurosci* 2013, **33**:8288–8300.

[68] Varazzani C, San-Galli A, Gilardeau S, Bouret S: **Noradrenaline and dopamine neurons in the reward/effort trade-off: A direct electrophysiological comparison in behaving monkeys**. *J Neurosci* 2015, **35**:7866-7877.

[69] Kishida KT, Saez I, Lohrenz T, Witcher MR, Laxton AW, Tatter SB, White JP, Ellis TL, Phillips PEM, Montague PR: **Subsecond dopamine fluctuations in human striatum encode superposed error signals about actual and counterfactual reward**. *Proc Natl Acad Sci USA* 2016, **113**:200-205.

••[70] von Neumann J, Morgenstern O: *The Theory of Games and Economic Behavior*. Princeton University Press; 1944. •• The classic and genius origin of Expected Utility Theory, and cardinal utility under risk, still valid after all those years with some additions (e. g. Prospect Theory).

[71] Caraco T, Martindale S, Whitham TS: **An empirical demonstration of risk-sensitive foraging preferences**. *Anim Behav* 1980, **28**:820-830.

• [72] Stauffer WR, Lak A, Schultz W: **Dopamine reward prediction error responses reflect marginal utility**. *Curr Biol* 2014, **24**:2491-2500. • The conceptually important step from objective and subjective reward value coding to formal economic utility signalling.

• [73] Genest W, Stauffer WR, Schultz W: **Utility functions predict variance and skewness risk preferences in monkeys**. **Proc Natl Acad Sci USA** 2016, **113**:8402-8407. • A behavioural study in monkeys showing the unique capacity of utility functions to predict choices under different forms of risk.

[74] Tsai H-C, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, Deisseroth K: **Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning**. *Science* 2009, **324**:1080-1084.

[75] Witten IB, Steinberg EE, Lee SY, Davidson TJ, Zalocusky KA, Brodsky M, Yizhar O, Cho SL, Gong S, Ramakrishnan C, Stuber GD, Tye KM, Janak PH, Deisseroth K: **Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement**. *Neuron* 2011, **72**:721-733.

[76] Adamantidis AR, Tsai H-C, Boutrel B, Zhang F, Stuber GD, Budygin EA, Touriño C, Bonci A, Deisseroth K, de Lecea L: **Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior**. *J Neurosci* 2011, **31**:10829-10835.

[77] Kim KM, Baratta MV, Yang A, Lee D, Boyden ES, Fiorillo CD: **Optogenetic mimicry of the transient activation of dopamine neurons by natural reward is sufficient for operant reinforcement**. *PLoS ONE* 2012, **7**:e33612.
[78] Ilango A, Kesner AJ, Keller KL, Stuber GD, Bonci A, Ikemoto S: **Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion**. *J Neurosci* 2014, **34**:817-822.
[79] Zweifel LS, Parker JG, Lobb CJ, Rainwater A, Wall VZ, Fadok JP, Darvas M, Kim MJ, Mizumori SJ, Paladini CA, Philipps PEM, Palmiter R: **Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior**. *Proc Natl Acad Sci USA* 2009, **106**:7281–7288.
[80] Saddoris MP, Sugam JA, Stuber GD, Witten IB, Deisseroth K, Carelli RM: **Mesolimbic dopamine dynamically tracks, and is causally linked to, discrete aspects of value-based decision making.** *Biol Psychiatry* 2015, **77**:903-911.
[81] van Zessen R, Phillips JL, Budygin EA, Stuber GD: **Activation of VTA GABA neurons disrupts reward consumption**. *Neuron* 2012, **73**:1184–1194.
[82] Han X, Qian X, Bernstein JG, Zhou HH, Franzesi GT, Stern P, Bronson RT, Graybiel AM, Desimone R, Boyden ES: **Millisecondtimescale optical control of neural dynamics in the nonhuman primate brain**. *Neuron* 2009, **62**:191–198.
[83] Diester I, Kaufman MT, Mogri M, Pashaie R, Goo W, Yizhar O, Ramakrishnan C, Deisseroth K, Shenoy KV: **An optogenetic toolbox designed for primates**. *Nat Neurosci* 2011, **14**:387–397.
[84] Cavanaugh J, Monosov IE, McAlonan K, Berman R, Smith MK, Cao V, Wang KH, Boyden ES, Wurtz RH: **Optogenetic inactivation modifies monkey visuomotor behavior**. *Neuron* 2012, **76**:901–907.
[85] Galvan A, Hu X, Smith Y, Wichmann T: **In vivo optogenetic control of striatal and thalamic neurons in non-human primates**. *PLoS ONE* 2012, **7**:e50808.
[86] Galvan A, Hu X, Smith Y, Wichmann T: **Effects of optogenetic activation of corticothalamic terminals in the motor thalamus of awake monkeys**. *J Neurosci* 2016, **36**:3519–3530.
[87] Gerits A, Farivar R, Rosen BR, Wald LL, Boyden ES, Vanduffel W: **Optogenetically induced behavioral and functional network changes in primates**. *Curr Biol* 2012, **22**:1722–1726.
[88] Jazayeri M, Lindbloom-Brown Z, Horwitz GD: **Saccadic eye movements evoked by optogenetic activation of primate** V1. *Nat Neurosci* 2012, **15**:1368–1370.
[89] Ohayon S, Grimaldi P, Schweers N, Tsao DY: **Saccade modulation by optical and electrical stimulation in the macaque frontal eye field**. *J Neurosci* 2013, **33**:16684–16697
[90] Dai J, Brooks DI, Sheinberg DL: **Optogenetic and electrical microstimulation systematically bias visuospatial choice in primates**. *Curr Biol* 2014, **24**:63–69.
[91] Stauffer WR, Lak A, Yang A, Borel M, Paulsen O, Boyden E, Schultz W: **Dopamine neuron-specific optogenetic stimulation in Rhesus macaques**. *Cell* 2016, **166**:1564-1571.
[92] Arsenault JT, Rima S, Stemmann H, Vanduffel W: **Role of the primate ventral tegmental area in reinforcement and motivation**. *Curr Biol* 2014, **24**:1347–1353.
[93] Kupchik YM, Brown RM, Heinsbroek JA, Lobo MK, Schwartz DJ, Kalivas PW: **Coding the direct/indirect pathways by D1 and D2 receptors is not valid for accumbens projections**. *Nat Neurosci* 2015, **18**:1230-1232.
[94] Kravitz AV, Tye LD, Kreitzer AC: **Distinct roles for direct and indirect pathway striatal neurons in reinforcement**. *Nat Neurosci* 2012, **15**:816-818.
[95] Yttri EA, Dudman JT: **Opponent and bidirectional control of movement velocity in the basal ganglia**. *Nature* 2016, **533**: 402-406.
[96] Tai L-H, Lee AM, Benavidez N, Bonci A, Wilbrecht L: **Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value**. *Nat Neurosci* 2012, **15**:1281-1289.

[97] Lee HJ, Weitz AJ, Bernal-Casas D, Duffy, BA, Choy MK, Kravitz AV, Kreitzer AC, Lee JH: **Activation of direct and indirect pathway medium spiny neurons drives distinct brain-wide responses**. *Neuron* 2016, **91**:412-424.

Figure 1

Click here to download high resolution image

Figure 1

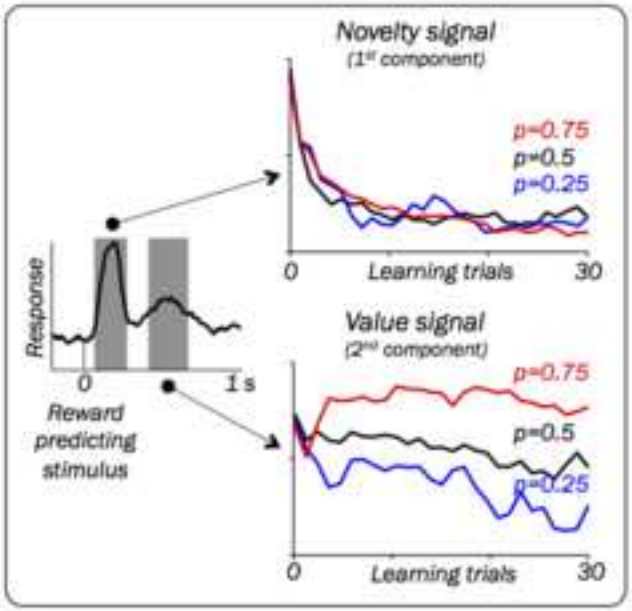

Figure 2

Click here to download high resolution image

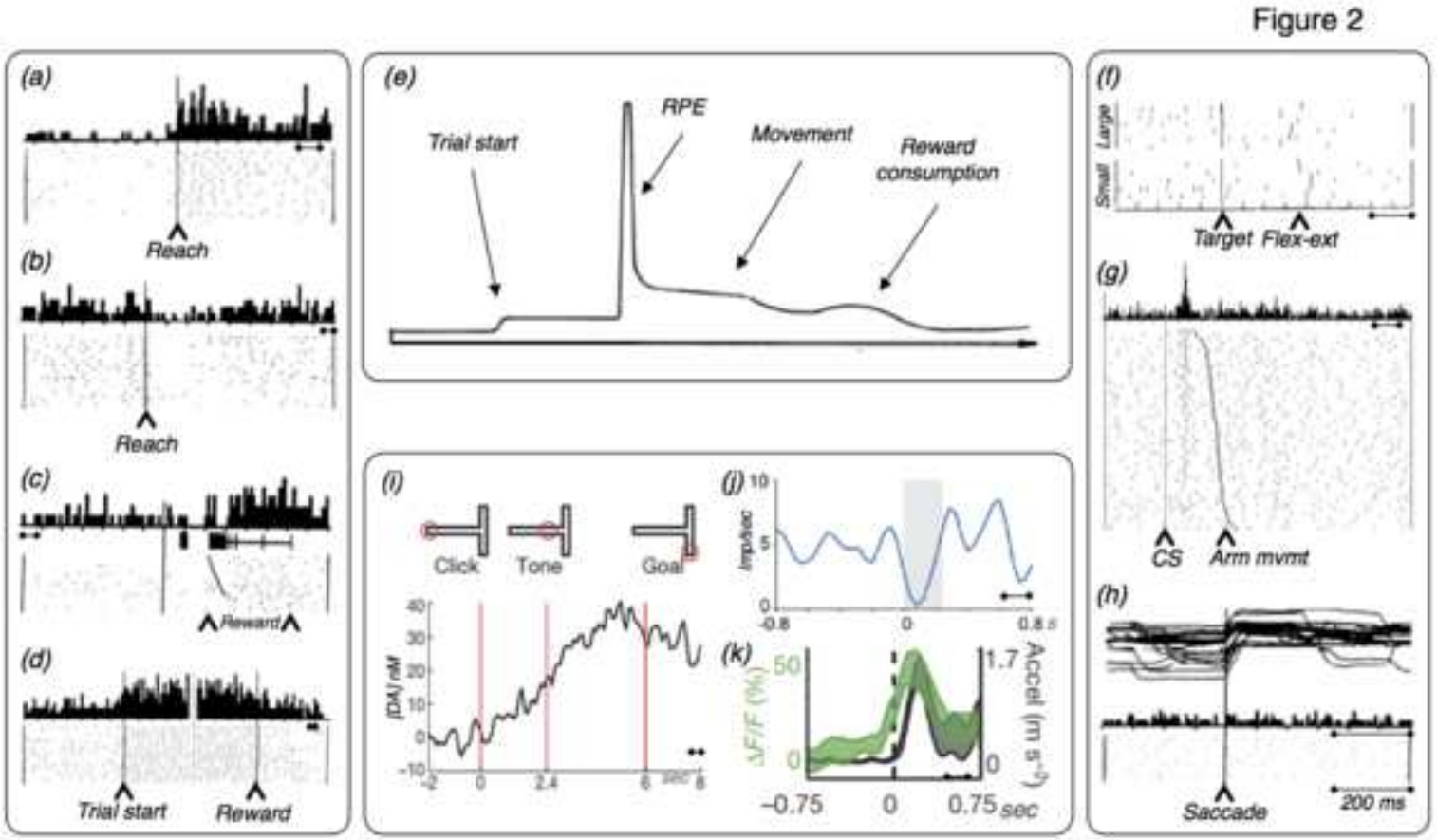

Figure 3

Click here to download high resolution image

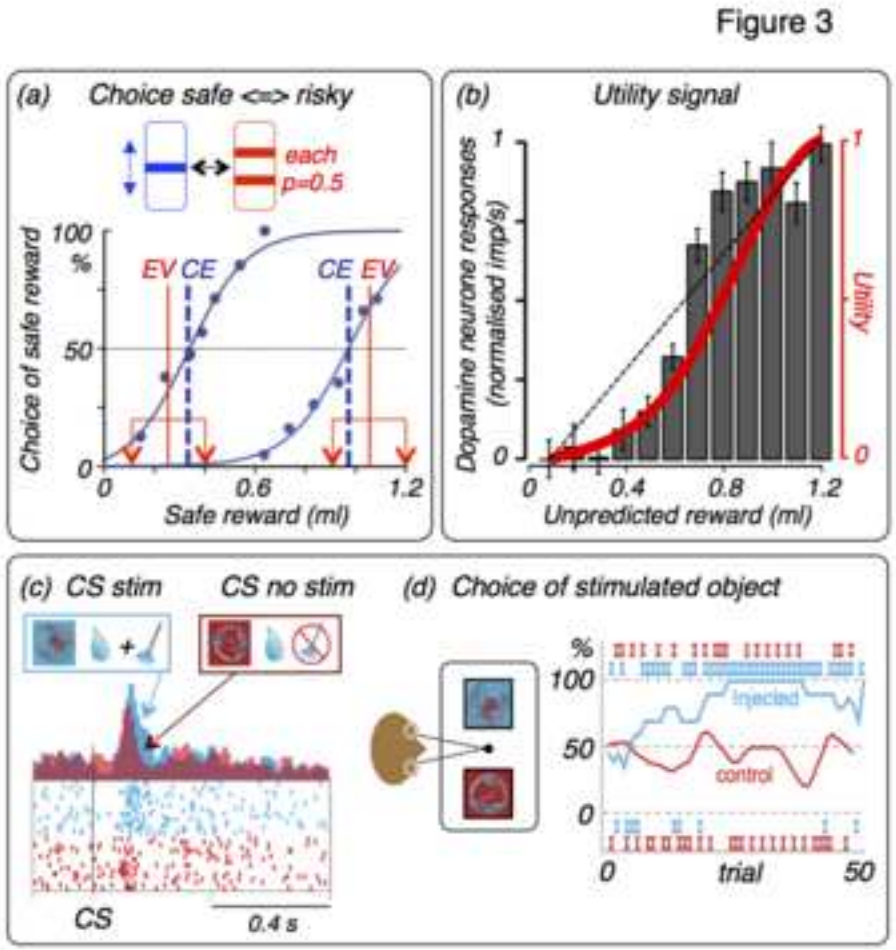

Figure 4

Click here to download high resolution image