Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine

Jerylin O Gan[1,2,5], Mark E Walton[1,3,5] & Paul E M Phillips [1,2,4]

Reward-predicting cues evoke activity in midbrain dopamine neurons that encodes fundamental attributes of economic value, including reward magnitude, delay and uncertainty. We found that dopamine release in rat nucleus accumbens encodes anticipated benefits, but not effort-based response costs unless they are atypically low. This neural separation of costs and benefits indicates that mesolimbic dopamine scales with the value of pending rewards, but does not encode the net utility of the action to obtain them.

For individuals to prosper in diverse environments, they need to use predictive sensory information to optimize outcomes in a flexible manner. Decision-making processes weigh the benefits of a reward with the cost of obtaining it to determine the overall subjective value (utility) of the transaction[1,2]. Dopamine is a neural substrate that has been heavily implicated in this valuation process. Midbrain dopamine neurons encode fundamental economic parameters pertaining to predicted rewards (magnitude, probability, delay and uncertainty) in their firing rate[3–6] and innervate areas that have been implicated in economic decision-making (prefrontal cortex, amygdala, dorsal striatum and nucleus accumbens)[7–9]. Moreover, dopamine in the nucleus accumbens core (NAcc) enables animals to respond to cues and overcome effortful response costs[10,11]. However, to fully understand decision-making computations encoded by the mesoaccumbens dopamine pathway, we need to deconstruct the nature of the valuation signal: specifically, how it accounts for changes in anticipated costs and benefits.

Rats were trained on decision-making tasks (**Supplementary Fig. 1**) that independently manipulated either benefits or cost. We employed fast-scan cyclic voltammetry (see **Supplementary Methods** and **Supplementary Fig. 2**) to record phasic dopamine transmission in NAcc (**Supplementary Fig. 3**) while rats performed these tasks. All of the procedures on animals were approved by the University of Washington Institutional Animal Care and Use Committee. Rats were trained to select between a reference option (16 lever presses for 1 food pellet) and an alternative that differed in either the reward magnitude (4 or 0 food pellets, benefit conditions) or response requirement (2 or 32 lever presses, cost conditions) (see **Supplementary Methods**). Cues signaling the availability of the reference and/or alternative options were presented either separately in forced trials or simultaneously in

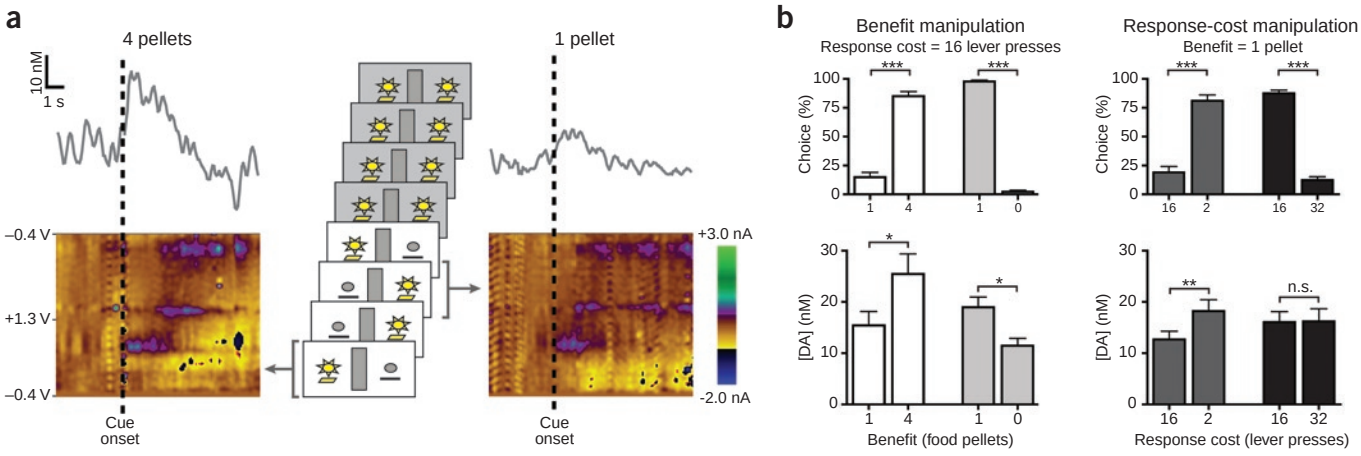


Figure 1 Decision making following manipulation of benefits or costs. (**a**) Example trials in the benefit condition. Center schematic represents cue lights (yellow star, active; gray circle, inactive) and levers (trapezoid, present; line, retracted) flanking the food magazine. Each frame represents response options on one trial (white background, forced; gray background, choice). The outside panels are representative examples of dopamine release evoked by presentation of cue (dashed line) predicting the availability of a response option resulting in four (left) or one (right) food pellets. The color plots provide electrochemical information for these examples with voltammetric scans plotted on the *y* axis, time of consecutive scans on the *x* axis and electrochemical current represented by color. (**b**) Post-criterion choice behavior (top) and cue-evoked dopamine release (bottom) across sessions in benefit and cost conditions. Data are mean ± s.e.m. * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.0001$. DA, dopamine.

[1]Department of Psychiatry & Behavioral Sciences, [2]Graduate Program in Neurobiology and Behavior, University of Washington, Seattle, Washington, USA. [3]Department of Experimental Psychology, University of Oxford, Oxford, UK. [4]Department of Pharmacology, University of Washington, Seattle, Washington, USA. [5]These authors contributed equally to this work. Correspondence should be addressed to P.E.M.P. (pemp@uw.edu).

Received 9 October; accepted 2 November; published online 10 November 2009; doi:10.1038/nn.2460

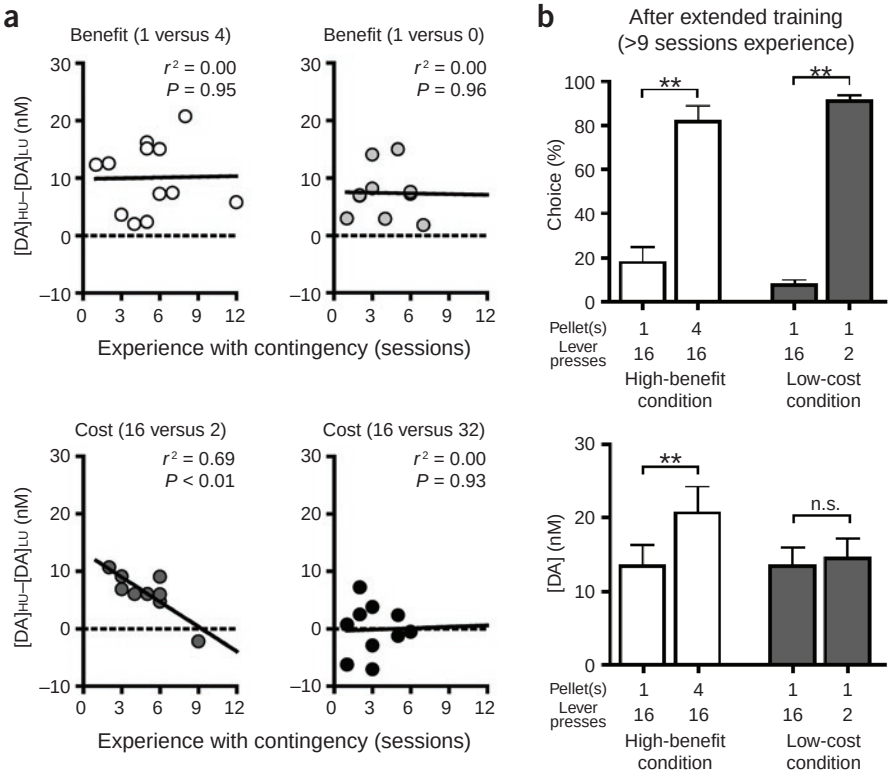


Figure 2 Effect of behavioral history on dopamine release. (**a**) Differences in cue-evoked dopamine release between the high- and low-utility options ($[DA]_{HU} - [DA]_{LU}$) against behavioral history. (**b**) Post-criterion choice behavior (left) and cue-evoked dopamine release (right) for the high-benefit (4 food pellets for 16 lever presses, left) or low-cost (1 food pellet for 2 lever presses, right) option in rats given extended training (>9 sessions) with either contingency before testing. Data are mean ± s.e.m. * $P < 0.05$, ** $P < 0.01$.

choice trials (**Fig. 1a** and **Supplementary Fig. 1**). Forced trials allowed the evaluation of cue-evoked dopamine for one option without the confound of another option being present and choice trials provided a measure of behavioral preference. Data were evaluated after the rats reached a behavioral criterion, choosing one option on ≥75% of choice trials. To prevent side-bias, we always reversed the assignment of high-/low-utility options to the two levers from the previous session and included counterbalanced sessions for each contingency pair in the analysis.

Across all contingency pairs, the rats consistently chose the option with the highest benefit or lowest cost (**Fig. 1b**, see **Supplementary Fig. 4** for rate to criterion). Subjective preference was also evident on post-criterion forced trials where response latencies were significantly faster to higher-benefit or lower-cost options (all $P < 0.001$; **Supplementary Fig. 4**). Furthermore, when the high-benefit (4 pellets for 16 lever presses) and the low-cost (1 pellet for 2 lever presses) options were presented as concurrent choices in a decision-making session, the rats were indifferent, demonstrating equivalent utility (**Supplementary Fig. 5**). Thus, not only was the utility of reward options successfully modulated as expected by both benefit and cost conditions (that is, increased utility conferred to the option with greater benefit or lower cost), the additional utility conferred by increased benefits was equivalent to that conferred by decreased costs.

Despite predictable behavior, cue-evoked NAcc dopamine release did not track utility under all conditions. Manipulating reward magnitude led to a corresponding increase (main effect of reward size, $F_{1,5} = 15.61$, $P = 0.01$) or decrease ($F_{1,4} = 19.88$, $P = 0.01$) in cue-evoked dopamine compared with the reference option (**Fig. 1b** and **Supplementary Fig. 6**). Manipulations of response cost, on the other hand, did not always alter dopamine release. When the response cost of the alternative was increased, there was no difference in dopamine release between the reference and alternative option (main effect of response cost, $F_{1,4} = 0.05$, $P = 0.84$; **Fig. 1b**), despite the strong behavioral preference for the reference option. When the response cost was reduced, there was greater dopamine release to the low-cost cue than to the reference ($F_{1,4} = 25.38$, $P = 0.007$), but this was only significant in the first of two counterbalanced sessions in each rat (session × option interaction, $P = 0.03$, $F_{1,4} = 10.92$; **Supplementary Fig. 6**). *Post hoc* tests indicated that this effect was driven by a reduction in dopamine release to the low-cost cue ($P = 0.0006$), but not the reference cue ($P = 0.20$), across sessions.

To further investigate across-session effects, we performed regression analysis between utility encoding and experience with any alternative contingency before recording. Experience-related changes in cue-evoked dopamine release were only observed in the reduced-cost condition, in which the preferential dopamine release for the low-cost cue diminished over time (Pearson's $r = -0.830$, $P = 0.005$, $n = 9$; Spearman's rho = –0.817, $P = 0.007$; **Fig. 2a**). Additional experimentation with a cohort of rats that were given more experience (>9 sessions) with the high-benefit option before recording verified that both behavioral preference and preferential encoding of the higher benefits was maintained with extended training ($P = 0.007$, $t = 4.08$, degrees of freedom = 6, $n = 7$ session; **Fig. 2b**). Conversely, in a parallel experiment with the low-cost option, cue-evoked dopamine release did not preferentially encode the low-cost option after additional experience before recording ($P = 0.16$, $t = 1.55$, degrees of freedom = 8, $n = 9$ sessions), even though behavioral preference was preserved (**Fig. 2b**). These data are consistent with the notion that, although preferential encoding of high benefit by dopamine release is stable over training, low costs are only preferentially encoded early in training. Further analyses of the neurochemical data with respect to contextual framing, choice trials (**Supplementary Fig. 7**) and within-session learning (**Supplementary Fig. 8**) are included in the **Supplementary Results**.

When making sound economic choices, one must consider a reasonable cost to obtain an outcome on the basis of its perceived benefit. The data presented here indicate that phasic NAcc dopamine transmission reliably reflects the magnitude of the benefit, but only correlates with effort-discounted utility in situations in which the response cost is both novel and better than the reference. Incorporating these findings with those of previous studies showing that dopamine enables effortful responses, we reason that representation of reward magnitude by phasic dopamine provides

a threshold to determine worthwhile cost expenditures in familiar situations[10–12]. Moreover, in novel situations, dopamine provides an additional opportunistic mechanism for exploitation of low-cost rewards that become available unexpectedly[12,13]. Thus, we found a dissociation between dopaminergic encoding of anticipated costs and benefits, indicating that, although dopamine release in the nucleus accumbens scales with the value of a pending reward, it is not sufficient to describe the net utility of the action to obtain it.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We would like to thank S. Ng-Evans for invaluable technical support, C. Akers and S. Barnes for assistance, and J. Clark, S. Sandberg and M. Wanat for helpful comments. This work was funded by the National Institutes of Health (R01-MH079292 and R21-AG030775 to P.E.M.P.) and a Wellcome Trust Advanced Training Fellowship (M.E.W.). J.O.G. was supported by the National Institute of General Medical Sciences (T32-GM007270, Kimelman).

AUTHOR CONTRIBUTIONS

M.E.W. and P.E.M.P. conceived the study. J.O.G. and M.E.W. collected and analyzed the data. All authors contributed to experimental design and preparation of the manuscript.

Published online at http://www.nature.com/natureneuroscience/.
Reprints and permissions information is available online at http://www.nature.com/reprintsandpermissions/.

1. Stephens, D.W. & Krebs, J.R. *Foraging Theory* (Princeton University Press, Princeton, New Jersey, 1986).
2. Walton, M.E., Kennerley, S.W., Bannerman, D.M., Phillips, P.E.M. & Rushworth, M.F. *Neural Netw.* **19**, 1302–1314 (2006).
3. Fiorillo, C.D., Tobler, P.N. & Schultz, W. *Science* **299**, 1898–1902 (2003).
4. Roesch, M.R., Taylor, A.R. & Schoenbaum, G. *Neuron* **51**, 509–520 (2006).
5. Morris, G., Nevet, A., Arkadir, D., Vaadia, E. & Bergman, H. *Nat. Neurosci.* **9**, 1057–1063 (2006).
6. Kobayashi, S. & Schultz, W. *J. Neurosci.* **28**, 7837–7846 (2008).
7. Glimcher, P.W., Dorris, M.C. & Bayer, H.M. *Games Econ. Behav.* **52**, 213–256 (2005).
8. Knutson, B., Delgado, M.R. & Phillips, P.E.M. in *Neuroeconomics: Decision Making and the Brain* (eds. Glimcher, P.W., Camerer, C.F., Fehr, E. & Poldrack, R.A.) 389–406 (Academic Press, London, 2008).
9. Floresco, S.B., St Onge, J.R., Ghods-Sharifi, S. & Winstanley, C.A. *Cogn. Affect. Behav. Neurosci.* **8**, 375–389 (2008).
10. Salamone, J.D., Correa, M., Farrar, A. & Mingote, S.M. *Psychopharmacology (Berl.)* **191**, 461–482 (2007).
11. Fields, H.L., Hjelmstad, G.O., Margolis, E.B. & Nicola, S.M. *Annu. Rev. Neurosci.* **30**, 289–316 (2007).
12. Phillips, P.E.M., Walton, M.E. & Jhou, T.C. *Psychopharmacology (Berl.)* **191**, 483–495 (2007).
13. Redgrave, P. & Gurney, K. *Nat. Rev. Neurosci.* **7**, 967–975 (2006).

Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine

Jerylin O. Gan [1, 2, 5], Mark E. Walton [1, 4, 5] and Paul E. M. Phillips [1, 2, 3]

[1] Department of Psychiatry & Behavioral Sciences
[2] Graduate Program in Neurobiology & Behavior
[3] Department of Pharmacology
University of Washington
Seattle, WA 98195
U.S.A.
[4] Department of Experimental Psychology
University of Oxford
Oxford, OX1 3UD,
U.K.
[5] These authors contributed equally to the work

Correspondence: pemp@uw.edu

Supplementary Methods (inc. Supp. fig. 1-3), Results (inc. Supp. fig. 4-8), References

Nature Neuroscience: doi:10.1038/nn.2460

SUPPLEMENTARY MATERIALS

Nature Neuroscience: doi:10.1038/nn.2460

Supplementary Methods

Animals

All procedures were approved by the University of Washington Institutional Animal Care and Use Committee. Thirty-five naïve male Sprague-Dawley rats (Charles River, CA, 3-8 months old during testing) were used for this experiment. Seventeen animals contributed to the data reported here. All other animals were excluded based upon histology (10 animals), electrical issues such as connector malfunction or electrode saturation (4 animals) or failing to meet criterion for dopamine detection (4 animals; see *Recording Sessions* for details). Animals were maintained on a twelve-hour light/dark cycle (lights on 0700) and were group housed during initial habituation and training but individually housed following surgery. All testing was carried out during the light phase. During the training and testing periods, access to food was restricted to a total of ~12-16 g per day, consisting of the reward pellets gained during testing supplemented by lab chow given at the end of the day, such that rats' weights were kept at 85-90% of their free-feeding weight. Water was available *ad libitum* while animals were in their home cages.

Behavioral training

Testing was carried out in operant chambers (30.5 x 24.1 x 29.2 cm; Med Associates, VT, USA) with sloped inserts between the floor and walls (63° towards the levers and magazine, and the back wall, 52° towards the sides). Each chamber was housed within a custom-built sound-attenuating cabinet ventilated with a fan. Each chamber was fitted with two retractable levers on either side of an extra-tall food magazine into which 45-mg food pellets (Bioserv, NJ, USA) could be dispensed. Above each lever was a stimulus light, which could act as a visual cue, and the chamber could be illuminated by a 2.8-W house light located at the top of the wall opposite the levers and food magazine. The food magazine was fitted with an infrared beam that could signal when animals entered the receptacle and could also be illuminated by an internal light.

Habituation and training were comparable to that performed in previous studies of operant cost-benefit decision making[1,2]. In brief, following initial habituation to the chambers, rats experienced a 60-minute session in which a single reward, cued by the magazine light, was dispensed under a variable interval schedule (every 40-80s with a 60s mean). On the following sessions, animals were trained to lever press for reward on a fixed ratio (FR) 1 schedule. The house light was illuminated throughout, and either the left or right lever (counterbalanced across animals) was extended and its associated cue light illuminated throughout the session. To

Nature Neuroscience: doi:10.1038/nn.2460

facilitate responding in some animals, a few food pellets were placed behind the extended lever such that their odor was evident but the pellets themselves were unobtainable.

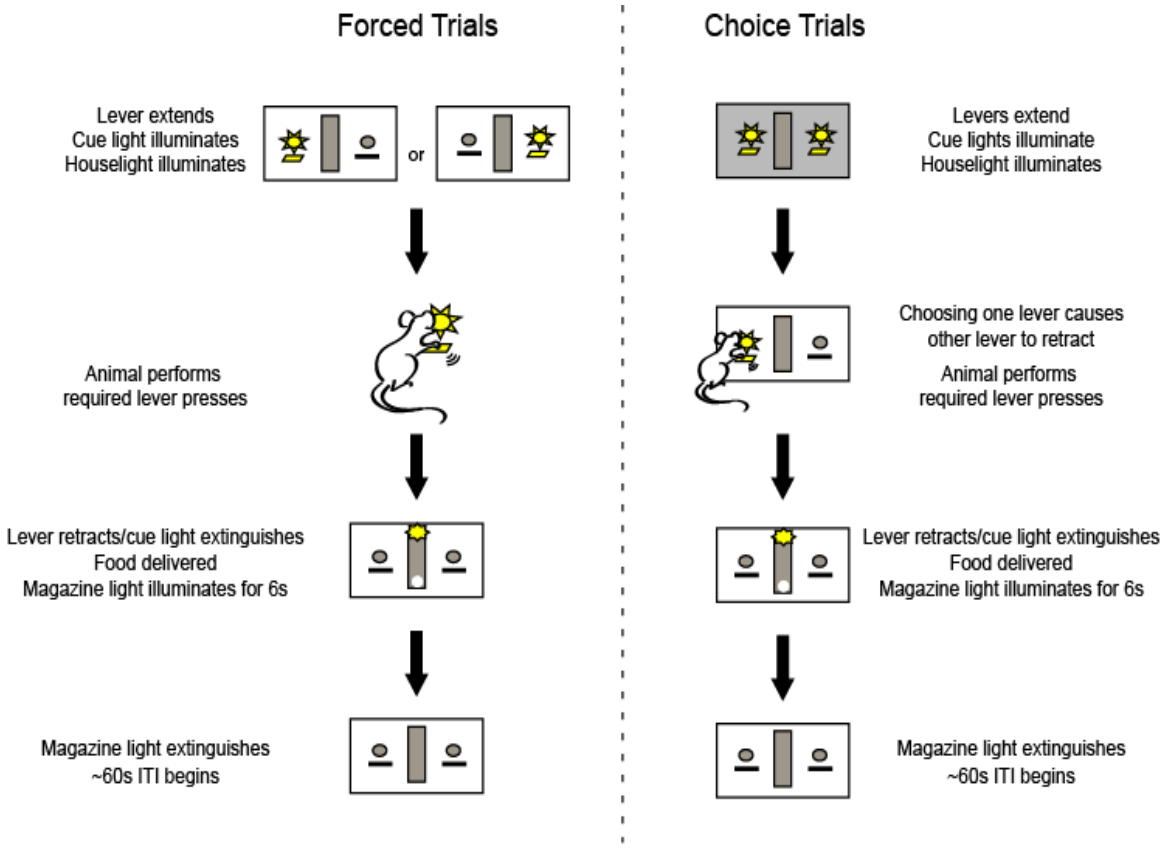
Once animals reliably responded on both levers, the paradigm was changed so completing the response requirement caused the lever to retract and the associated cue light to extinguish. At the same time, reward was delivered and the magazine-light was illuminated. Six seconds after food delivery, the magazine-light was extinguished and the intertrial-interval (ITI) began. The start of a subsequent trial was signaled by illumination of one of the two cue lights and simultaneous extension the associated lever. In these "forced" trials (where only one of the two response options was available), the response cost was increased on each lever across sessions up to a maximum of sixteen lever presses for a single pellet. This response cost (16 lever presses) and reward (1 pellet) is subsequently referred to throughout as the "reference" option. Once animals responded on both levers with the reference response requirement across 80 trial sessions, they subsequently underwent surgery to allow for *in vivo* voltammetric recording.

Decision-making sessions

Following recovery from surgery, rats were reintroduced to the behavioral task described above. Once pre-surgery levels of performance were achieved, the animals were introduced to new contingencies where the benefit or the cost was altered from the reference (16 lever presses for 1 food pellet). These contingencies consisted of four or zero food pellets for sixteen lever presses (benefit manipulations) or one food pellet for two or thirty-two lever presses (cost manipulations). In each session, the altered contingency was assigned to one lever with the reference assigned to the other and remained fixed for the entire session. To avoid side-biased habit formation, the lever assigned to the high-value option was reversed at the start of each session.

Reference and alternative options were presented independently in "forced" trials or concurrently in "choice" trials. Forced trials ensured that the animal experienced both the preferred and non-preferred contingencies throughout the session while choice trials permitted assessment of the animal's subjective preference. Sessions were comprised of repeating blocks of four forced trials (each option presented twice in pseudo-random order) followed by four choice trials. A schematic of the protocol used throughout testing can be seen in Supp. fig. 1 and Fig. 1a.

Nature Neuroscience: doi:10.1038/nn.2460



Supplementary figure 1. Schematic of a forced (left hand panel) or choice trial (right hand panel).

The start of each trial (forced or choice) was signaled by the illumination of the house light, presentation of the lever(s) and illumination of the associated cue light(s). During choice trials, the first lever press caused the other lever to retract and its cue light to extinguish, eliminating the unselected option for that trial. Completion of the response requirement on the selected lever resulted in reward delivery. At this time, the lever was retracted, the cue light was extinguished, the magazine light was illuminated, and the appropriate reward magnitude was delivered to the magazine. After six seconds, the house and magazine lights were extinguished and an inter-trial interval commenced. The inter-trial interval was sixty seconds minus the time taken to complete the response requirement for the completed trial, ensuring that the overall rate of reward delivery throughout the session was independent of choice and response rates. If animals did not make a lever-press response within ten seconds from the start of a trial, all lights were extinguished for a "time out" of sixty seconds.

On each session animals learn the assignment of the contingencies to the levers, as evidenced by development of a preference for one lever during choice trials. Preference is inferred when a behavioral criterion was reached, defined as choosing one option ≥75% of the last twelve choice

Nature Neuroscience: doi:10.1038/nn.2460

trials. For example, an animal reached the behavioral criterion when it chooses 4 pellets over 1 pellet in nine out of the last twelve choice trials. Decision-making sessions continued for 6-8 blocks after animals reached this criterion or a maximum of 120 trials. No additional training was provided to teach the animals to choose between the alternatives. However, for each condition, all animals completed at least two (side-counterbalanced) decision-making sessions to criterion while tethered to the voltammetry recording equipment prior to the first session of voltammetric data acquisition.

To prevent our results being influenced by the order of testing, half of the animals started by performing a benefit condition and the other half, a cost condition. The order (alternative option = higher/lower utility than the reference option) and side (reference option = left or right lever) of the cost-benefit contingencies was counterbalanced across animals.

Test of utility equivalence between the high-benefit and low-cost contingencies

Both the high-benefit (4 pellets for 16 presses) and low-cost (1 pellet for 2 presses) options were preferred over the reference option (1 pellet for 16 presses) (see Results). However, these data do not tell us the relative utility of these options compared to each other. To test whether the utility conferred by the increased benefit was equivalent to that conferred by the decreased cost, eight rats took part in further cost-benefit behavioral experiments where the high-benefit and low-cost options were compared directly. The high-benefit and low-cost contingencies were assigned to the left and right levers counterbalanced across animals for a first session and reversed on a second session. During these sessions, animals were tethered to the voltammetry recording equipment during testing to mimic the conditions during recording sessions, although electrochemical data were not acquired.

Assignment of a behavioral criterion to assess a learned preference for one option was not pertinent in this experiment because it was reasonable that a strong preference to one contingency would not prevail. Therefore, animals were pre-trained with 16 forced trials (8 for each contingency) to provide experience with the pairing comparable to that for the pre-criterion trials of a decision-making session where one contingency is paired with the reference option. Thirty minutes after pre-training, animals were tested in a session consisting of blocks of trials similar to those previously described, up to a maximum of 56 trials. Animals were tested on this utility equivalence experiment after either ≤9 training sessions (n=5) or extended training of >9

Nature Neuroscience: doi:10.1038/nn.2460

sessions (n=5) of experience with the high-benefit or low-cost contingencies (in separate sessions paired with the reference option).

Surgical procedures

Following habituation and initial operant training, animals underwent surgical preparation for in vivo voltammetry using an aseptic technique, following the University of Washington Institutional Animal Care and Use Committee guidelines. All rats were anesthetized with ~5% isoflurane and maintained during surgery with ~2-3% isoflurane. They were placed in a stereotaxic frame, the scalp was swabbed with 10% iodine, bathed with a mixture of lidocaine (0.5 mg/kg) and bupivicaine (0.5 mg/kg), and an incision with made over the midline to expose the cranium. After the head was leveled between bregma and lambda, holes were drilled for 3 anchor screws and a reference electrode, along with 2 others bilaterally above the NAcc (at +1.3 mm anterior and ±1.3 mm lateral to bregma). The NAcc was targeted (rather than the adjacent shell region) as this has been suggested to be the critical site where dopamine allows animals to overcome effort constraints[3]. In-house constructed carbon fiber microelectrodes for long-term chronic recordings were lowered into position (+6.8-7.0 mm ventral to dura), and these, along with an Ag/AgCl reference electrode, were attached to a voltammetric amplifier. Voltammetric components along with a headpost were secured with cranioplastic cement. Rats were given an injection of 5mg/kg carprofen mixed in with 3ml ringer's solution immediately following surgery and again 12 hours later. The animals were allowed between 7-14 days to recover with food and water freely available before being food deprived again prior to further behavioral training and testing.

Recording sessions

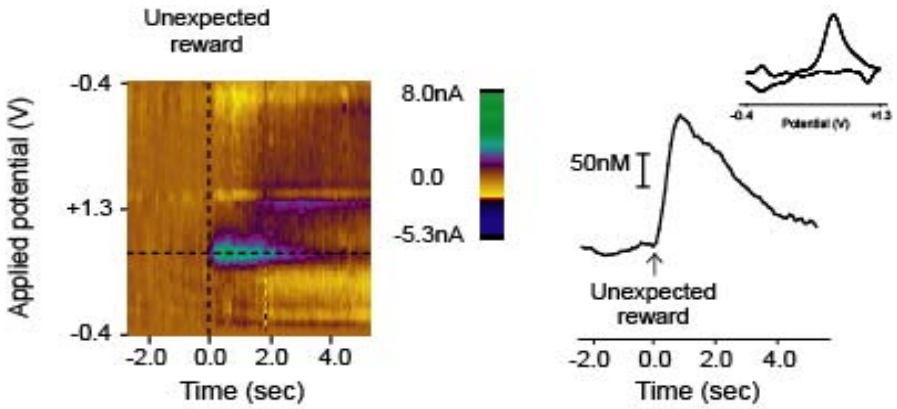
During experimental recording sessions, the chronically-implanted carbon-fiber microelectrodes were connected to a head-mounted voltammetric amplifier for dopamine detection by fast-scan cyclic voltammetry as described in detail elsewhere[4]. In brief, the potential applied to the carbon fiber was ramped from -0.4 V (vs Ag/AgCl) to +1.3 V and back at a rate of 400 V/s during a voltammetric scan and held at -0.4 V between scans. Scans were repeated at a frequency of 10 Hz throughout the session. The application of this triangular waveform causes redox reactions in electrochemically active species at the carbon fiber (including dopamine: ~+0.7 V and -0.3 V peak oxidation and reduction potentials respectively) which can be measured as changes in current. The average current from the scans obtained in the second prior to cue

Nature Neuroscience: doi:10.1038/nn.2460

presentation was subtracted from the current generated in each scan within a trial to yield background-subtracted signals[5,6].

To further ensure that recording electrodes were able to reliably detect behaviorally-evoked dopamine, we measured the neurochemical response to a food pellet delivered to the magazine without forewarning at the start and end of each session. This procedure has been shown to consistently increase burst firing in midbrain dopaminergic neurons[7] and also to elicit dopamine release in the nucleus accumbens[5] (Supp. fig. 2). The inclusion criterion for neurochemical recording sessions was electrochemically verifiable dopamine release for unexpected food-pellet delivery both before and after the session. This criterion was not met for four animals which were excluded from the study.

Chemical verification was achieved by obtaining high correlation of the cyclic voltammogram (electrochemical signature) to that of a dopamine standard (correlation coefficient $r^2 \geq 0.75$ by linear regression). The only other analyte known to closely approximate the chemical signature of dopamine is norepinephrine. However, the norepinephrine tissue content in the NAcc is only 2-20% of that for dopamine[2,8] and electrode sensitivity to norepinephrine is approximately half of its sensitivity to dopamine[4]. Therefore, it highly unlikely that norepinephrine contributes to any signals observed in the current experiment.



Supplementary figure 2. Example response following delivery of an unexpected food reward. Left-hand panel shows the background-subtracted recorded current change time-locked to delivery of the reward. Color plot is a two-dimensional representation of a series of cyclic voltammograms across time. Dopamine oxidation is visualized as green peaks at the bottom third of the color plot. Right-hand panel shows change in oxidative currents over time at the peak sensitivity to dopamine for this electrode (+0.71 V), converted to dopamine concentration using its calibration factor. The inset panel is the background subtracted cyclic voltammogram for this response (current versus applied potential) taken 0.8 s after reward delivery, which is consistent with the electrochemical signature for dopamine (r^2=0.95).

Nature Neuroscience: doi:10.1038/nn.2460

Data analysis

Animals included in the study contributed two side-counterbalanced recording sessions for a given cost-benefit contingency (e.g., 4 pellets assigned to left lever, 1 pellet assigned to right lever in one session, and 4 pellets assignment to right lever, 1 pellet assigned to left lever in another). These sessions were treated as a within-subjects repeated measure. All other factors were treated as between-subjects measures, even though in seven rats, the same animals contributed to the data from separate cost-benefit contingencies. Analysis of extracellular dopamine concentration was restricted to the period of 2 seconds following cue onset, prior to reward delivery, on post-criterion forced trials. Dopamine signals on trials where no lever-press response was made within the 10 second response window were excluded to ensure that the data only reflected trials where animals had perceived the cues.

Voltammetric data analysis was carried out using software written in LabVIEW. Electrochemical signals were low-pass filtered at 2,000Hz. Individual cyclic voltammograms (electrochemical current-voltage plots) were used for chemical identification. The current at the peak dopamine oxidation potential across successive voltammograms was used for dopamine quantification. Any noise spikes of >±1.5 nA greater than the signal in both 100ms time-bins before and after the time point were manually removed, and the data were smoothed using a 0.5-s moving average.

Estimation of dopamine concentration

The main statistical tests in this work were within-session comparisons and so are unaffected by determination of the absolute concentration of dopamine. Nonetheless, it is more intuitive to present these data as estimated dopamine concentrations rather than raw voltammetric currents. For histological verification of recording sites, electrolytic lesions were made via the recording electrode as described above. This procedure renders electrodes unsuitable for post-implantation assessment of sensitivity. Thus, electrode sensitivity was estimated by extrapolation from a cohort of electrodes (matched to background current) through which a lesion was not made. Control electrodes (n=15) were implanted for an equivalent period to experimental electrodes and underwent post-implantation assessment of sensitivity *in vitro*. Electrode background currents generated during recording sessions were used to verify comparability to those obtained during electrode calibration. Notably, conversion to dopamine concentration did not change any of the reported effects, either within or between sessions.

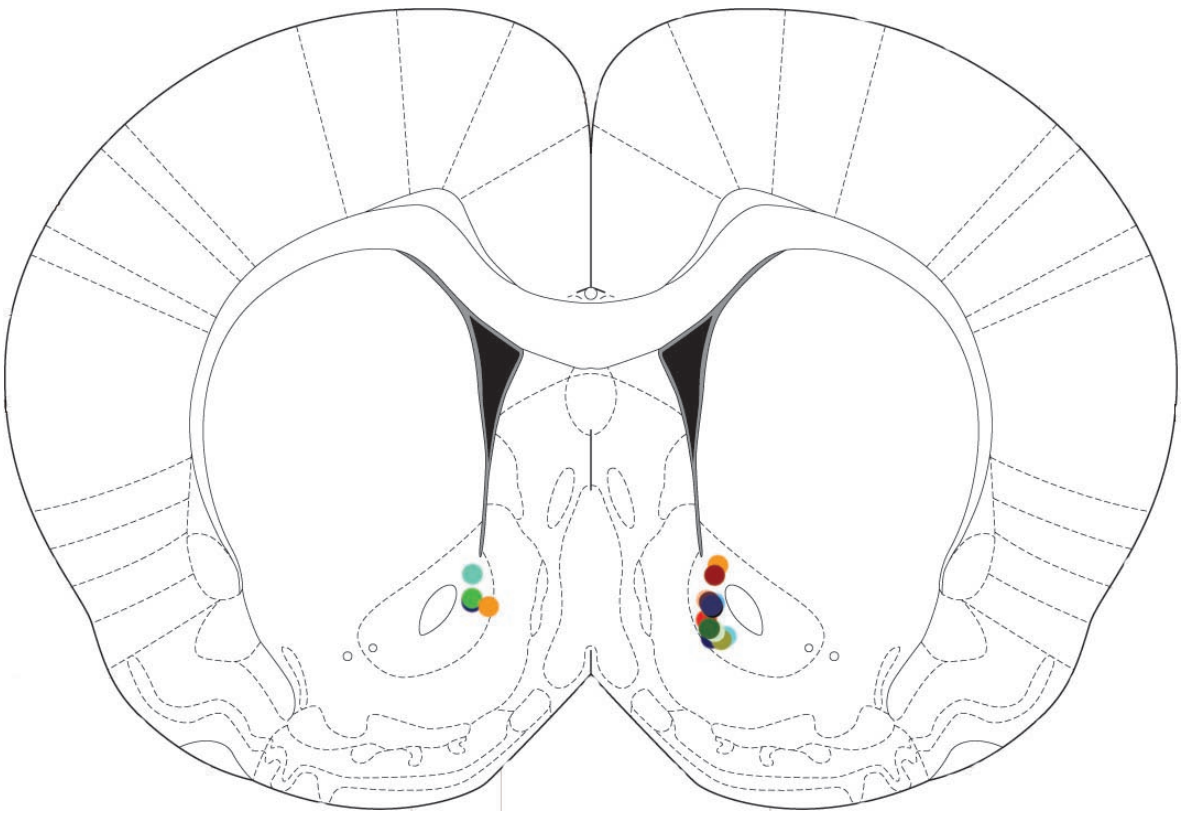
Nature Neuroscience: doi:10.1038/nn.2460

Supplementary Results

Histology

Following completion of the experimental sessions, animals were anesthetized with ketamine/xylazine (100 mg/kg) and the recording site was marked by making a small electrolytic lesion at the electrode tip by passing a current (~70μA) through the carbon fiber microelectrode for twenty seconds. Animals were subsequently perfused transcardially with physiological saline and then with 4% paraformaldehyde in phosphate-buffered saline, before the brains were removed and post-fixed in a paraformaldehyde solution. The brains were then placed in 30% sucrose solution in phosphate-buffered saline for 48 h, flash frozen, and sectioned coronally (30 μm). All sections were mounted and stained with cresyl violet.

The majority of recording locations were in the medial NAcc (Supp. fig. 3). The electrode for one animal was in the adjacent ventromedial shell and for another was on the boundary of the core and the shell, and both were therefore removed from the analyses. Nonetheless, their voltammetric data was similar to those from the NAcc and so their removal did not markedly alter the pattern of results described in the main text (data not shown).

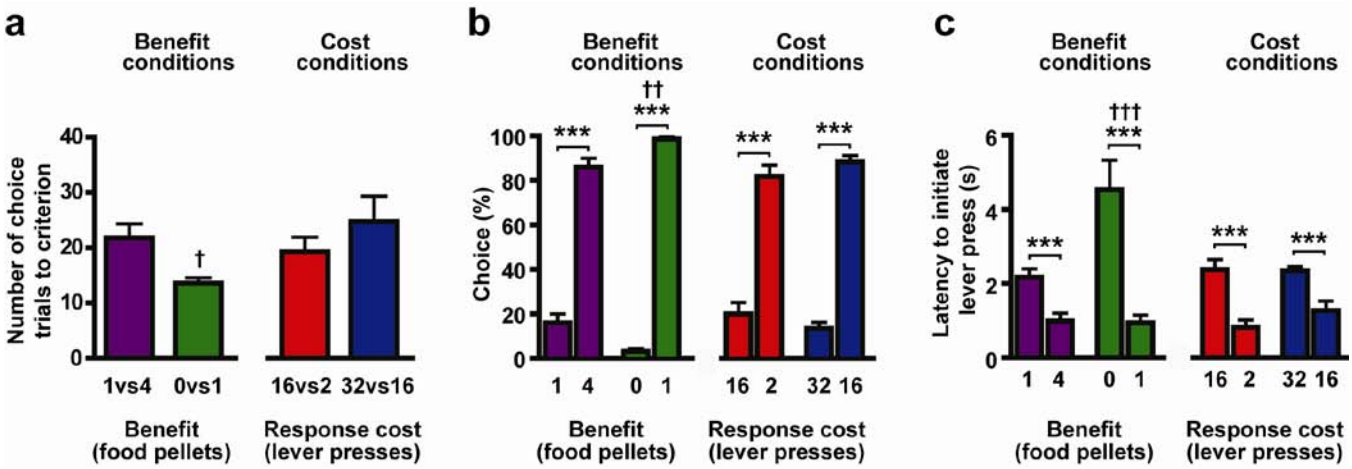


Supplementary figure 3. Locations of the carbon fiber recording electrodes within the NAcc.

Nature Neuroscience: doi:10.1038/nn.2460

Behavior: voltammetric recording sessions

Three behavioral metrics were analyzed from recording sessions: (i) number of trials to criterion, (ii) post-criterion choice allocation and (iii) response latencies on post-criterion forced trials. All three measures demonstrated that animals reliably preferred the option with greater benefits or lower cost in each condition. There was no significant difference in the number of trials to behavioral criterion between the two cost conditions or the high-benefit condition (Mann-Whitney test: all comparisons $p>0.3$, n=10-12 sessions; Supp. fig. 4a). However, rats took significantly fewer trials to reach criterion when the reward was reduced to zero ($p<0.05$ versus other conditions, n=10; Supp. fig. 4a). All animals continued to allocate their choices preferentially to the option with the higher benefit or lower cost in post-criterion trials (Supp. fig. 4b). There was no difference in choice performance between either cost condition or the high-benefit condition but the preference was strongest in the low-benefit condition (main effect of group: $F_{3,17}=5.37$, $p=0.01$; post-hoc tests, $p<0.05$ for all comparisons of lower benefit session with the other sessions; all other comparisons $p>0.16$). Choice performance can also be reliably indexed by reaction times on forced trials[1]. Post-criterion, rats were significantly faster to select the higher benefit or lower cost option in all conditions ($F_{1,17}=52.75$, $p<0.001$), though this difference was again particularly marked when reward on the alternative was reduced to zero ($F_{3,17}=5.22$, $p=0.01$) with animals responding significantly slower to the cue predicting zero rewards ($p<0.01$; Supp. fig. 4c). Based on these three behavioral criteria, we conclude that the utility of reward options were modulated in both benefit and cost conditions (i.e. increased utility conferred to the option with greater benefit or lower cost).

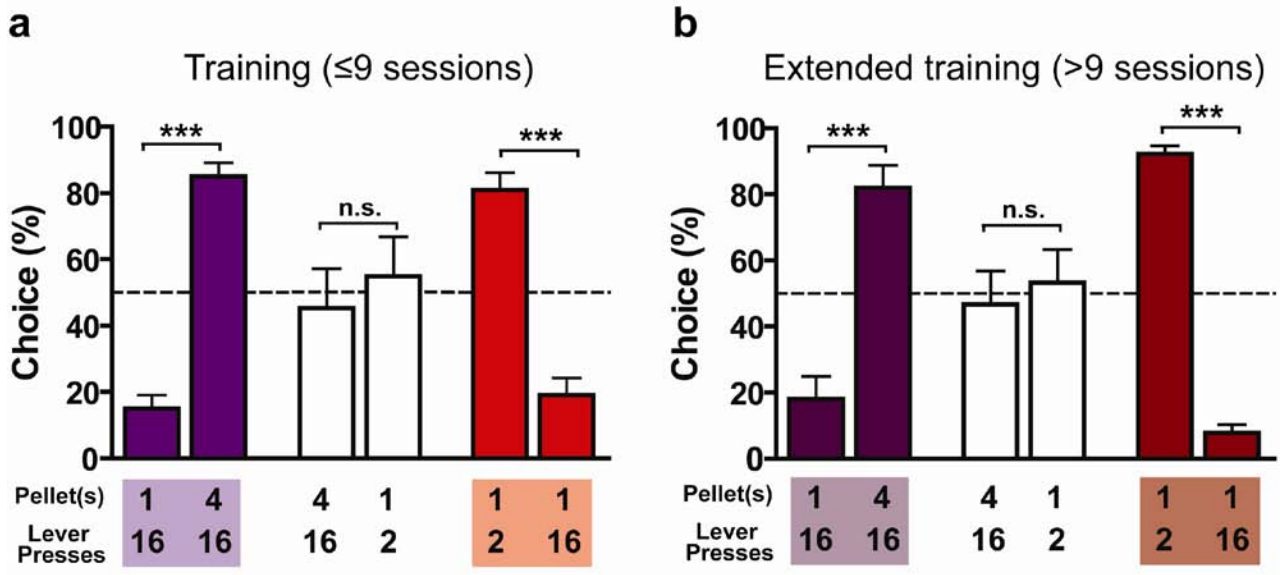


Supplementary figure 4. Behavioral data. **(a)** Number of choice trials to reach the criterion of choosing the high reward / low cost option on ≥75% of trials. **(b)** Percentage of choices allocated to each option post-criterion. **(c)** Latency to make an initial lever press response on post-criterion forced trials. (For all data: n=5-6; ***$p<0.001$ versus paired option; †$p<0.05$, ††$p<0.01$, †††$p<0.001$ versus all other conditions).

Nature Neuroscience: doi:10.1038/nn.2460

Behavior: test of utility equivalence between the high-benefit and low-cost contingencies

As demonstrated above, increased utility of a reward was conferred by increasing the benefit (number of food pellets) or decreasing the response cost (number of lever presses). These manipulations altered behavior in a comparable manner as assessed by learning rate, response latency and choice when presented concurrently with a reference reward. To test directly whether high-benefit and low-cost conditions yield equivalent utility, these two conditions were compared directly in a behavioral experiment. Animals that had ≤9 sessions of training in all conditions chose the higher-benefit at roughly the same rate as the lower-cost (p=0.71, Supp. fig. 5a). Similarly, extensively trained animals still chose either option at the same rate (p=0.76; Supp. fig. 5b). This indifference of choice leads us to believe that regardless of animals' experience and by extension, lever-pressing aptitude, the utility conferred by the two manipulations are approximately equal. Therefore, different patterns of dopamine release between the high-benefit and low-cost conditions are not a result of differences in conferred utility.



Supplementary figure 5. Intercontingency choices after <9 training sessions. **(a)** Choice behavior following training with ≤9 previous exposures to the contingencies. Bars represent percentage of choices allocated to: reference option vs higher reward (n=6), higher reward vs lower effort (n=5), lower effort vs reference effort (n=5). **(b)** Choice behavior following "extended" training regime which consisted of >9 previous exposures to one of the contingencies. Bars represent percentage of choices allocated to: reference option vs higher reward (n=3), higher reward vs lower effort (n=5), lower effort vs reference effort (n=4). (n.s., not significant; ***, p<0.0001).

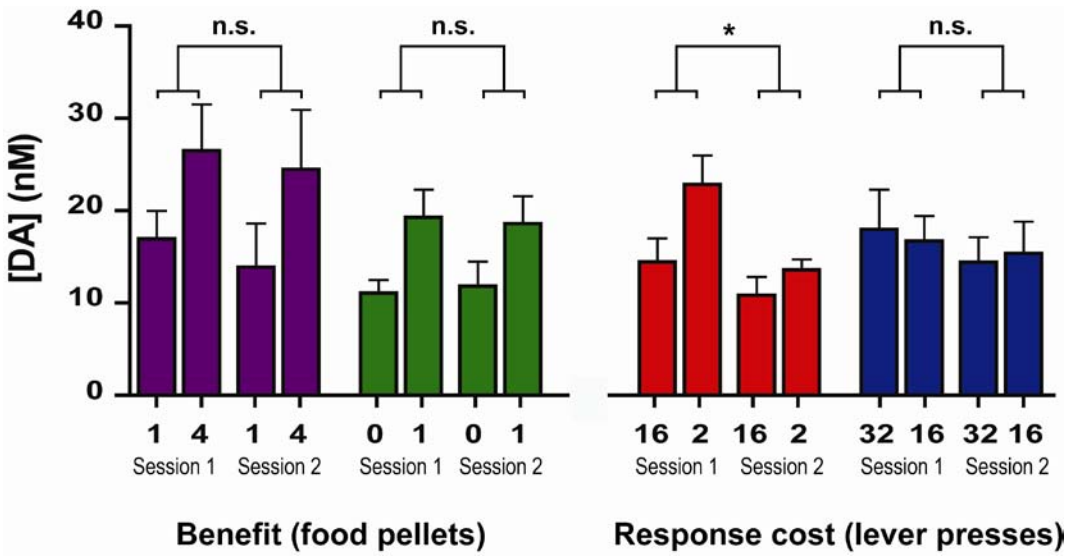
Neurochemistry: lever-side and session effects

For a given cost-benefit contingency, two recording sessions were collected from each animal. Therefore, it is possible that there could be differences in dopamine release between these

Nature Neuroscience: doi:10.1038/nn.2460

repeated sessions. The sessions were counterbalanced so that for half the animals the reference option was on the left lever for the first session and the right for the second, and vice versa for the other half of the animals. Because of this counterbalanced design, any effects of contingency side can be disambiguated from effects of session order. When the data were analyzed by presentation side (for a given cost-benefit contingency, dopamine release for an option on the right lever versus dopamine release to the same option on the left lever), this factor did not significantly affect the magnitude of cue-evoked dopamine transmission for any condition (all $F<1.26$, $p>0.27$). Therefore, any changes in dopamine release across sessions cannot be attributed to the side on which a particular cost-benefit option was presented.

When the data were analyzed by session order (for a given cost-benefit contingency, dopamine release for a cue in the first session versus dopamine release to the same cue in the second session), this factor significantly affect the magnitude of cue-evoked dopamine transmission in the low-cost option, where dopamine release was attenuated on the second session ($p=0.03$). There was no session order effect for any other condition (all $F<0.78$, $p>0.42$; Supp. fig. 6).



Supplementary figure 6. Average cue-evoked peak dopamine for the first and second of two contingency-counterbalanced sessions. (n.s., not significant; *, $p<0.05$).

Neurochemistry: contextual framing

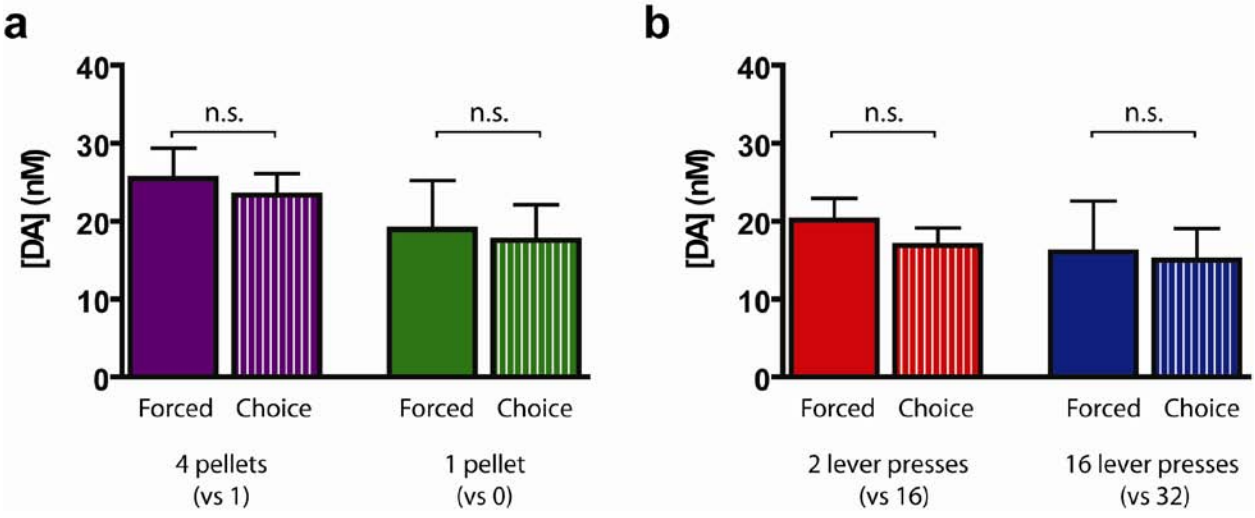
To test whether the reference option was regarded differently based upon context[9], dopamine release to the reference option on post-criterion forced trials across all conditions were compared. Regardless of whether the reference option conferred higher or lower utility, presentation of the

Nature Neuroscience: doi:10.1038/nn.2460

cue predicting the reference elicited similar release of dopamine (main effect of condition: $F_{3,17}$=0.786, p=0.52).

Neurochemistry: forced versus choice trials

While the focus of our investigation was on post-criterion forced trials, voltammetric data were also recorded (i) on choice trials and (ii) while the animal was learning to choose between the cost-benefit contingencies. When comparing the peak amount of dopamine on the post-criterion choice trials (where the high utility option was subsequently chosen) against the peak dopamine on high utility forced trials, there was no statistical difference between the cue-evoked dopamine on forced and choice trials (main effect of trial type or interaction between trial type and group: F<2.3, p>0.15; Supp. fig. 7). There were too few post-criterion low net value choice trials to gain a reliable estimate of changes in dopamine concentration. While this rules out that cue-evoked dopamine release reflects the average value of all available options, this data set cannot arbitrate between models which advocate that dopamine signals the value of the chosen option or the highest available value option[10,11].



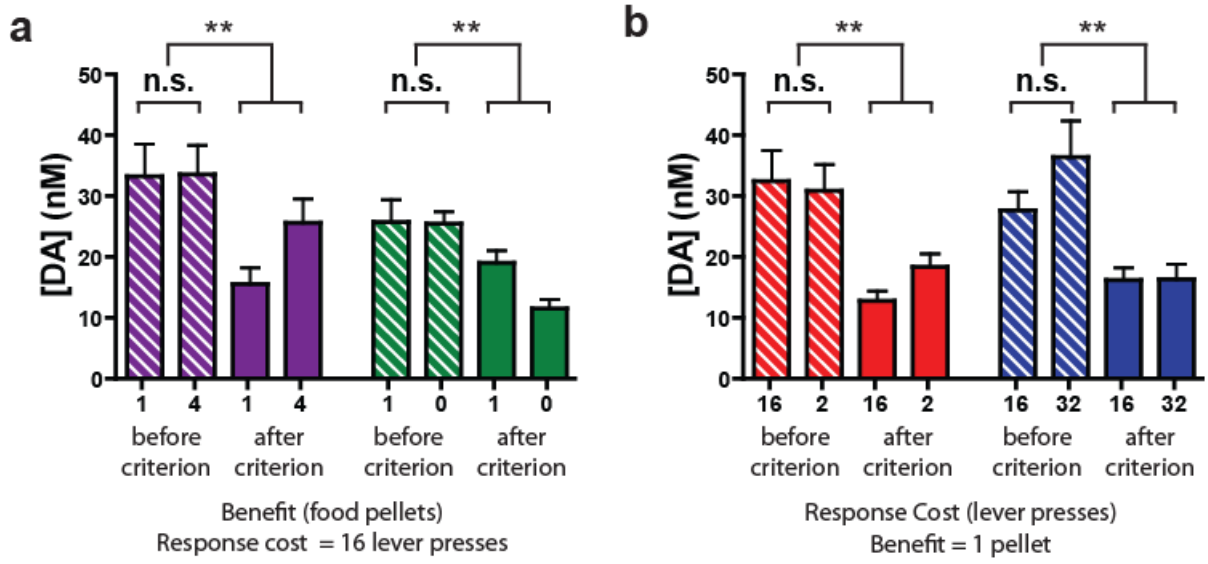
Supplementary figure 7. Comparison of post-criterion cue-evoked dopamine release on high utility forced trials and on choice trials where the high utility option was chosen in **(a)** benefit conditions; **(b)** cost conditions. There was no difference between the measured dopamine concentration on the forced and choice trials in any block (n.s., not significant).

Neurochemistry: within-session learning

Contingencies were assigned to the levers at the start of each session. We considered the animals had learned these assignments once they achieved a behavioral criterion where they chose one option in at least nine of the last twelve choice trials. To investigate how the dopamine

Nature Neuroscience: doi:10.1038/nn.2460

signal changed as contingencies were learned, we compared cue-evoked dopamine transmission at the start of the session (first two blocks of forced trials) to that after the behavioral criterion was reached. At the start of the session, there was no significant difference in dopamine release between cues in any of the conditions (main effect of cost-benefit contingency or interaction between contingency x group: both $F<3.2$, $p>0.09$; Supp. fig. 8). When these early trials were directly compared with the post-criterion data, for all conditions, there was a significant interaction between contingency and learning, indicating that selective cue encoding develops as the criterion is reached ($F_{1,17}=26.38$, $p<0.01$). There was also a main effect of learning across all conditions, with dopamine concentrations higher at the start of the session ($F_{1,17}=28.99$, $p<0.01$), possibly reflecting a higher motivational state or additional novelty bonuses at the start of sessions[12-14].



Supplementary figure 8. Comparison of cue-evoked dopamine release at the beginning of sessions to post-criterion cue-evoked dopamine release. **(a)** Average cue-evoked dopamine to the first two blocks of forced trials (striped) and to trials after criterion (solid) when response cost is manipulated. **(b)** Average cue-evoked dopamine to the first two blocks of forced trials (striped) and to trials after criterion (solid) when the benefit was manipulated. (n.s., not significant; **, $p<0.01$).

Nature Neuroscience: doi:10.1038/nn.2460

References for Supplementary Materials

1. Walton, M.E., Kennerley, S.W., Bannerman, D.M., Phillips, P.E.M. & Rushworth, M.F. Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. *Neural Netw* **19**, 1302-1314 (2006).
2. Walton, M.E*., et al.* Comparing the role of the anterior cingulate cortex and 6-hydroxydopamine nucleus accumbens lesions on operant effort-based decision making. *Eur J Neurosci* **29**, 1678-1691 (2009).
3. Salamone, J.D., Correa, M., Farrar, A. & Mingote, S.M. Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology* **191**, 461-482 (2007).
4. Phillips, P.E.M., Robinson, D.L., Stuber, G.D., Carelli, R.M. & Wightman, R.M. Real-time measurements of phasic changes in extracellular dopamine concentration in freely moving rats by fast-scan cyclic voltammetry. *Methods Mol Med* **79**, 443-464 (2003).
5. Roitman, M.F., Stuber, G.D., Phillips, P.E.M., Wightman, R.M. & Carelli, R.M. Dopamine operates as a subsecond modulator of food seeking. *J Neurosci* **24**, 1265-1271 (2004).
6. Phillips, P.E.M., Stuber, G.D., Heien, M.L., Wightman, R.M. & Carelli, R.M. Subsecond dopamine release promotes cocaine seeking. *Nature* **422**, 614-618 (2003).
7. Schultz, W. Multiple dopamine functions at different time courses. *Annu Rev Neurosci* **30**, 259-288 (2007).
8. Garris, P.A., Collins, L.B., Jones, S.R. & Wightman, R.M. Evoked extracellular dopamine in vivo in the medial prefrontal cortex. *J Neurochem* **61**, 637-647 (1993).
9. Tobler, P.N., Fiorillo, C.D. & Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642-1645 (2005).
10. Morris, G., Nevet, A., Arkadir, D., Vaadia, E. & Bergman, H. Midbrain dopamine neurons encode decisions for future action. *Nat Neurosci* **9**, 1057-1063 (2006).
11. Roesch, M.R., Calu, D.J. & Schoenbaum, G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci* **10**, 1615-1624 (2007).
12. Martinez, V., Walton, M.E., Gan, J.O. & Phillips, P.E.M. Phasic dopamine transmission attenuates following reward devaluation in a dual-reward detection task. In *Monitoring Molecules in Neuroscience* (ed. P.E.M. Phillips, S.G. Sandberg, S. Ahn & A.G. Phillips) 265-266 (UBC Institute of Psychiatry, Vancouver, BC, 2008).
13. Kakade, S. & Dayan, P. Dopamine: generalization and bonuses. *Neural Netw* **15**, 549-559 (2002).
14. Redgrave, P. & Gurney, K. The short-latency dopamine signal: a role in discovering novel actions? *Nat Rev Neurosci* **7**, 967-975 (2006).

Nature Neuroscience: doi:10.1038/nn.2460