Dopamine prediction error responses integrate subjective value from different reward dimensions

Armin Lak[1,2], William R. Stauffer[1], and Wolfram Schultz

Department of Physiology, Development, and Neuroscience, University of Cambridge, Cambridge CB2 3DY, United Kingdom

Edited by Michael E. Goldberg, Columbia University College of Physicians and Surgeons, New York, NY, and approved December 24, 2013 (received for review November 20, 2013)

Prediction error signals enable us to learn through experience. These experiences include economic choices between different rewards that vary along multiple dimensions. Therefore, an ideal way to reinforce economic choice is to encode a prediction error that reflects the subjective value integrated across these reward dimensions. Previous studies demonstrated that dopamine prediction error responses reflect the value of singular reward attributes that include magnitude, probability, and delay. Obviously, preferences between rewards that vary along one dimension are completely determined by the manipulated variable. However, it is unknown whether dopamine prediction error responses reflect the subjective value integrated from different reward dimensions. Here, we measured the preferences between rewards that varied along multiple dimensions, and as such could not be ranked according to objective metrics. Monkeys chose between rewards that differed in amount, risk, and type. Because their choices were complete and transitive, the monkeys chose "as if" they integrated different rewards and attributes into a common scale of value. The prediction error responses of single dopamine neurons reflected the integrated subjective value inferred from the choices, rather than the singular reward attributes. Specifically, amount, risk, and reward type modulated dopamine responses exactly to the extent that they influenced economic choices, even when rewards were vastly different, such as liquid and food. This prediction error response could provide a direct updating signal for economic values.

electrophysiology | primate | behavioral economics | reinforcement learning

Prediction errors represent the difference between predicted and realized outcomes. As such they are an ideal way to learn through everyday experiences (1). These experiences include making value-based (economic) choices between different rewards and evaluating the outcome of such decisions. Some of the most common economic decisions we face are between rewards that lack a common quality for comparison. To facilitate consistent choices between them, such rewards should be evaluated on a common scale of value (2–4). Thus, an ideal way to facilitate and reinforce economic decisions is to encode the prediction error directly in terms of subjective value. Midbrain dopamine neurons encode a reward prediction error (5–7) that is sufficient to cause learning (8, 9). These neurons receive inputs from several brain areas that encode subjective value and project axons to every brain structure implicated in economic choice (10–21). Therefore, dopamine neurons are ideally positioned to broadcast a teaching signal that directly updates economic values.

Economic preferences between alternatives that vary in one attribute are easily determined by the magnitude of the attribute (22). For instance, larger rewards are preferred over smaller ones, high reward probability is preferred over low reward probability, and reward delivered after a short delay is preferred over the same reward delivered after a long delay. Subjective preferences can only be isolated from the underlying reward attributes for rewards that vary in more than one dimension. Although previous studies showed that dopamine responses reflected magnitude, probability, and delay (23–25), it remained unclear how dopamine responses would reflect individuals' subjective preferences among rewards that vary along multiple dimensions. These rewards cannot be ordered according to objective metrics; rather, the subjective rankings of such rewards can only be inferred by observing an individual's choices (26). If those choices are complete and transitive, then the individual behaved as if she was maximizing subjective value (26). Completeness demonstrates that the individual had preferences (or was indifferent) between all of the rewards, whereas the transitive property provides strong evidence that different reward attributes were integrated onto a common scale, and that choices were made by selecting the highest rank on that scale (27). Therefore, the activity of neurons that encode subjective value should reflect the transitive ordering of rewards that vary in more than one attribute.

Here, we varied the amount, risk, and type of rewards and used different behavioral economic paradigms to infer the subjective value of these rewards from monkeys' choices. These tests provided a means to quantify the influence of singular reward attributes on a common value scale. In close correspondence to the behaviorally defined reward value, the dopamine prediction error response reflected the integrated subjective value derived from different rewards rather than distinguishing between their attributes.

Results

Behavioral Measurement of Value. We measured subjective value by observing the choices two monkeys made between different reward predicting cues (Fig. 1*A*, Fig. S1*A*, and *Methods*). Each cue predicted black currant juice (juice 1), strawberry juice (juice 2; orange juice for monkey B), or an equiprobable ($P = 0.5$ each) gamble between juices 1 and 2 (Fig. 1*B*). Moreover, each cue predicted either a certain amount of juice (0.45 mL) or an equiprobable gamble between a small and large amount (0.3 or

Significance

Most real-world rewards have multiple dimensions, such as amount, risk, and type. Here we show that within a bounded set of such multidimensional rewards monkeys' choice behavior fulfilled several core tenets of rational choice theory; namely, their choices were stochastically complete and transitive. As such, in selecting between rewards, the monkeys behaved as if they maximized value on a common scale. Dopamine neurons encoded prediction errors that reflected that scale. A particular reward dimension influenced dopamine activity only to the extent that it influenced choice. Thus, vastly different reward types such as juice and food activated dopamine neurons in accordance with subjective value derived from the different rewards. This neuronal signal could serve to update value signals for economic choice.

Author contributions: A.L., W.R.S., and W.S. designed research; A.L. and W.R.S. performed research; A.L. and W.R.S. analyzed data; and A.L., W.R.S., and W.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]A.L. and W.R.S. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: arminlak@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1321596111/-/DCSupplemental.

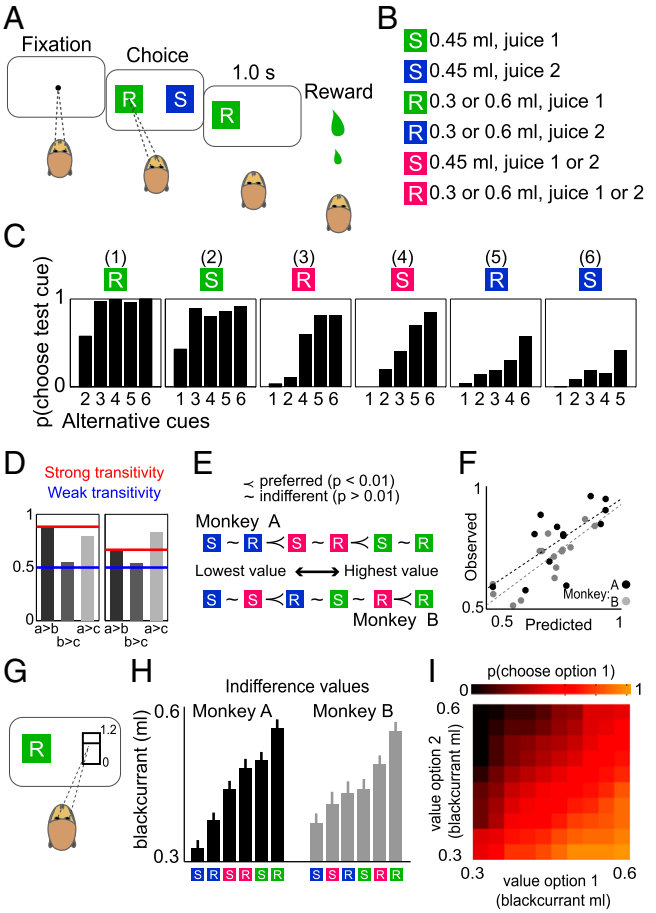

Fig. 1. Monkeys choose as if they are maximizing economic value gained from rewards. (*A*) Binary choice task. From left to right: the animal fixated on a spot to start the trial, chose one of two cues, and received the chosen reward. (*B*) Six cues predicting varying juice type and risk (R, risky magnitude: equiprobable gamble between 0.3 or 0.6 mL; S, safe magnitude, 0.45 mL; green, juice 1, blue, juice 2; red, equiprobable gamble between juice 1 and 2). (*C*) Choice probabilities in monkey A. Each plot indicates the probability of choosing the test cue (shown on the top of the plot) over alternative cues (other five cues used in the experiment). For display purpose, cues were ordered according to the overall probability with which they were selected. See Fig. S1*B* for monkey B. (*D, Left*) An example of choice probabilities across three cues that satisfies weak stochastic transitivity. (*D, Right*) An example of choice probabilities across three different cues that satisfies weak and strong stochastic transitivity. See Table S1 for all cue combinations. (*E*) The preference order of cues in both monkeys. ≺ indicates choice probabilities significantly greater than 0.5 ($P < 0.01$), suggesting choice preference. ~ corresponds to choice probabilities not statistically different from 0.5 ($P > 0.01$), suggesting choice indifference. (*F*) Correlations between observed and predicted choice probabilities (*SI Methods*). (*G*) PEST followed the same task sequence as the choice task shown in *A*. Over consecutive trials, animals chose between one of the cues and a "currency cue" that explicitly indicated a specific amount of black currant juice that varied trial by trial. The procedure terminated when the animal was indifferent between the two options. For a full description see Fig. S2 and *Methods*. (*H*) Indifference points indicating subjective values for each cue in units of black currant juice (averages from 25 and 40 PEST measurements per cue in monkey A and B, respectively). Error bars indicate SEM. (*I*) Choice probabilities measured in the binary choice task as a function of the estimated numerical value acquired through PEST. Data are combined from both animals and include all choices.

0.6 mL). Thus, in this experiment we used the combination of the animals' preference for different juices and risk levels to elicit a range of values in reward-predicting cues. The resulting choice ratios demonstrated that the animals had a clear bias in their selection probabilities and thus preferred some combinations of reward attributes over others (Fig. 1*C* and Fig. S1*B* for monkey B).

To infer economic value from choice probability, the choices must be complete and transitive (26–28). We had imposed nominal completeness by offering the animals choices between all pairs of cues. To measure actual completeness for all pairs of options a and b, we tested how close the choice behavior approached the ideal choice behavior defined by $P(a,b) + P(b,a) = 1$, where $P(a,b)$ indicates the probability of choosing option a over b (28). We found that the animals' average error rates in making a successful choice for a given pair of cues was 1% for all pairs. Importantly, none of the error rates was statistically different from the others ($P > 0.05$, ANOVA). Thus, both animals expressed well-defined preferences between all possible pairs of cues. Next, we quantified the transitivity of their choices. Stochastic transitivity comes in two strengths: weak and strong (*SI Methods*), and both strive to establish consistency among different binary choices (26). Thus, for options a, b, and c, if the individual probabilistically selects a over b, and b over c, she should then probabilistically select a over c. If $P(a,c)$ is greater than 50%, then the choices satisfy weak stochastic transitivity (Fig. 1*D*, *Left*). When $P(a,c)$ is greater than or equal to the maximum of $P(a,b)$ and $P(b,c)$, then the choices satisfy both the weak and the strong axioms (Fig. 1*D*, *Right*). The choices of both monkeys satisfied the weak axiom of stochastic transitivity in all 20 three-cue combinations ordered according to the animals' probabilistic choices (cue a over b, and cue b over c), and the strong axiom in 15 of 20 and 14 of 20 combinations, respectively (Table S1). Satisfying completeness and transitivity established that the animals' choices were consistent with an ordinal ranking of the cues explained by a monotonic, weakly increasing value function.

Monkey A preferred juice 1 over juice 2 and was risk-neutral. Monkey B preferred juice 1 over juice 2 but also preferred risky over safe options. Both animals associated the six cues with three levels of subjective value, although with different ordering in each animal. As such, although they preferred some of the cues to others, they showed choice indifference among some pairs of cues (Fig. 1*E*; $P < 0.01$, binomial test). Both animals showed stable behavior in their valuation over multiple sessions of testing (Fig. S1*C*). Moreover, the saccadic response times showed weak inverse relationship with the subjective value of the chosen cue (Fig. S1*D*).

To determine whether the animals assigned menu-invariant values to each cue, we examined the relationship between the strength of the stochastic preferences (27). We predicted the probability of choosing cue a over c from the observed probabilities of choosing cue a over b and cue b over c and then compared the predicted with the observed probability of choosing cue a over c in each combination (*SI Methods*). The predicted and observed choice probabilities were highly correlated (Fig. 1*F*; $\rho = 0.68$ and 0.82, $P < 0.01$ for monkeys A and B, respectively, Pearson's correlation). These results suggest that when making a choice between rewards the animals assigned stable and menu-invariant values to options and probabilistically selected the option with the highest value on that scale.

Because any monotonically increasing utility function would be consistent with the measured value rankings, these tests did not reveal numerical values. However, a numerical scale is vital to quantify the influences of individual reward properties on common scale value and provides a measure for analyzing neuronal data. To obtain such a scale, we used a psychometric choice method, parameter estimation by sequential testing (PEST) (29, 30). We used black currant juice as a common reference and measured the amount of this juice that was subjectively equivalent to the value of each cue (Fig. 1*G*, Fig. S2 *A* and *B*, and *Methods*). Therefore, milliliters of black currant juice at choice indifference provided a numerical estimate of value for each cue (Fig. 1*H* and Fig. S2 *C* and *D*). The choice probabilities from the binary choices reflected the values acquired psychometrically, suggesting that the two methods for measuring economic value were broadly consistent (Fig. 1*I*). The larger the difference was between the measured values, the greater was the probability of choosing the higher-valued option ($\rho = 0.90$, $P = 0.01$ and $\rho = 0.93$, $P = 0.007$ in monkeys A and B, respectively, Pearson's

correlation). These converging behavioral results allowed us to define the value of the reward-predicting cues on a common scale and in a convenient currency (milliliters of black currant juice).

Cue Responses. To unambiguously relate the neuronal responses to the value of specific rewards, we recorded from single dopamine neurons in a task in which the reward-predicting cues were displayed individually. Of 80 dopamine neurons tested with different risks and juice types, 77 responded to unpredicted reward and were further analyzed (Figs. S3 and S4 and *Methods*). Each trial started with a central fixation point (FP), to which the animal shifted its gaze. Subsequently, one of the six cues appeared in the center of the monitor for 1.5 s and extinguished upon reward delivery (Fig. 2*A* and *SI Methods*). Because the intertrial interval (ITI) duration was random and the animals could not predict trial onset, the reward prediction during the ITI was close to zero, whereas the value at the FP was the average value of all trial types. Thus, the prediction error at the FP was the constant difference between the average value of all trial types and the prediction of zero. When one of the cues appeared, the prediction shifted from the average value at the FP to the specific value at the cue.

Dopamine neurons displayed identical phasic activations at FP onset in all trial types that reflected the uniform prediction error at the FP (Fig. 2*B* and Fig. S5, $P > 0.2$, one-way ANOVA across trial types in both monkeys; see *SI Results* for saccadic response times to FP). Onset of the cue coincided with a robust dopamine response that reflected the value predicted by the cue. Cues with the lowest values elicited substantial neuronal depressions (Fig. 2*C*, *Left*), whereas cues with highest values induced pronounced activations (Fig. 2*C*, *Right*). Importantly, cues with similar values elicited indistinguishable neuronal responses, despite predicting rewards with different attributes (Fig. 2*C* and Fig. S6*A*). This pattern of neuronal activity was clearly visible at the population level (Fig. 2D and Fig. S6*A*). The dopamine neurons did not encode the objective properties of the upcoming reward (Fig. 2*D*, *Inset* and Fig. S6*B*). These observations suggested a close relationship between the neuronal encoding and the behavioral manifestation of subjective value.

We quantified the prediction error using the numerical estimation of the economic value in milliliters of black currant juice (Fig. 1*F*). For example, the size of the prediction error with the most preferred cue was 0.14 mL of black-currant juice (0.56 mL – 0.45 mL, Fig. 2*A*). The neuronal responses to cues were highly correlated with the numerical values (Fig. 2*E*). Cues eliciting a negative prediction error in value resulted in graded depressions, whereas cues eliciting a positive prediction error induced graded activations in dopamine neurons. This positive relationship with the prediction error in economic value was significant in two thirds of single neurons (Fig. 2*F*; $P < 0.05$, linear regression) and in the population of all sampled neurons in both animals (Fig. 2*G*; $R^2 = 0.27$ and 0.28, respectively, $P < 0.0001$). Other factors that might influence the activity of dopamine neurons include reinforcement history (7, 31) and motivational decline owing to satiety effects (32). However, the neuronal recordings took place after the animals were highly trained in a stable environment, and thus learning should be minimal. Accordingly, cues and rewards presented in the previous trials had negligible effects on the neuronal responses (Fig. S7 *A–D* and *SI Results*). Moreover, rewards accumulated over a session had only a small effect on the behavioral performance and on the neuronal responses to FP in monkey B (Fig. S7*E*, $P = 0.001$ and $P = 0.008$ for behavior and neurons, respectively, linear regression). Thus, the neuronal prediction error responses to the cues reflected the subjective value of the predicted reward.

In four trial types, juice type and risk varied orthogonally. We used receiver operating characteristics (ROC) analysis to quantify

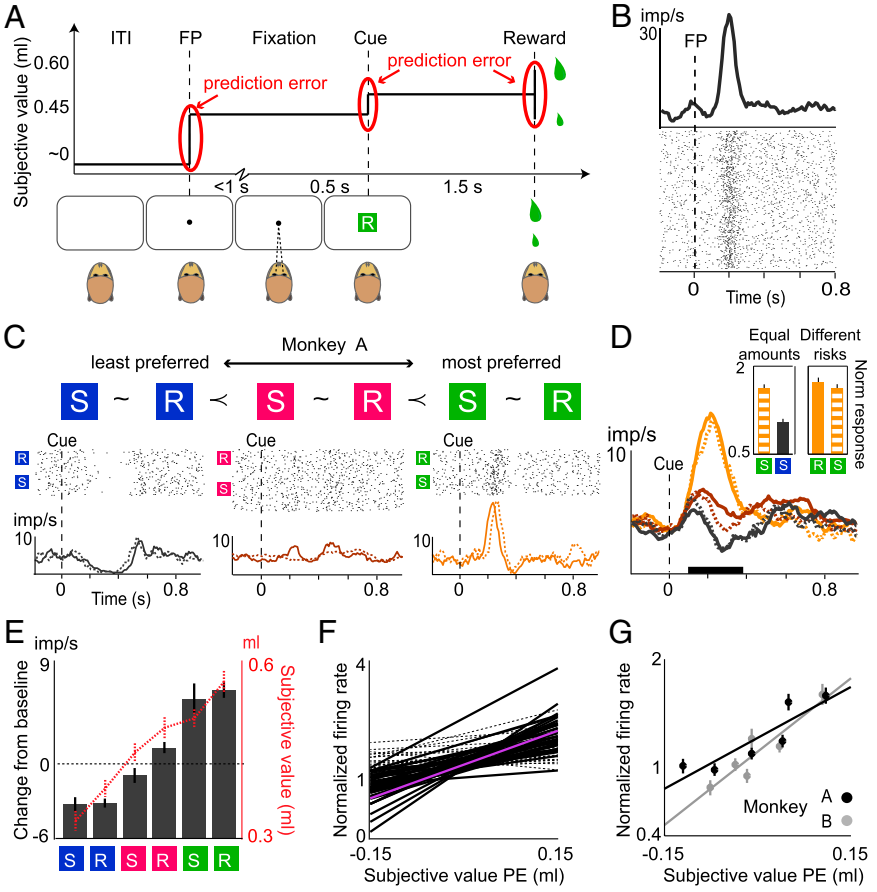

Fig. 2. The dopamine response encodes a prediction error in economic value. (*A*) Neuronal recording task and schematic representation of prediction errors during a trial. Following successful fixation for 0.5 s, one of the six cues was presented. Rewards ocurred 1.5 s after cue onset. (*B*) Responses of single dopamine neuron from monkey A to fixation spot (*Upper*: peristimulus time histograms (PSTH); *Lower*: rastergrams). These responses were identical regardless of trial type (Fig. S5). (*C*) Responses of single dopamine neuron shown in *B* to cue onset (*Upper*: cues and rastergrams; *Lower*: PSTH). Solid PSTH corresponds to R (risky) cue; dashed PSTH represents S (safe) cue. Responses to cues with the same subjective value are shown in the same panel. (*D*) Averaged neuronal population responses in monkey A. Line and color conventions are as in *C*. (*Inset*) The averaged neuronal population activity for singular reward attributes. The thick horizontal bar indicates the epoch used in statistical analysis. Data for monkey B are shown in Fig. S6. (*E*) Relationship between example neuron, shown in *C*, and numerical estimates of subjective value (Fig. 1*H*). Bars represent change in the firing rate in epoch indicated in Fig. 2*C*. Red line indicates average subjective values. Error bars are SEM across trials (black) and PEST measurements (red). (*F*) Linear regressions of normalized activities of all neurons (from both animals) onto subjective value prediction errors. Solid lines indicate regression slopes significantly different from zero. Nonsignificant regressions are shown by dotted lines. The purple line is the regression slope of the example neuron shown in *C*. (*G*) Regression of normalized neuronal population responses to cues onto subjective value prediction error.

the effect of these attributes on both the subjective values and the neuronal response to cues (*SI Methods*). In monkey A, this analysis revealed good discrimination between the distributions of values arising from the different juices but largely overlapping value distributions arising from different risk levels (Fig. 3*A*, *Left*; $P < 0.0001$ and $P > 0.3$ for juice type and risk, respectively, bootstrap test). Correspondingly, the neuronal responses in this animal were largely sensitive to juice type but not risk (Fig. 3*A*, *Right*; $P < 0.0001$ and $P > 0.5$ for juice type and risk, respectively, bootstrap test). In monkey B, the behavioral sensitivity to both juice type and risk was paralleled by the neuronal sensitivity to these two attributes (Fig. 3*B*, *Left*, behavior: $P < 0.0001$ and $P < 0.008$; Fig. 3*B*, *Right*, neurons: $P < 0.0001$ and $P < 0.0001$ for juice type and risk, respectively, bootstrap test). ROC analysis of individual neuronal responses revealed that 63% of all tested neurons reflected the behavioral sensitivity to different reward attributes in a subject-specific manner (Fig. 3*C*). This result mirrors the 66% of neurons whose activity was significant in single linear regressions on subjective value (Fig. 2*F*). Because monkey A was risk-neutral in the tested gambles, we could not conclude whether the dopamine responses integrated the preference for risk. Therefore, we measured risk preference in a different gamble with higher expected value and higher risk (equiprobable gamble between 0.1 and 1.2 mL). The subjective value of the gamble was significantly larger than the subjective value of a safe reward with equal expected value ($P < 0.0001$, Wilcoxon signed-rank test). Accordingly, the cue responses in a subset of five dopamine neurons were larger for the gamble than for the safe reward and thus reflected the higher subjective value of risky compared with safe rewards observed in behavioral tests (Fig. S8, $P = 0.0002$, Wilcoxon rank-sum test across trials). Together, these findings indicate that singular reward attributes modulated dopamine responses exactly to the extent that they influenced economic choices. Thus, the majority of dopamine neurons encode subjective value by integrating the behavioral weight of different reward attributes into a common scale of value.

To investigate whether results presented so far extend beyond liquid rewards, in a separate set of experiments we measured preferences between food and juice rewards of varying magnitude (*SI Methods*) and then recorded the responses from 20 additional dopamine neurons to cues that predicted these rewards. Binary choices among cues were stochastically transitive (in all possible cue combinations) and revealed that the tested food rewards were preferred to small juice rewards, but less preferred than large juice rewards (Fig. 4*A*, $P < 0.01$, binomial test). Psychometric estimation of the numerical subjective values correlated well with the binary choices ($\rho = 0.99$, $P < 0.0001$, Pearson's correlation). Dopamine responses scaled with these subjective values, independent of the identity or physical properties of these rewards, and demonstrated a highly significant positive relationship with prediction error in economic value in both single neurons (Fig. 4*B*, 15 of 20 neurons, $P < 0.05$, linear regression) and the population (Fig. 4*C*, $R^2 = 0.45$, $P < 0.0001$, linear regression). Thus, the dopamine responses reflected the common scale value of rewards with very different physical properties.

Reward Responses. We next investigated whether the dopamine responses to rewards reflected the preferences the animal has over different rewards. We tested this hypothesis by varying type and amount of the juice rewards, both together and separately. When there was an equiprobable gamble between amounts (0.3 or 0.6 mL) but juice type was fully predicted, the prediction error derived only from juice amount. Dopamine neurons showed the typical positive and negative prediction error responses for juice amount, independently of juice type (Fig. 5*A*). However, when neither juice type nor amount was fully predicted, the prediction error resulted from both attributes. Accordingly, the graded positive and negative prediction error responses reflected both juice type and amount (Fig. 5*B*). This influence of juice type on prediction error responses was also apparent in the population responses (Fig. S9*A*, $P < 0.0001$, bootstrap test) even when juice

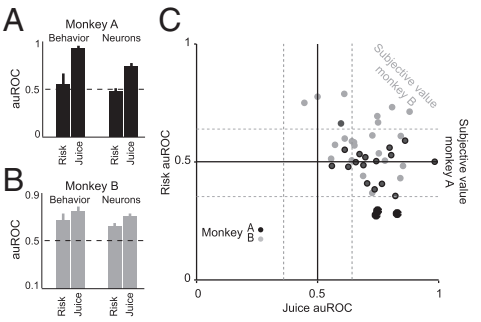

Fig. 3. Dopamine responses reflect the behavioral integration of reward attributes. (*A*) Areas under the ROC curve (auROC) for distributions of numerical subjective values (acquired through PEST) and neuronal responses with orthogonal variations of juice type and risk. An auROC >0.5 indicates higher subjective value (neuronal response) for preferred juice (compared with nonpreferred juice) or preferred risk (compared with nonpreferred risk). Error bars represent SEM across days of behavioral testing and neurons. (*B*) As in *A* for monkey B. (*C*) Scatter plot of auROC measures of individual neurons for juice type and risk. For display purpose, only neurons that were significant in the linear regression (Fig. 2*F*) are shown. Dotted lines indicate the 95% bootstrapped confidence intervals estimated using single-cell data.

amount was fully predicted (Fig. S9 *B* and *C*, $P < 0.0001$, Wilcoxon rank-sum test). Regression of neuronal responses to all tested rewards onto the prediction errors in economic value was highly significant in both animals (Fig. 5*C*, $R^2 = 0.17$ and 0.16, respectively, $P < 0.0001$). These data demonstrate that the prediction error response at the time of reward encodes the integrated value derived from reward type and amount. Thus, the reward responses mirrored the value encoding of the cue responses. As such they would be an eligible substrate for training economic value.

Discussion

This study investigated the relationship between dopamine prediction error responses and economic value. We demonstrated that the choices monkeys made among rewards of different types and different attributes satisfied the fundamental requirements

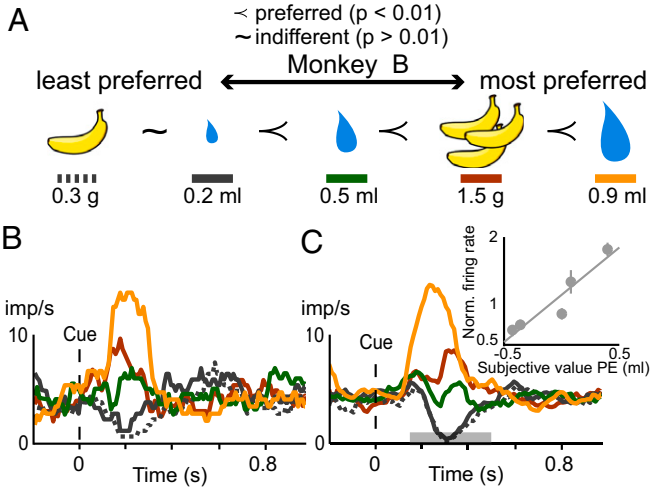

Fig. 4. Dopamine responses encode economic values of liquid and food rewards. (*A*) The preference order of monkey B among cues predicting different amounts of juice (black currant) or food (mashed mixture of banana, chocolate, and hazelnut). < and ∼ are as in Fig. 1*E*. (*B*) Response of single neuron (from monkey B) to the cues. Line and color conventions are as in *A*. (*C*) As in *B* for averaged neuronal population responses. (*Inset*) Regression of normalized neuronal population responses to cues onto subjective value prediction errors measured using PEST.

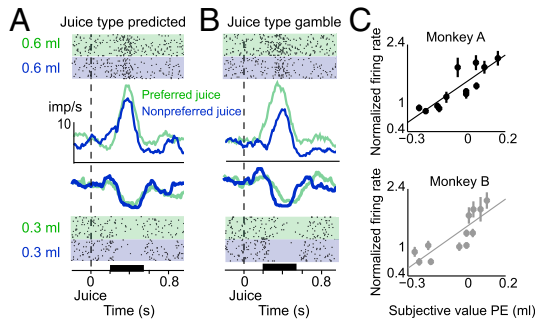

Fig. 5. Dopamine responses encode the economic value, not the identity or magnitude, of primary rewards. (*A*) Prediction error responses to juice amount when juice type was fully predicted. Rastergrams and PSTHs from the single neuron shown in Fig. 2. Green and blue PSTHs show responses to preferred and nonpreferred juice, respectively. (*B*) Prediction error responses to juice type and amount, when neither attribute was fully predicted. The same conventions as in *A*. Thick horizontal bars in *A* and *B* indicate the epoch used in the subsequent statistical analysis. (*C*) Regression of normalized neuronal population responses to rewards on common scale prediction errors.

of economic valuation, completeness and transitivity. Monkeys chose as if they were maximizing value on a common scale that reflected the integrated value of different reward attributes, such as amount, risk, and reward type. The dopamine neurons, at the time of the cues and at the time of the reward itself, encoded the integrated subjective values rather than changes in the objective properties of rewards. This coding scheme held not only when cues predicted different attributes of liquid rewards (amount, risk, and type) but also when the experiment involved rewards that were physically very different (liquid versus food). These neuronal responses paralleled the individual weightings used to construct value from reward attributes during choices.

We estimated subjective value with two independent techniques: binary choice that revealed the preferred option probabilistically over several trials and psychometric measurement of subjective equivalents that directly estimated the numerical values of the cues in units of one currency (29, 30). Binary choice is well tested, and the strong degree of choice transitivity provided a reliable ordinal value ranking of the rewards. The psychometric procedure provided a direct estimate of numerical value. When an individual is indifferent between two goods, then by definition those goods have the same numerical value. Hence, the indifference point of the psychometric procedure provided a simple estimate of numerical value in units of juice volume. The high degree of correlation ($\rho \geq 0.90$) between the choice probabilities and the numerical value estimates is strong evidence that our measurements were reliable. Thus, these measures provided a strong foundation for examining the encoding of prediction errors in the space of economic value.

We used a nonchoice task for the neuronal recording to attain the best isolation of prediction errors in value. Multiple lines of evidence suggest that the value of the cues was stable between choice (behavioral) and nonchoice (recording) tasks. First, the animals displayed the same ranking of cues regardless of the type of choice task (Figs. 1*I* and 4). Second, the preference of animals was stable over several days of testing (Fig. S1*C*). Third, and most importantly, the values were menu-invariant within the choice set (Fig. 1*F*). That is, the value of a cue was independent of the value of the other options. These results suggest that the economic values were, to a great extent, stable across time and behavioral context. The linear regression analysis demonstrated that the neuronal responses in the nonchoice task reflected the value function derived from behavioral choices (Figs. 2*G* and 4*C*). In other words, if phasic dopamine activity encodes economic value, then the value represented by dopamine neurons in a nonchoice context is consistent with the values used for choice.

Brain structures processing subjective value could provide the input used by dopamine neurons to compute prediction errors. Anatomical studies in rodents have demonstrated that dopamine neurons receive inputs from a variety of regions involved in reward processing, including the lateral hypothalamus, the amygdala, and the orbitofrontal cortex (OFC) (20). Neurons in the lateral hypothalamus code subjective reward value (33) and reflect internal states, such as satiety (34). These neurons receive inputs from the prefrontal cortex (PFC) (35), possibly conveying subjective value signals to dopamine neurons. Furthermore, GABAergic neurons of the ventral tegmental area, which directly inhibit dopamine neurons (36), could signal expected reward (37). Thus, this circuit could relay some of the observed subjective value signals to dopamine neurons (38, 39).

Given the evidence that dopamine neurons encode economic value, what role might this activity play in brain function? Dopamine neurons project axons to every structure that has been implicated in economic decision making. These include the striatum, OFC, medial prefrontal cortex, anterior cingulate cortex, amygdala, and parietal cortex (12). In these structures, dopamine is released most efficaciously by stimulation with a pattern that most closely resembles the phasic prediction error responses (40). The release of dopamine in downstream structures following valued events will affect the synaptic strength in these structures. Neurons in the striatum encode action values during economic choices (14). These signals are likely to be stored and updated through corticostriatal synapses on the dendrites of medium spiny neurons (41). Current views suggest that dopamine prediction error signals might differentially modulate the efficacy of these synapses in a time-dependent manner via subtypes of dopamine receptors (42). Accordingly, optogenetic stimulation of striatal dopamine D1 and D2 receptor-expressing neurons introduce opposing biases in the distribution of choices (43, 44). Thus, in this hypothetical model, the prediction error signal could be used to update the stored action values and thus affect subsequent choices.

Although dopamine neurons project axons to the prefrontal cortex, less is known about the role of their prediction error signal there. Most studies of frontal dopamine function infused D1 receptor (D1R) agonists and antagonists. These manipulations simulated changes in baseline dopamine levels produced by tonic dopamine activity, rather than dopamine transients produced by prediction error responses. Nevertheless, a recent study has demonstrated that disruption of D1R activity in PFC results in attenuation of value-based learning of cue-response contingencies and can lead to noneconomical choice patterns, such as perseveration (45). Furthermore, dopamine actions seem to extend beyond learning. D1R-mediated dopamine activity enhances the suppression of working memory fields in the PFC, sculpting and sharpening the response of the PFC network for greater task fidelity (46, 47). Moreover, the optimal dopamine levels seem to be task-dependent. Higher dopamine levels are useful for tasks that demand greater fidelity (47). Although dopamine neurons do not show enhanced activity during delay periods (48), the size of the cue-evoked prediction error response could function to set the PFC network to its optimal state, based on the subjective value of the task. Algorithmically, this would be akin to adjusting the slope of the softmax rule. Thus, dopamine prediction errors broadcast to the cortex might contribute to modulating economic behavior.

In summary, dopamine neurons encode prediction errors on a common scale of values derived from several reward attributes such as amount, risk, and type of reward. The demonstrated economic value coding would allow the dopamine signal to play a simple and rapid role in updating value-based decision mechanisms in target structures.

Methods

Animals. The Home Office of the United Kingdom approved all experimental procedures. Two male monkeys (*Macaca mulatta*) weighing 13.4 and 13.1 kg were used in the experiment. Neither animal had been used in any prior study.

Binary Choice Task. A central fixation spot indicated onset of each trial. After successful gaze fixation, two randomly drawn cues appeared to the left and right on the monitor (Figs. 1*A* and 4*A* and Fig. S1). The animal had 1 s to indicate its choice by shifting its gaze to the center of the chosen cue and holding it there for another 0.5 s. The chosen reward was delivered at offset of the chosen cue.

PEST. PEST was used to measure the amount of black currant juice that was subjectively equivalent to the different rewards. The rules governing the PEST procedure were adapted from Luce (30). Each PEST sequence consisted of several consecutive trials during which one of the cues was presented as a choice option against the currency cue. The currency cue consisted of a horizontal bar on a vertical axis (Fig. 1*G* and Fig. S2*A*). The height of the bar was a safe and explicit indicator of the volume of the black currant juice. The procedure converged by locating currency offers on either side of the true indifference value. The indifference value was measured by averaging the last two currency offers. We repeated the PEST procedure several times for each of the cues, which allowed us to compare the distribution of indifference values acquired for each cue (Fig. S2*D*) and measure the area under the ROC curve (auROC, Fig. 3).

Identification and Recording of Dopamine Neurons. We recorded the extracellular activity of single dopamine neurons within the substantia nigra and in the ventral tegmental area. We localized the positions relative to the recording chamber using X-ray imaging and functional properties of surrounding cell groups (Fig. S3). We identified discharges from putative dopamine neurons using the following classic criteria: (*i*) polyphasic initially positive or negative waveforms followed by a prolonged positive component, (*ii*) relatively long durations (>2.5 ms measured at 100 Hz high-pass filter), and (*iii*) irregular firing at low baseline frequencies (fewer than eight spikes per second). Most neurons that met these criteria showed the typical phasic activation after unexpected reward (Fig. S4), which was used as a fourth criterion for inclusion. We did not test the responses of neurons to unexpected aversive stimuli. However, if any of the dopamine neurons encoded motivational salience, rather than reward value, they should have responded to unexpected reward (as a motivationally salient event) and thus have been sampled in our recordings (49). In total, we recorded 100 neurons and 96 met the criteria listed above and were used in our data analysis.

Analysis of Neuronal Data. The analysis of neuronal data used defined time windows that included the major positive and negative response components following fixation spot onset (100–400 ms), cue onset (80–340 ms and 150–500 ms in animals A and B, respectively), juice delivery (200–550 ms), and unexpected juice outside of the task (100–400 ms). Control time windows had identical durations and preceded immediately each respective task event. For normalization, in each neuron we divided the neuronal activity in each time window by the ensemble average activity of the neuron in the control window.

ACKNOWLEDGMENTS. We thank Christopher Harris, Camillo Padoa-Schioppa, Shunsuke Kobayashi, Martin O'Neill, Fabian Grabenhorst, Kelly Diederen, and Raymundo Baez Mendoza for discussions. This work was supported by the Wellcome Trust, European Research Council, and National Institutes of Health Caltech Conte Center.

1. Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA).
2. Montague PR, Berns GS (2002) Neural economics and the biological substrates of valuation. *Neuron* 36(2):265–284.
3. Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* 9(7):545–556.
4. Padoa-Schioppa C (2011) Neurobiology of economic choice: A good-based model. *Annu Rev Neurosci* 34:333–359.
5. Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275(5306):1593–1599.
6. Nakahara H, Itoh H, Kawagoe R, Takikawa Y, Hikosaka O (2004) Dopamine neurons can represent context-dependent prediction error. *Neuron* 41(2):269–280.
7. Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47(1):129–141.
8. Tsai HC, et al. (2009) Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* 324(5930):1080–1084.
9. Steinberg EE, et al. (2013) A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci* 16(7):966–973.
10. Williams SM, Goldman-Rakic PS (1993) Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. *Cereb Cortex* 3(3):199–222.
11. Lynd-Balta E, Haber SN (1994) The organization of midbrain projections to the ventral striatum in the primate. *Neuroscience* 59(3):609–623.
12. Lewis DA, et al. (2001) Dopamine transporter immunoreactivity in monkey cerebral cortex: Regional, laminar, and ultrastructural localization. *J Comp Neurol* 432(1):119–136.
13. Sugrue LP, Corrado GS, Newsome WT (2004) Matching behavior and the representation of value in the parietal cortex. *Science* 304(5678):1782–1787.
14. Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. *Science* 310(5752):1337–1340.
15. Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. *Nature* 441(7090):223–226.
16. Matsumoto M, Matsumoto K, Abe H, Tanaka K (2007) Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci* 10(5):647–656.
17. Pesaran B, Nelson MJ, Andersen RA (2008) Free choice activates a decision circuit between frontal and parietal cortex. *Nature* 453(7193):406–409.
18. Kennerley SW, Behrens TEJ, Wallis JD (2011) Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat Neurosci* 14(12):1581–1589.
19. Louie K, Grattan LE, Glimcher PW (2011) Reward value-based gain control: Divisive normalization in parietal cortex. *J Neurosci* 31(29):10627–10639.
20. Watabe-Uchida M, Zhu L, Ogawa SK, Vamanrao A, Uchida N (2012) Whole-brain mapping of direct inputs to midbrain dopamine neurons. *Neuron* 74(5):858–873.
21. Peck CJ, Lau B, Salzman CD (2013) The primate amygdala combines information about space and value. *Nat Neurosci* 16(3):340–348.
22. Green L, Myerson J (2004) A discounting framework for choice with delayed and probabilistic rewards. *Psychol Bull* 130(5):769–792.
23. Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299(5614):1898–1902.
24. Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. *Science* 307(5715):1642–1645.
25. Kobayashi S, Schultz W (2008) Influence of reward delays on responses of dopamine neurons. *J Neurosci* 28(31):7837–7846.
26. Luce RD, Suppes P (1965) Preference, utility, and subjective probability. *Handbook of Mathematical Psychology*, eds Luce RD, Bush EE, Galanter E (Wiley, New York), Vol 3, pp 252–410.
27. Luce RD (1959) *Individual Choice Behavior: A Theoretical Analysis* (Wiley, New York).
28. Davidson D, Marschak J (1958) Experimental tests of a stochastic decision theory. Technical Report No. 17 (Stanford Univ, Stanford, CA).
29. Taylor MM, Creelman CD (1967) PEST: Efficient Estimates on Probability Functions. *J Acoust Soc Am* 41(4A):782–787.
30. Luce RD (2000) *Utility of Gains and Losses: Measurement-Theoretic and Experimental Approaches* (Lawrence Erlbaum Assoc, Mahwah, NJ).
31. Bromberg-Martin ES, Matsumoto M, Nakahara H, Hikosaka O (2010) Multiple timescales of memory in lateral habenula and dopamine neurons. *Neuron* 67(3):499–510.
32. Rossi MA, Fan D, Barter JW, Yin HH (2013) Bidirectional modulation of substantia nigra activity by motivational state. *PLoS ONE* 8(8):e71598.
33. Ono T, Nakamura K, Nishijo H, Fukuda M (1986) Hypothalamic neuron involvement in integration of reward, aversion, and cue signals. *J Neurophysiol* 56(1):63–79.
34. Burton MJ, Rolls ET, Mora F (1976) Effects of hunger on the responses of neurons in the lateral hypothalamus to the sight and taste of food. *Exp Neurol* 51(3):668–677.
35. Ongür D, An X, Price JL (1998) Prefrontal cortical projections to the hypothalamus in macaque monkeys. *J Comp Neurol* 401(4):480–505.
36. Tan KR, et al. (2012) GABA neurons of the VTA drive conditioned place aversion. *Neuron* 73(6):1173–1183.
37. Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N (2012) Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482(7383):85–88.
38. Hong S, Jhou TC, Smith M, Saleem KS, Hikosaka O (2011) Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. *J Neurosci* 31(32):11457–11471.
39. van Zessen R, Phillips JL, Budygin EA, Stuber GD (2012) Activation of VTA GABA neurons disrupts reward consumption. *Neuron* 73(6):1184–1194.
40. Gonon FG (1988) Nonlinear relationship between impulse flow and dopamine released by rat midbrain dopaminergic neurons as studied by in vivo electrochemistry. *Neuroscience* 24(1):19–28.
41. Reynolds JN, Hyland BI, Wickens JR (2001) A cellular mechanism of reward-related learning. *Nature* 413(6851):67–70.
42. Shen W, Flajolet M, Greengard P, Surmeier DJ (2008) Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* 321(5890):848–851.
43. Kravitz AV, et al. (2010) Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* 466(7306):622–626.
44. Tai L-H, Lee AM, Benavidez N, Bonci A, Wilbrecht L (2012) Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat Neurosci* 15(9):1281–1289.
45. Puig MV, Miller EK (2012) The role of prefrontal dopamine D1 receptors in the neural mechanisms of associative learning. *Neuron* 74(5):874–886.
46. Williams GV, Goldman-Rakic PS (1995) Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature* 376(6541):572–575.
47. Vijayraghavan S, Wang M, Birnbaum SG, Williams GV, Arnsten AFT (2007) Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nat Neurosci* 10(3):376–384.
48. Schultz W, Apicella P, Ljungberg T (1993) Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J Neurosci* 13(3):900–913.
49. Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459(7248):837–841.