Current Biology, Vol. 15, 594–602, April 12, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.cub.2005.02.052

Single Units in the Pigeon Brain Integrate Reward Amount and Time-to-Reward in an Impulsive Choice Task

Tobias Kalenscher,[1,*] Sabine Windmann,[1] Bettina Diekamp,[1] Jonas Rose,[2] Onur Güntürkün,[1] and Michael Colombo[2]
[1]Department of Biopsychology
Institute of Cognitive Neuroscience
GAFO 05-618
Ruhr-University Bochum
Universitätsstraße 150
44780 Bochum
Germany
[2]Department of Psychology
Goddard Laboratories
University of Otago
Post Office Box 56
95A Union Street
Dunedin
New Zealand

Summary

Background: Animals prefer small over large rewards when the delays preceding large rewards exceed an individual tolerance limit. Such impulsive choice behavior occurs even in situations in which alternative strategies would yield more optimal outcomes. Behavioral research has shown that an animal's choice is guided by the alternative rewards' subjective values, which are a function of reward amount and time-to-reward. Despite increasing knowledge about the pharmacology and anatomy underlying impulsivity, it is still unknown how the brain combines reward amount and time-to-reward information to represent subjective reward value.

Results: We trained pigeons to choose between small, immediate rewards and large rewards delivered after gradually increasing delays. Single-cell recordings in the avian Nidopallium caudolaterale, the presumed functional analog of the mammalian prefrontal cortex, revealed that neural delay activation decreased with increasing delay length but also covaried with the expected reward amount. This integrated neural response was modulated by reward amount and delay, as predicted by a hyperbolical equation, of subjective reward value derived from behavioral studies. Furthermore, the neural activation pattern reflected the current reward preference and the time point of the shift from large to small rewards.

Conclusions: The reported activity was modulated by the temporal devaluation of the anticipated reward in addition to reward amount. Our findings contribute to the understanding of neuropathologies such as drug addiction, pathological gambling, frontal lobe syndrome, and attention-deficit disorders, which are characterized by inappropriate temporal discounting and increased impulsiveness.

*Correspondence: tobias.kalenscher@ruhr-uni-bochum.de

Introduction

Impulsive choice—that is, the preference for a small, immediate reward over a large, delayed one—is a key element in a wide range of pathological behaviors, such as drug addiction, pathological gambling, attention-deficit/hyperactivity disorder, and prefrontal dysfunction. The crucial deficit seems to involve a reduced sensitivity to delayed response outcomes in relation to immediate ones [1–3]. Although knowledge about the pharmacology and neuroanatomy of impulsiveness is expanding, its neural basis is still poorly understood [1–9].

A range of classic psychological studies on impulsive choice behavior of pigeons reveals that the preference for small, immediate or large, delayed rewards is determined by their relative subjective reward values. The subjective value of a reward is not only a function of its amount, but is also inversely proportional to the delay between response and reward [8, 10–13]. Mazur [11] described the relation between reward value, amount, and delay-to-reward with a function in which the subjective value of a reward decreases hyperbolically with increasing delay duration:

$$V = \frac{A}{1 + \gamma D} \quad (1)$$

where V is the subjective reward value, A is the fixed reward amount, D is the delay, and γ is a discount factor that describes the individually different impact of delay on reward value and determines the slope of the decay function.

This equation describes how time-to-reward and reward amount are combined into subjective reward value. Numerous studies indicate that single units in the mammalian orbitofrontal cortex (OFC) and dorsolateral prefrontal cortex process the constituents of this computation; that is, neurons provide the neural mechanism for the estimation of interval time [14], their activity can be modulated by the amount of an expected reward [15, 16], and they encode the relative reward value [17]. However, the neural mechanisms by which reward amount and time-to-reward are combined and translated into subjective reward value are unknown.

To address this issue, we trained pigeons to choose between a small, immediate reward and a large reward that was initially delivered after the same short delay but was delayed further as the session progressed. This schedule resulted in a typical within-session preference shift (PS) from the large (delayed) to the small (immediate) reward. We examined single-neuron activity in the Nidopallium caudolaterale (NCL) and related the neural responses to the animals' reward preference. The NCL is a pallial multimodal association structure that is suggested to be functionally equivalent to the mammalian prefrontal cortex [18–21] and has recently been related to the temporal organization of choice behavior [22]. Note that the NCL was named Neostriatum caudolaterale prior to the revision of the nomenclature [21].

Our main goal was to find single neurons that integrate delay-to-reward and reward-amount information to represent subjective reward value as expressed in Equation 1. More precisely, if single neurons are indeed modulated by subjective reward value, then according to Equation 1, the following hypotheses about the activity of single units involved in the processing of subjective reward value should hold:

(1) For constant delays and reward amounts across several trials, the subjective reward value and, hence, neural activity should remain constant.
(2) For constant reward amounts but increasing delays, the subjective reward value decays. Accordingly, neural activity of the same neuron should decrease with increasing delay.
(3) For constant delays but different reward amounts (reward A > B), the subjective value of reward A is larger than the value of reward B. Consequently, neural activity related to the choice of reward A should be higher than activity related to B.

In this article, we report that single neurons in the NCL show evidence of integrating reward amount and delay-to-reward information. We conclude that their activity was modulated by subjective reward value, as predicted by these hypotheses. Furthermore, the integrated neural response covaried with the time point of the preference shift from large, delayed rewards to small, immediate ones.

Results

Procedure

Eight pigeons were trained in the delayed-reward choice paradigm depicted in Figure 1. They had the choice between two pecking keys associated with different reward amounts and delay times. A fixed ratio response on one key led to a small reward (2 s access to food) that was delivered after a fixed short delay (always 1.5 s between response and delivery), and a fixed ratio response on the other key resulted in a large reward (4 s access to food) that was also delivered after 1.5 s at the beginning of a session. As the session progressed, however, the delay preceding the large reward was increased successively in a blockwise fashion in individually different increments. The increments were determined for each pigeon separately to account for interindividual differences in impulsiveness (see the Supplemental Data available with this article online for details).

Behavioral Performance

The pigeons showed a within-session PS (defined as the $block_i$ in which the large reward is chosen in $\geq 50\%$ of the trials, whereas it is chosen in <50% of the trials in $block_{i+1}$) that is typical for this design. Figure 2A shows a representative example of the animals' behavior in one single session. This pigeon predominantly chose the large reward when the delay was small, but once the delay exceeded an individual tolerance limit, its preference shifted to the small, short-delayed reward. This biased preference pattern was confirmed by the group data of the three PS-preceding and the three PS-following blocks, averaged across all animals and aligned to each individual PS (Figure 2B). Note that the three PS-preceding blocks will be labeled –3, –2, and –1 in relation to individual PS occurrence throughout this paper unless otherwise stated; accordingly, the three blocks following the PS will be labeled 1, 2, and 3.

In general, there was large interindividual variability in the maximum delay tolerance because some pigeons were more impulsive than others (see Table S1). The mean PS occurred between the third and fourth block, and the mean maximum delay tolerance (DT) was 10.4 s (standard error of the mean: 0.8; range: 1.5–48 s).

Neurons Show Delay Activity Related to Reward Anticipation

We recorded single-unit activity from 159 NCL neurons. One hundred fourteen cells showed significantly increased activity in reward anticipation during the delay period (71.7%, all $p < 0.05$, paired t test; for details on further task-related activations, refer to Supplemental Data). Figure 3A displays a representative neuron with sustained delay activity. Typically, these units showed an initial short phasic peak at delay onset and then gradually increased their activity until it stabilized at a final level.

To compare the level of the sustained response activation across the blocks, we examined the ranked percent change of delay activity in a time window of 1000–1500 ms after delay onset as a measure of activation magnitude—that is, toward the end of the maximum delay duration that all conditions had in common and where the final level of delay activity had already been reached (see gray area in Figure 3).

Selection of Neurons of Interest

We hypothesized the existence of single units that encode the subjective reward value as a function of delay length and reward amount. The hypotheses outlined above allowed derivation of quantitative predictions of expected neural-activity pattern. We selected units of interest from all sustained delay units if they met the following criteria:

(1) Before PS: $A_{-3} > A_{-2} > A_{-1}$ with $A_{-3} >> A_{-1}$
(2) After PS: $A_1 \cong A_2 \cong A_3$

where A_i is the mean activation magnitude in the indexed block. ">>" was defined as "a difference larger than two standard errors," and "$\cong$" was defined as "within the range of one standard error."

Of the 114 neurons with sustained delay activity, 22 had to be excluded prior to criterion testing because pigeons showed no or no definite PS in the recorded session. Of the remaining 92 units, 14 (15.22%) met these criteria.

It was necessary to show that the probability of a false-positive detection was < 0.05 to reject the hypothesis that our criteria were met by chance. We therefore performed a bootstrap analysis to determine the likelihood of detecting 14 out of 92 neurons whose activity pattern was task unrelated and fully random [23]. The analysis was done for each of the 92 neurons separately by reassigning each neural response during the delay and the baseline phase to a randomly selected

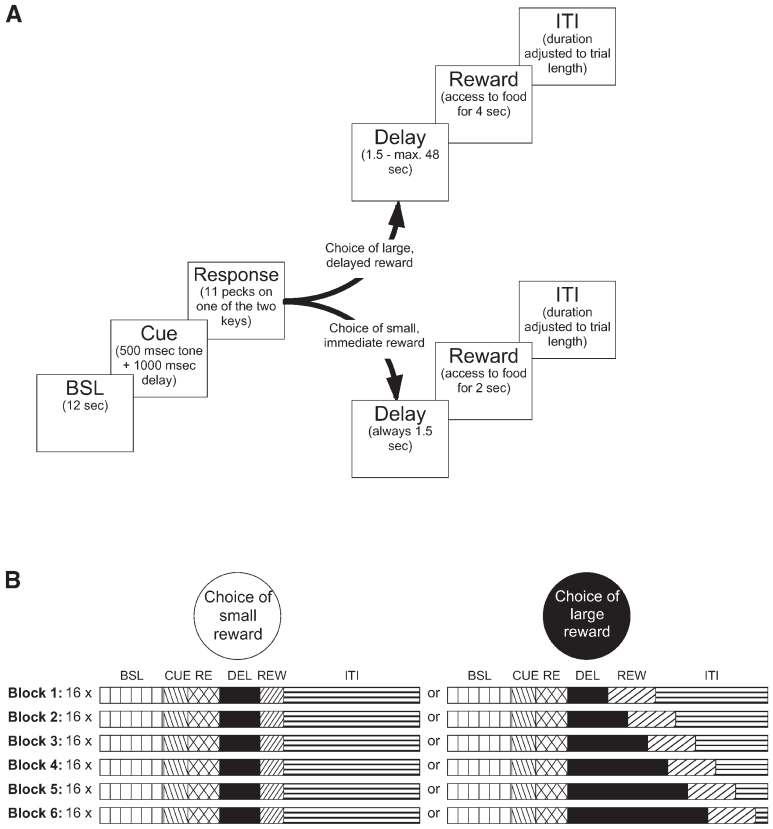

Figure 1. Delayed-Reward Choice Task

Sequence of intervals within a trial (A) and block design (B). Each trial started with a baseline (BSL) interval, followed by a cue phase (CUE). By responding on one of the two keys during the RESPONSE phase (RE), pigeons chose between a small, short-delayed REWARD (REW) and a large reward preceded by increasing DELAYS (DEL, incremented in individually different steps). Total trial length within one session was always constant and independent of choice and variations of delay and reward times. The length of the intertrial interval (ITI) was individually adjusted on a trial-by-trial basis depending on delay and reward duration. A session consisted of a succession of a varying number of blocks of 16 trials, including the first two trials, which were forced choice trials (1 × large, delayed and 1 × small, short-delayed reward in randomized order). Delays preceding large rewards were increased between the blocks. A session was terminated three blocks after a PS or after a maximum of ten blocks (note that this figure displays a session of six blocks).

trial with a uniformly distributed randomizer. After reassignment, the activation-magnitude pattern was computed as described above. This was done for every neuron, and we determined how many of the 92 shuffled data sets met our criteria. This was repeated 1000 times. Across the 1000 repetitions, our criteria were met on average in 3.81 cases (± 0.06 SEM, 4.1%), and 14 hits (15.22%) were met in none of the 1000 draws. The probability of false-positively detecting 14 out of 92 neurons was therefore $p < 0.001$. We concluded that the observed data pattern far exceeded what would have been expected by chance.

Single Neuron Activity Is Modulated by Time-to-Reward and Reward Amount

Figure 3 shows an example of a neuron that met the above criteria. The mean firing rate of this unit shows that, when the pigeon chose the large reward before the PS (Figure 3A), the level of sustained delay activity was maximal in block –3, in which the delay was minimal, and decreased across the blocks as the delay preceding large rewards increased. In small-reward trials after the PS (blocks 1 to 3, in which small rewards were delivered after an invariantly short delay; Figure 3B), the same neuron showed no evidence of a systematic variation in discharge rates across the blocks. Hence, this unit's activity was modulated by the duration of the delay.

Statistical analysis of the population response on the basis of all cells (Figure 4A) confirmed that neural activation magnitude indeed significantly decayed across the PS-preceding blocks (only large-reward choices; activation in block –3 > block –2 > block –1; all $p < 0.01$; Kolmogorov-Smirnov test). Moreover, there was no significant difference in activation magnitude between the blocks after the PS (only small-reward choices; block 1 versus block 2, block 1 versus block 3, and block 2 versus block 3; all $p > 0.2$; Kolmogorov-Smirnov test). Thus, consistent with our first hypothesis and our selection criteria, the neural response rate remained unchanged across trials when delay and amount were both invariant (post-PS blocks 1 to 3), and, consistent with our second hypothesis, the neural response rate decreased with increasing delay and constant reward amount (pre-PS blocks –3 to –1).

The analysis reveals, on the one hand, that the reported units adjusted their response rate to interval duration. On the other hand, this explanation alone is insufficient to fully capture the observed response pattern, as shown by the comparison of trials that differed only in reward amount and not in delay duration. If neural responses were solely modulated by delay duration, one would expect no difference between the activation in post-PS trials (small reward, short delay of 1.5 s) and block –3, in which the delay preceding the large reward was likewise 1.5 s (large reward, same short delay). However, we found that large reward activity in block –3 was significantly higher than the small reward activity in all post-PS blocks (only the eight sessions in which the delay in block –3 was 1.5 s; see Sup-

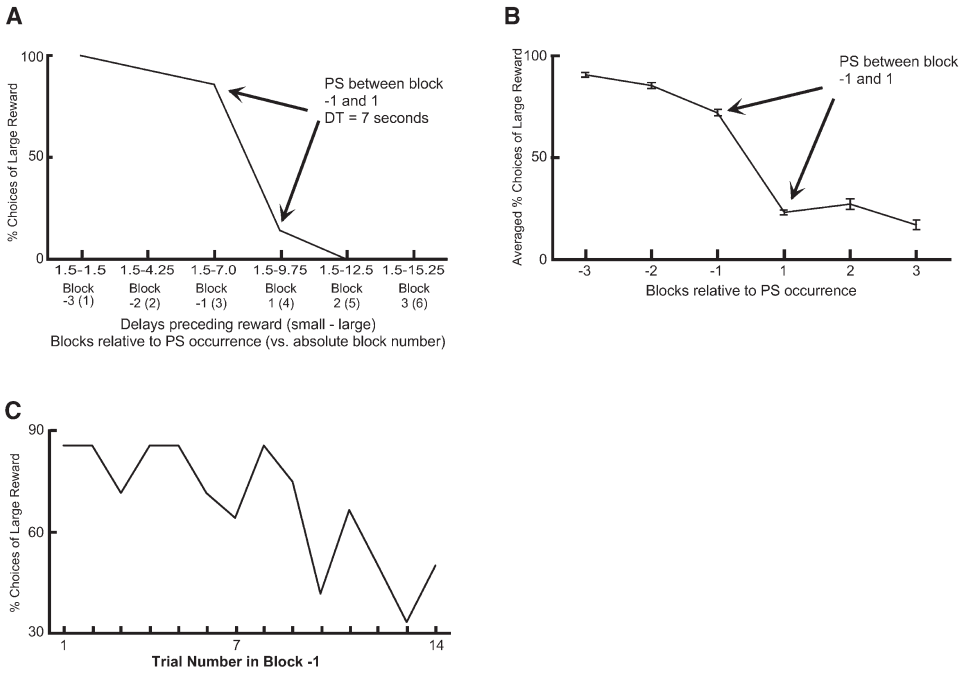

Figure 2. Behavioral Results

(A) Example of delay-dependent preference for large rewards of one pigeon. This pigeon preferred the large reward until the individual delay-tolerance limit (DT) of 7 s was reached, then shifted its preference and went on to almost exclusively opt for the small, short-delayed reward. The x axis depicts delay combinations (in s) as well as the absolute and relative block numbers with regard to PS occurrence.

(B) Mean large-reward choice frequency (± SEM), averaged across all sessions and pigeons and aligned to each individual preference shift (only the three blocks preceding and following the individual PS are displayed). Sessions with lacking or unclear PS were excluded from the group data.

(C) Percentage, shown for each trial in block –1, of sessions in which pigeons chose the large reward (calculated from the 14 sessions in which the neurons of interest were recorded). The frequency of large-reward choices decreased significantly across the trials (regression analysis). Such a trial-dependent decrease in choice frequency was not observed in any of the other blocks. The preference shift starts at about trial number eight. The fact that it did not occur immediately at the beginning of the block presumably results from the fact that pigeons needed to sample the current delay/reward-amount configuration for some trials before changing the preference bias.

plemental Data; activity in block –3 versus block 1, block 2, and block 3; all $p < 0.01$; Kolmogorov-Smirnov test; Figure 4B). Therefore, delay duration alone cannot fully account for the variation in response magnitude; instead, the cells additionally encoded reward amount. This conclusion is consistent with the third hypothesis, which states that for identical delay durations, the magnitude of the neural activity related to large-reward choices should be higher than activity related to small-reward choices.

The Relative Activation Magnitude Correlates with the Pigeons' Preference for the Large, Delayed or Small, Immediate Reward

Most interestingly, we found that changes in the neural activity pattern coincided with the occurrence of the PS: The activation related to large-reward choices in block –3 was significantly higher than the activation of small-reward choices in all post-PS blocks (block –3 versus all post-PS blocks; all $p < 0.001$; Kolmogorov-Smirnov test), whereas the activation in block –2 did not significantly differ from any of the post-PS blocks (block –2 versus all post-PS blocks; all $p > 0.44$; Kolmogorov-Smirnov test). Block –1 was the first block in which the large-reward-related response magnitude significantly dropped below the activation magnitude related to the small-reward choices (block –1 versus all post-PS blocks; all $p < 0.05$; Kolmogorov-Smirnov tests).

A detailed behavioral analysis of the choice distribution within each block revealed that the PS actually occurred within block –1, in which the frequency of large-reward choices significantly decreased across trials (standardized $\beta = -0.795$; $R^2 = 0.632$; $F[1,12] = 20.65$; $p < 0.001$; regression analysis). This decrease was not observed in any of the other blocks, –3, –2, 1, 2, and 3 (all $F[1,12] < 0.45$; all $p > 0.52$; regression analysis), in which pigeons showed a highly consistent preference bias. This suggests that block –1 was the only block in which pigeons gradually shifted their preference from the large to the small reward (Figure 2C). Because block –1 was also the first block in which the delay activity related to large-reward choices dropped below the activity related to small-reward choices, we conclude that the difference between large- and small-reward-related activity coincided with the time point of the PS occurrence.

In conclusion, the neural discharge pattern was modulated by a combination of delay duration and reward amount. We propose that this pattern resulted from the

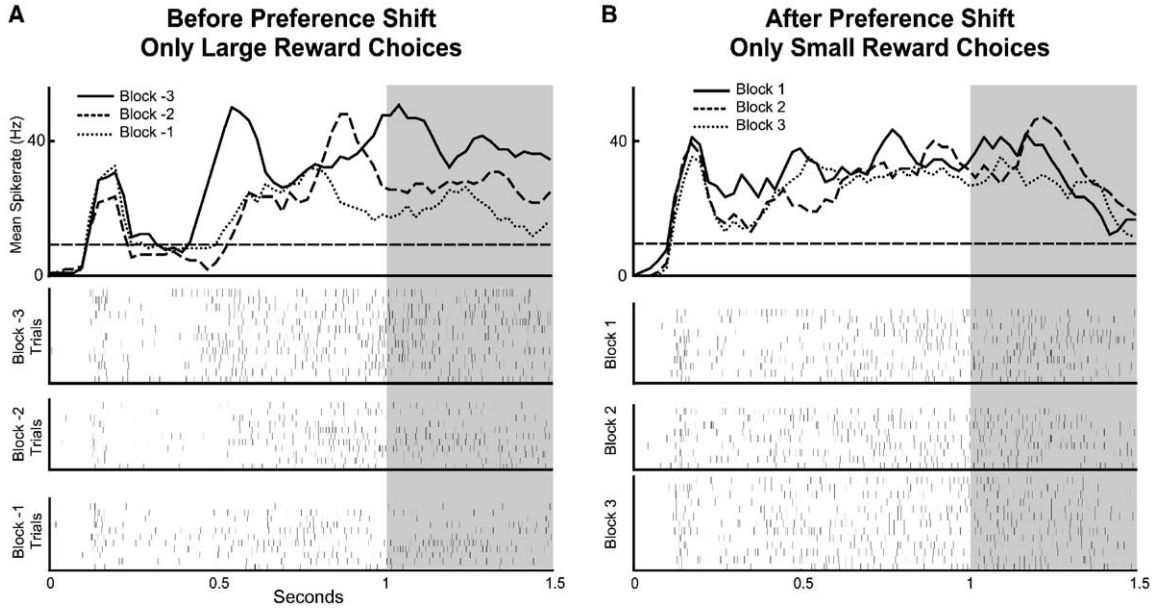

Figure 3. One Unit Encoding the Subjective Reward Value

The mean filtered firing rate (mean peristimulus time histograms, bin size 25 ms, averaged across all trials per block and then smoothed with a moving average filter, width of 4 bins) and raster plots are shown for one neuron during the first 1500 ms of the delay in the three blocks preceding the PS ([A], large-reward choices) and three blocks following the PS ([B], small-reward choices). The horizontal dotted line marks the averaged baseline spike rate (mean spikes per bin). The gray area delineates the window used for statistical analysis (see Figure 4A).

(A) This neuron exhibits the two characteristic response features typical of delay neurons: a short, phasic response with a latency of approximately 100 ms and a duration of approximately 100 ms and a long, sustained response with varying latency and a duration of at least 1000 ms. Spike rate during the sustained response was highest three blocks prior to the PS, where the delay was minimal and then decreased across blocks –2 and –1 as the delay duration increased.

(B) The same neuron shows no systematic and block-dependent variation in discharge rate. (A) and (B) illustrate the delay dependence of the neuron's response rate: When the delay increases, the activity decreases, and when the delay is constant, the activity is constant, too.

integration of delay and reward amount information yielding a composite subjective reward value. Moreover, the difference between neural responses associated with large-reward choices and small-reward choices correlated with preference bias and time point of PS occurrence. (Note: In this analysis, neural activation values were classified into blocks –3 and –1 in sessions in which the PS occurred already after two blocks, and data in block –2 were considered missing. However, exactly the same significance pattern was found when neural data were assigned to block –2 and –1, with data in block –3 considered missing. Also, the same significance pattern is found when sessions are included in which the PS occurred after only one block.)

The Neural Response Pattern Can Be Better Approximated by a Hyperbolic Model than by Linear or Exponential Models

For support of our conclusion, it is important to show that the hyperbolic model in Equation 1 is an adequate approximation of the actual data, in contrast to more simple equations, such as a linear function or an exponential function. We therefore contrasted two alternative models with the hyperbolic model. We expected that the observed neural data would be better explained by

$$V = \frac{A}{1 + \gamma D} + c \quad (2)$$

which adds a scaling constant c to Equation 1, than by the linear function

$$V = \gamma DA + c \quad (3)$$

or by the exponential function

$$V = Ae^{(-\gamma D)} + c \quad (4)$$

with the same variable labels as in Equation 1. For each of the 14 neurons, we individually fitted functions of Equations 2, 3, and 4 to the observed response magnitude values. The fit was done for each neuron separately by feeding delay and reward-amount information from every relevant block into the equation. The individual decay factor γ and the scaling constant c were then estimated for each function and each neuron with a least-square method to best fit the activation values of that neuron. Figure 4C illustrates for an exemplar neuron that the hyperbolic and exponential models provided a better approximation of this neuron's responses than the linear model. To compare the goodness of fits, we computed the Akaike information criterion (AIC) across all neurons and for each model. In general, the lower the relative AIC value, the better the fit. Results showed that for the given dataset, the hyperbolic

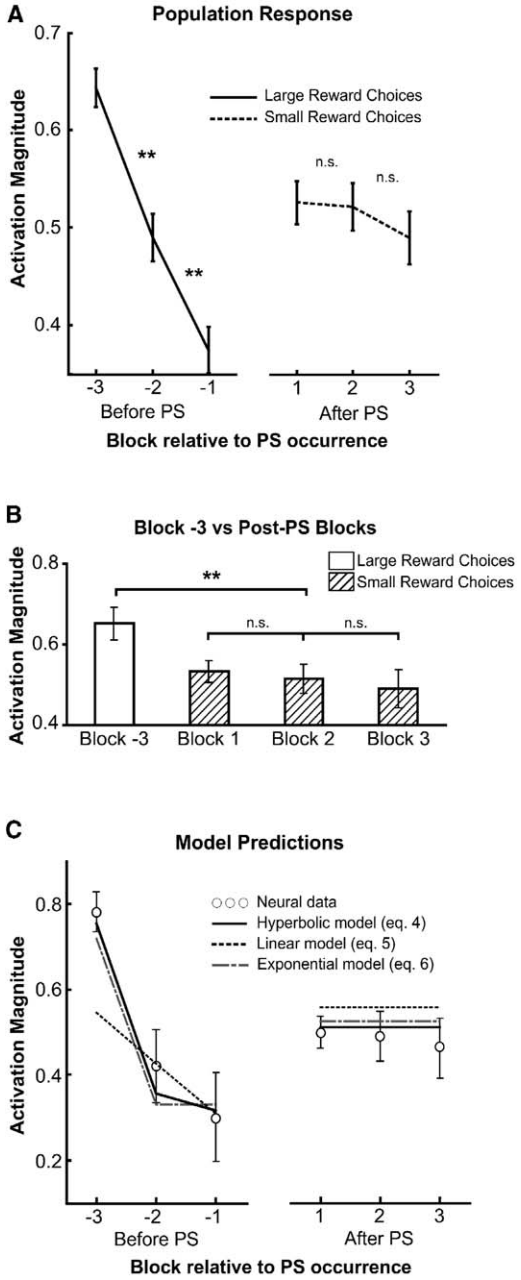

Figure 4. Single Neurons Encoding Delay Length (A) and Reward Amount (B) and Model Predictions for an Exemplar Neuron (C)

(A). This figure shows the activation pattern, averaged across all 14 neurons of interest (± SEM). The solid line represents the magnitude of neural activation in relation to large-reward choices preceding the PS, and the dotted line represents the neural activation in relation to small-reward choices following the PS. As indicated in Figure 3, the neural response magnitude significantly decreased across the blocks preceding the PS when delay length increased, and it remained roughly constant across the blocks that followed the PS when delay duration was constant.

(B). This panel shows the neural activity in only those sessions in which the delay in block –3 was 1.5 s (means ± SEM). The blank block represents the activation magnitude in block –3 trials, and the striped blocks represent the activation magnitudes in trials in the post-PS blocks 1, 2, and 3. This figure thus compares trials with equal delay durations (always 1.5 s delay in all blocks) but different reward amounts (large in block –3 and small in blocks 1, 2, and 3). The response magnitude in block –3 was significantly higher than the response magnitudes in the blocks following the PS, indicating that in addition to delay length, the neurons also encoded reward amount.

(C). The hyperbolic nature of the delay-dependent decay becomes more evident when an exemplar single neuron is examined. The dots and error bars represent the observed neural activations of this neuron, the solid line represents the delay- and reward-dependent subjective reward value as predicted by the hyperbolic model in Equation 2, the dotted line accordingly represents the respective data as predicted by the linear model in Equation 3, and the partly dotted line represents the data as predicted by the exponential model in Equation 4. The superiority of the hyperbolic and exponential fits over the linear fit shows that Equations 2 and 4 provided a better approximation to this neuron than Equation 3 did. For the entire neural population, the hyperbolic model clearly provided the best approximation.

model was clearly superior to both the linear and the exponential model ($AIC_{hyperbolic}$ = 97.9; AIC_{linear} = 209.66; and $AIC_{exponential}$ = 270.63). This superiority suggests that subjective reward value and neural firing rate were both influenced by delay and reward amount in a hyperbolic fashion, as predicted by Equation 2 rather than by Equation 3 or 4.

Discussion

Irrational Decision-Making and Impulsivity

Decision-making involves predicting and weighing the consequences of each response alternative and then selecting one of the alternatives on the basis of the relative value of the predicted outcomes. Inconsistently with assumptions made by many economic theories, such as utility theory, humans and other animals often do not make perfectly rational decisions aimed at optimizing the expected outcomes. On the contrary, social, biological, or emotional heuristics frequently bias decisions toward apparently irrational strategies [24, 25]. Here, we show an example of a simple but prominent irrational choice behavior: impulsive decision making—that is, the preference for a small, immediate reward over a large, delayed one. In our design, the total trial duration was constant across all trials within a session and independent of the animal's choice. Choosing the large reward on all trials would have therefore maximized the total amount of reward obtained per time and session. Because the animals nevertheless switched to the short-delayed small reward at some point during the session, they displayed a suboptimal strategy that resulted from the devaluation of temporally distant rewards. A similar behavioral bias has also been demonstrated for humans [25, 26].

Single Neurons Integrate Reward Amount and Delay Information to Encode the Subjective Reward Value

In the present study, we aimed to reveal the neural representation of the subjective value of anticipated rewards as a function of temporal discount and reward amount. We recorded activity of single forebrain units in pigeons while the animals waited for the upcoming rewards. By investigating the level of sustained neuronal delay activity, we found single units that reduced their activity with increasing delay durations, as pre-

dicted by our second hypothesis. Consistent with our first hypothesis, the same units did not show any significant activity variation when reward amount and delay-to-reward were held constant. Conversely, a comparison of the activation magnitude in trials with equal delays but different reward amounts revealed that the same units were more active when the animal expected a large rather than small reward amount, as predicted by our third hypothesis. This suggests that the delay activity was modulated by a combination of anticipated delay duration and reward amount. Finally, the difference in neural activation between large, delayed and small, immediate rewards correlated with the animal's current preference bias and the time point of the preference shift: When the neural activity during anticipation of large rewards was significantly higher than or equal to the activity during anticipation of small rewards, the animals preferred the large, delayed reward. However, once the large-reward activity dropped below the small-reward activity, the PS was induced, and pigeons started preferring the small, immediate reward.

Generally, preference for either of the two choice alternatives depended on the subjective value of the anticipated reward, which was a function of expected reward amount and waiting time. Examination of single-cell activity in this study thus showed that the neural responses during reward anticipation were modulated by both delay duration and reward amount, as predicted by classical temporal-discounting behavioral laws [10–13] that link subjective reward value to preference behavior. Although the subjective value of the anticipated reward is only one variable among others determining future choice behavior, we suggest that its representation essentially contributes to evaluating response alternatives [3, 27–31].

As an alternative explanation to the subjective-reward-value hypothesis, the reported neural activation pattern might merely be the consequence of a decrease in neural excitability over the course of a session—for example, as a measure of time lapse since session onset. This explanation, however, is unlikely because the units showed the activation decrease only before, but not after, the PS. In fact, there was a significant activation increase between block –1 and 1 in the middle of a session—a phenomenon that is inconsistent with the assumption of a continuous reduction in excitability. It is likewise inconceivable that the neurons tracked the time until, and scheduled the occurrence of, a habitual PS that developed for reasons other than the successively increasing delay of the large reward. First, it is unclear why the pigeons should acquire and maintain such a habit of systematically shifting their preference because this shift has never been reinforced by our paradigm (owing to the fixed trial durations; see Figure 1). On the contrary, our design even discouraged the establishment of a habitual PS because the received reward per time, trial, and session would be maximal if no PS occurred. Hence, the only apparent reason why pigeons nevertheless show the PS is the benefit of reducing emotional distress associated with waiting for the large reward. Second, the time point of the PS varied largely within pigeons (see Table S1). If the pigeons had developed a habit of shifting preference after a particular amount of time, one would expect a more temporally stable PS time point. Taken together, we find time-based arguments theoretically unfounded and empirically inconsistent with our data. We hence conclude that our interpretation of the neural activity as a correlate of subjective reward value is the only plausible explanation for our results.

Numerous studies reported that prefrontal activation in primates [16, 17, 27, 28, 32, 33] and humans [29, 34, 35] occurred during the delay between response and reward, and therefore after the decision had been made. This is also what we found in the present study. Naturally, this activity might reflect (or at least be associated with) cortical correlates of the emotional arousal induced by reward expectation [3, 34, 36], perhaps in order to prepare the animal for goal-directed actions. A more cognitive explanation is provided by Montague and Berns [27], who extended the prediction-error signal-learning theory by Schultz, Montague, and colleagues [28, 30] and gave a computational account of why postresponse activity may be important for subsequent decision-making. They proposed that delay activity corresponded to the integration of qualitatively different entities—in the present study reward amount and time-to-reward—to represent the anticipated choice consequence in an internal currency. This enables an organism to continuously evaluate the subjective value of the upcoming reward and decide whether to stick with the current choice in the future or choose alternative response options. Whether these two accounts refer to the same or to different cortical processes remains an open question. In any event, they both suggest that the animal needs a mechanism to integrate reward amount and time-to-reward to avoid response conflict and behave adaptively.

A Neural Network for the Integration of Reward and Delay Information in Impulsive Choice Behavior

We suggest that information on time-to-reward and reward amount is transformed into a common subjective metric in order to guide future choice behavior. This integration process is likely to take place in the frontal cortex or analog avian structures, but where the to-be-integrated information comes from and how the different structures interact during decision-making need to be investigated further. Evidence suggests that non-primary thalamic neurons encode information about reward quality, reward amount, and time-to-reward [37] and feed this information into the brain's reward-circuit structures, including amygdala, ventral tegmental area (VTA), cingulate cortex, lateral intraparietal area, perirhinal cortex, or nucleus accumbens (NAc), all of which communicate at least indirectly with prefrontal areas [28, 30, 31, 38–40]. The basolateral amygdala (BLA) and the NAc in particular appear to be key structures for controlling impulsive choice behavior in both mammals [7, 9] and birds [41]. Whereas both structures are important for maintaining the incentive value of a reward across a delay [7, 9], the BLA plays a role in updating the current, but not the referential stimulus-reward value [42], and the NAc seems to encode temporal proximity to reward delivery and reward amount [43]. Our results, in conjunction with the finding that lesions in the rodent OFC decrease impulsive choice in a de-

layed-reward paradigm [9], suggest that fronto-striatal and fronto-amygdaloid loops are essential for updating the subjective reward value with increasing delays.

Conclusions

In this study, we have shown that the reward-anticipating activity of single forebrain neurons is modulated by a combination of delay duration and reward amount. We argue that this compound is a correlate of the subjective reward value and plays a role in biasing the animal's preference to either stay with the large, delayed reward or switch to the immediate, small reward. Abnormally high weighting of the time component during this integration process might lead to an accelerated temporal discount of reward value or a reduced delay tolerance, resulting in inappropriate impulsive decision-making—a short-sighted, irrational choice behavior that is typical for neuropathologies such as addictive gambling, drug abuse, attention-deficit disorder, and frontal lobe syndrome.

Experimental Procedures

Task

The task is illustrated in Figure 1 and described in greater detail in the Supplemental Data.

Electrophysiological Recordings

Eight pigeons were prepared for extracellular single-cell recordings in awake, freely moving animals with a procedure described elsewhere [44, 45]. One microdrive per animal was chronically implanted at a lateral position within the borders of NCL, as defined by Kröner and Güntürkün [20]. The tips of the electrodes were inserted to reach the following coordinates (all dorsal-ventral coordinates in relation to brain surface and according to the pigeon brain atlas by Karten and Hodos [46]): A 4.5–6.5, L 7.5, and D 1.0–3.0. Each microdrive housed eight 25 μm formvar-coated nichrome wires. In each session, only two of the eight wires were used for recording. Neural activity was measured from the difference between one of the wires carrying a neural signal versus another wire, which served as the indifferent electrode, with minimal activity. Although sometimes several wires carried neural signals within a recording session, only one of these wires was used for recording, and after the session, the microdrive was advanced to exclude double recording from the same unit. Every weekday and in each pigeon, the electrodes were advanced by 40 μm, even when no recording took place, and the animal was returned to its homecage. The next recording session started at least 14 hr after advancement of the electrode to allow the compressed brain tissue to expand and ensure stable recordings. In six of the eight pigeons, the implants were removed after 6 weeks of recording and reimplanted in the opposite hemisphere. The minimum signal-to-noise ratio was 2:1; however, 113 of 159 neurons had a signal-to-noise ratio of 3:1 or higher. The average signal-to-noise ratio was 3.67:1.

During recording, the signals were continuously monitored with an oscilloscope and a speaker. All signals were first impedance matched through a field effect transistor (FET) headstage, amplified and filtered online, and stored on computer with standard CED amplifiers, AD converters, and Spike2 software (Cambridge Electronic Design, United Kingdom, 1401 plus system).

All subjects were kept and treated according to the University of Otago Code of Ethical Conduct for the Manipulation of Animals, and the research was approved by the University of Otago Animal Ethics Committee.

Data Analysis

Task-related activity was determined with paired t tests to test for a difference between delay and baseline activity (spikes/s). For the further analysis of units with sustained delay activity, only the three sessions preceding and following the PS were examined, and only large-reward choices before and small-reward choices after the PS were selected. For each trial, the ranked percent change in delay in relation to baseline activity was computed as the measure of activation magnitude, and the response pattern was evaluated with a Kolmogorov-Smirnov test. For details on the data analysis, refer to the Supplemental Data.

Supplemental Data

Supplemental Results and Discussion, detailed Experimental Procedures, two supplemental figures, and two supplemental tables are available at http://www.current-biology.com/cgi/content/full/15/7/594/DC1/.

Acknowledgments

This study was supported by a Royal Society of New Zealand Marsden grant (UOO012; M.C.) and grants from the Deutsche Forschungsgemeinschaft through the priority program Executive Functions (SPP 1107, Di 459/4-1, and Di 459/4-2; B.D., O.G., and T.K.) and the G.-A. Lienert Stiftung (T.K.). We would like to thank Martin Spiess for his great assistance with the model fits, Daniel Durstewitz for helpful discussions and comments, and Nicole Frost for great support and the initial training of the animals.

Received: September 23, 2004
Revised: February 10, 2005
Accepted: February 11, 2005
Published: April 12, 2005

References

1. Bizot, J., Le Bihan, C., Puech, A.J., Hamon, M., and Thiébot, M. (1999). Serotonin and tolerance to delay of reward in rats. Psychopharmacology (Berl.) *146*, 400–412.
2. Evenden, J.L. (1999). Varieties of impulsivity. Psychopharmacology (Berl.) *146*, 348–361.
3. Bechara, A., Tranel, D., and Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. Brain *123*, 2189–2202.
4. Evenden, J.L., and Ryan, C.N. (1996). The pharmacology of impulsive behaviour in rats: The effects of drugs on response choice with varying delays of reinforcement. Psychopharmacology (Berl.) *128*, 161–170.
5. Wade, T.R., de Wit, H., and Richards, J.B. (2000). Effects of dopaminergic drugs on delayed reward as a measure of impulsive behavior in rats. Psychopharmacology (Berl.) *150*, 90–101.
6. Cardinal, R.N., Robbins, T.W., and Everitt, B.J. (2000). The effects of d-amphetamine, chlordiazepoxide, alpha-flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. Psychopharmacology (Berl.) *152*, 362–375.
7. Cardinal, R.N., Pennicott, D.R., Sugathapala, C.L., Robbins, T.W., and Everitt, B.J. (2001). Impulsive choice induced in rats by lesions of the nucleus accumbens core. Science *292*, 2499–2501.
8. Kheramin, S., Body, S., Ho, M., Velazquez-Martinez, D.N., Bradshaw, C.M., Szabadi, E., Deakin, J.F., and Anderson, I.M. (2003). Role of the orbital prefrontal cortex in choice between delayed and uncertain reinforcers: A quantitative analysis. Behav. Processes *64*, 239–250.
9. Winstanley, C.A., Theobald, D.E.H., Cardinal, R.N., and Robbins, T.W. (2004). Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. J. Neurosci. *24*, 4718–4722.
10. Rachlin, H., and Green, L. (1972). Commitment, choice and self control. J. Exp. Anal. Behav. *17*, 15–22.
11. Mazur, J.E. (1984). Tests of an equivalence rule for fixed and variable reinforcer delays. J. Exp. Psychol. Anim. Behav. Process. *10*, 426–436.
12. Mazur, J.E. (1988). Estimation of indifference points with an adjusting-delay procedure. J. Exp. Anal. Behav. *49*, 37–47.

13. Grossbard, C.L., and Mazur, J.E. (1986). A comparison of delays and ratio requirements in self-control choice. J. Exp. Anal. Behav. *45*, 305–315.
14. Brody, C.D., Hernandez, A., Zainos, A., and Romo, R. (2003). Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. Cereb. Cortex *13*, 1196–1207.
15. Leon, M.I., and Shadlen, M.N. (1999). Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. Neuron *24*, 415–425.
16. Wallis, J.D., and Miller, E.K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. Eur. J. Neurosci. *18*, 2069–2081.
17. Tremblay, L., and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. Nature *398*, 704–708.
18. Mogensen, J., and Divac, I. (1982). The prefrontal "cortex" in the pigeon: Behavioral evidence. Brain Behav. Evol. *21*, 60–66.
19. Mogensen, J., and Divac, I. (1993). Behavioural effects of ablation of the pigeon-equivalent of the mammalian prefrontal cortex. Behav. Brain Res. *55*, 101–107.
20. Kröner, S., and Güntürkün, O. (1999). Afferent and efferent connections of the caudolateral neostriatum in the pigeon (Columba livia). J. Comp. Neurol. *407*, 228–260.
21. Reiner, A., Perkel, D.J., Bruce, L.L., Butler, A.B., Csillag, A., Kuenzel, W., Medina, L., Paxinos, G., Shimizu, T., Striedter, G., et al. (2004). Revised nomenclature for avian telencephalon and some related brainstem nuclei. J. Comp. Neurol. *473*, 377–414.
22. Kalenscher, T., Diekamp, B., and Güntürkün, O. (2003). Neural architecture of choice behaviour in a concurrent interval schedule. Eur. J. Neurosci. *18*, 2627–2637.
23. Manly, B.F.J. (1997). Randomization, Bootstrap and Monte Carlo methods in biology, Second Edition (Boca Raton: Chapman & Hall/CRC).
24. Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. J. Econ. Behav. Organ. *3*, 367–388.
25. Kahneman, D., and Tversky, A. (1984). Choices, values, and frames. Am. Psychol. *39*, 341–350.
26. Green, L., Myerson, J., and McFadden, E. (1997). Rate of temporal discounting decreases with amount of reward. Mem. Cognit. *25*, 715–723.
27. Montague, P.R., and Berns, G.S. (2002). Neural economics and the biological substrates of valuation. Neuron *36*, 265–284.
28. Schultz, W., Tremblay, L., and Hollerman, J.R. (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. Cereb. Cortex *10*, 272–284.
29. Breiter, H.C., Aharon, I., Kahneman, D., Dale, A., and Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. Neuron *30*, 619–639.
30. Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599.
31. Platt, M.L., and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. Nature *400*, 233–238.
32. Hikosaka, K., and Watanabe, M. (2000). Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. Cereb. Cortex *10*, 263–271.
33. Watanabe, M., Hikosaka, K., Sakagami, M., and Shirakawa, S. (2002). Coding and monitoring of motivational context in the primate prefrontal cortex. J. Neurosci. *22*, 2391–2400.
34. Critchley, H.D., Mathias, C.J., and Dolan, R.J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. Neuron *29*, 537–545.
35. Knutson, B., Fong, G.W., Adams, C.M., Varner, J.L., and Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. Neuroreport *12*, 3683–3687.
36. Patterson, J.C., Ungerleider, L.G., and Bandettini, P.A. (2002). Task-independent functional brain activity correlation with skin conductance changes: An fMRI study. Neuroimage *17*, 1797–1806.
37. Komura, Y., Tamura, R., Uwano, T., Nishijo, H., Kaga, K., and Ono, T. (2001). Retrospective and prospective coding for predicted reward in the sensory thalamus. Nature *412*, 546–549.
38. Shima, K., and Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. Science *282*, 1335–1338.
39. Baxter, M.G., Parker, A., Lindner, C.C., Izquierdo, A.D., and Murray, E.A. (2000). Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. J. Neurosci. *20*, 4311–4319.
40. Shidara, M., and Richmond, B. (2002). Anterior cingulate: Single neuronal signals related to degree of reward expectancy. Science *296*, 1709–1711.
41. Izawa, E., Zachar, G., Yanagihara, S., and Matsushima, T. (2003). Localized lesion of caudal part of lobus parolfactorius caused impulsive choice in the domestic chick: Evolutionarily conserved function of ventral striatum. J. Neurosci. *23*, 1894–1902.
42. Baxter, M.G., and Murray, E.A. (2002). The amygdala and reward. Nat. Rev. Neurosci. *3*, 563–573.
43. Izawa, E., and Matsushima, T. (2004). Neuronal representation of temporal proximity of anticipated reward in the avian striatum. Program No. 89.12 2004 Abstract Viewer/Itinerary Planner (Washington, DC: Society for Neuroscience).).
44. Colombo, M., Frost, N., and Steedman, W. (2001). Responses of ectostriatal neurons during delayed matching-to-sample behavior in pigeons (Columba livia). Brain Res. *917*, 55–66.
45. Bilkey, D.K., Russell, N., and Colombo, M. (2003). A lightweight microdrive for single-unit recording in freely moving rats and pigeons. Methods *30*, 152–158.
46. Karten, H.J., and Hodos, W. (1967). Stereotaxic Atlas of the Brain of the Pigeon (Columba livia) (Baltimore: Johns Hopkins University Press).