Planning activity for internally generated reward goals in monkey amygdala neurons

István Hernádi[1,2]†, Fabian Grabenhorst[1]*† and Wolfram Schultz[1]

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

[2] Present address: Department of Experimental Zoology and Neurobiology, Institute of Biology, University of Pécs, H-7624, Pécs, Hungary.

† These authors contributed equally to this work.

*Correspondence to: fg292@cam.ac.uk

The best rewards are often distant and can only be achieved by planning and decision-making over several steps. We designed a multi-step choice task in which monkeys followed internal plans to save rewards towards self-defined goals. During this self-controlled behavior, amygdala neurons showed future-oriented activity that reflected the animal's plan to obtain specific rewards several trials ahead. This prospective activity encoded crucial components of the animal's plan, including value and length of the planned choice sequence. It began on initial trials when a plan would be formed, reappeared step-by-step until reward receipt, and readily updated with a new sequence. It predicted performance, including errors, and typically disappeared during instructed behavior. Such prospective activity could underlie the formation and pursuit of internal plans characteristic for goal-directed behavior. The existence of neuronal planning activity in the amygdala suggests an important role for this structure in guiding behavior towards internally generated, distant goals.

The pursuit of distant rewards through planned behavior is a key function of the primate brain. As monkeys search their large habitats for the best foods, humans plan their careers towards the benefits of future rewards. Although planned, goal-directed behaviors can differ in timescale—from foraging across food patches to human economic saving—psychological and economic theories suggest a common principle[1-3]: the formation of an internal plan to obtain a distant goal, and its subsequent pursuit over several steps. In the present study, we investigated the neuronal mechanisms for reward-based planning by combining the advanced behavioral capacities of non-human primates with single-neuron recordings.

Planning functions have traditionally been ascribed to the frontal lobe[4-6]. Indeed, neurophysiological experiments have provided detailed accounts of frontal lobe activity during generation, execution and updating of movement plans[7-10]. Neuronal activity in frontal lobe and connected basal ganglia also precedes self-initiated movements[11-14], which constitute the effective means for carrying out a plan. Despite these advances, a fundamental question has remained unanswered: as planned behavior is typically motivated by the prospect of reward, what are the neural processes for directing action plans towards internally defined, distant reward goals?

We addressed this question by recording the activity of single neurons in the amygdala, a nuclear structure in the medial temporal lobe implicated in reward and emotion[15-22] with inputs to frontal lobe-basal ganglia systems involved in action planning[23]. We hypothesized that amygdala neurons might show planning activity related to internally generated reward goals and their value.

In addition to its well-known roles in emotion, the amygdala is an important component of the reward system[15,16,20]. In animals, amygdala neurons encode the value of sensory stimuli[15,21,22,24,25] and amygdala lesions impair reward-guided behaviors[16,17,26,27]. The human amygdala also processes rewards[18,28] and reward-based decisions[18,29,30], and amygdala damage is associated with decision impairments[31,32]. Accordingly, current theories view the amygdala as an associative learning and valuation system that regulates affective, cognitive and autonomic processes as well as decisions and behavior[15-19,21]. However, the amygdala's role in the pursuit of internally defined, distant rewards through planned behavior is still unexplored.

Here we show that during planned behavior, the primate brain generates future-oriented activity related to self-defined goals, which persists until a distant reward is received. We

recorded the activity of amygdala neurons while monkeys produced choice sequences to save rewards over several steps towards internal goals. We found amygdala neurons with prospective activity that reflected the animal's plan to obtain specific rewards by saving for a given number of steps. In different neurons, the activity reflected crucial components of the animal's plan, including the subjective value of the current plan ('sequence value') and the planned number of saving steps ('sequence length'). This planning activity began before the animal initiated a saving sequence and reoccurred with each step during pursuit of the plan, as the animal progressed towards reward. Such prospective neuronal activity seems suited to guide planned behavior over multiple steps towards distant reward goals.

RESULTS

Sequential reward-saving task

Two monkeys performed in a sequential reward-saving task in which they could follow internal plans towards obtaining reward at the end of a sequence of trials. On each trial (i.e. each step within a sequence) the animals freely chose to save juice reward for future consumption or spend the already saved amount (**Fig. 1a**). Consecutive save choices increased the available juice amount as determined by a given 'interest rate' (**Fig. 1b**, green, Eq. 1). Choices were made by a saccade towards the save or spend cue; pre-trained save cues indicated current interest rate. The animals freely determined the length of each saving sequence. This self-controlled and sequential task design allowed the animals to plan their behavior over multiple trials and anticipate final rewards more than 100 s in advance (up to 9 consecutive trials with ~12 s cycle time). Randomized cue positions precluded planning of left-right action sequences. To confirm the internal nature of planning, we also tested externally instructed 'imperative' save-spend sequences with comparable lengths.

Our task incorporated two key aspects of economic saving[1]: the internal formation of a plan to obtain a distant reward goal, and its pursuit over sequential choices. The final reward in a saving sequence corresponded to the animal's goal, and the sequence length corresponded to the means by which to achieve the goal. These features made the saving task a suitable model to investigate reward-based planning and goal pursuit[1-3].

As economic choices critically depend on value, testing the hypothesis of planning activity in reward neurons required us to determine the subjective values that the animals

associated with specific saving plans. These values depended not only on final reward amounts but also on expenditure related to sequence length: because higher rewards typically required longer sequences (determined by the current interest rate, **Fig. 1b** green), their value was compromised by temporal delay and physical effort. To capture these factors in a direct manner, we followed the general notion of standard economic choice theory that estimates subjective values from behavioral choices. We derived the value of different saving sequences by calculating the relative frequency with which the animals produced each sequence length within a given interest rate (**Fig. 1b**, black bars). Accordingly, for a given interest rate, a sequence had a higher subjective value if the animal chose it more frequently than other sequences. To account for reward magnitude differences between interest rates, choice frequencies for different sequences were weighted by associated reward magnitudes (**Fig. 1b**, green curves). Subjective values determined in this manner constituted a decision variable for the animals which we call 'sequence value' (i.e. the subjective value associated with a given saving sequence, **Fig. 1b**, magenta curves).

Sequence value differed from final reward magnitude as it was a non-monotonic function of sequence length; the shape of the value function depended on the relative frequency with which a sequence was chosen (**Fig. 1b**, magenta curves). By contrast, final reward magnitude increased monotonically with sequence length (**Fig. 1b**, green curves). Because sequence values were derived from the animals' relative choice frequencies, they effectively incorporated benefits related to reward amounts as well as expenditure related to waiting times and physical effort. Typically, sequence value functions increased with sequence length up to a peak, and then decreased with longer sequences that the animal chose less frequently, likely due to temporal discounting and physical effort cost. This non-linearity in value functions made it possible to distinguish neuronal coding of subjective sequence value from objective sequence length and reward magnitude. To control for valuations of save and spend choice options on single trials we also defined trial-by-trial subjective values ('spend value' and 'save value', Eq. 2,3, see Methods).

Behavioral data

Figure 1b shows the animals' relative choice frequencies for different saving sequences, calculated for different interest rate conditions. The animals saved more when interest was high:

mean saving lengths increased with higher interest rates (P = 0.003, linear regression, **Fig. 1b,c**, black, **Supplementary Fig. 1a**). Typically, at the beginning of a testing session, the animals adapted to the current interest rate within a few trials (**Supplementary Fig. 1b**). For interest rates with different rates of reward return, the animals' behavior approximated optimality by maximizing reward rate (**Supplementary Fig. 1c,d**). Control experiments confirmed that the animals adjusted their behavior even when interest changed without notification and tracked accumulated reward over consecutive trials (**Fig. 1c**, magenta; **Supplementary Fig. 1e,f**). Thus, saving was adaptive and internally controlled. It did not simply reflect conditioned or automated behavior.

To confirm that the animals planned saving sequences in advance, we examined trial-by-trial reaction times. Mean reaction times were shorter on spend trials compared to save trials (**Fig. 1d**, black vs. magenta data, z = –48.57, P = 1.0×10^{-10}, Wilcoxon test), suggesting higher motivation for immediately upcoming rewards. Reactions on spend trials were also faster after longer saving sequences, i.e. when the animals would obtain higher rewards (**Fig. 1d** black, $r = -$ 0.85, P = 0.007, linear regression), which demonstrated that the animals tracked internally the accumulated reward and were more motivated for higher amounts. Critically, reactions across consecutive save trials within a sequence, while the animals progressed towards their current goal, also depended on final sequence length, with faster reactions during longer sequences (r = –0.81, P = 0.009, **Fig. 1d** magenta). This suggested that the animals anticipated final reward outcomes several trials in advance, consistent with internally planned, goal-directed saving.

To confirm the behavioral importance of subjective values, we regressed trial-by-trial save-spend choices on subjective values using logistic regression (Eq. 4). 'Spend value' reflected the subjective value expected from spending on the current trial , whereas 'save value' reflected the average value expected in all future trials of the current saving sequence. Our main planning variable 'sequence value' corresponded to the spend value on the final trial of a sequence (i.e. the spend value actually chosen); accordingly, its influence on choice was captured by the spend value regressor. We used independent behavioral data for deriving subjective values (n = 5,600 trials) and for estimating logistic regression coefficients (n = 5,933 trials).

Logistic regression identified subjective values as main explanatory variables for saving behavior (**Fig. 1e**): higher spend values decreased the likelihood of saving on the current trial (negative beta in **Fig. 1e**), whereas higher save values increased saving (positive beta in **Fig. 1e**).

A stepwise logistic regression confirmed these results by selecting the key variables spend value and save value (both $P < 10^{-16}$). Regressing trial-by-trial reaction times on subjective values confirmed and extended these results: reaction times reflected subjective values with faster responses for higher sequence values ($P = 0.001$, multiple linear regression, Eq.5, **Supplementary Fig. 2a,b**), even as early as the initial saving trial ($P = 0.004$). Similar results were obtained from analysis of licking durations in animal A (**Supplementary Fig. 2c**).

Taken together, behavioral data confirmed subjective valuation of trial-by-trial choices, saving sequences, and final reward goals, consistent with internally planned saving.

Planning activity in amygdala neurons: single neuron data

While the animals saved rewards step-by-step towards self-defined goals, a striking group of amygdala neurons signaled the animals' internal saving plans multiple trials in advance. We refer to such prospective activity as 'planning activity' because its occurrence preceded the end of a saving sequence by several steps, and because it referred to a future event that was self-determined by the animal and existed only internally at the time of saving. Planning activity in different neurons reflected different components of the animal's plan: the subjective value of the planned sequence ('sequence value') or the planned number of saving steps ('sequence length'). Selection criteria for neurons with planning activity were task-related activity ($P < 0.01$, Wilcoxon test) and a significant regression coefficient for sequence value or sequence length ($P < 0.05$, multiple regression analysis, Eq. 6–9).

The neuron in **Fig. 2** had phasic trial-by-trial activity during the fixation period that was highest during sequences in which the animal would eventually spend on the fifth trial, and lower for shorter or longer sequences (**Fig. 2a**). This activity profile resembled closely the distribution of sequence values derived from the animal's choice preferences (**Fig. 2a** magenta curve): for this interest rate, five-trial sequences had the highest value as the animal chose them most frequently. Within trials, the prospective activity appeared during ocular fixation and continued beyond the cue period when a save-spend choice was made (**Fig. 2b**). Linear regression indicated a better relationship to sequence value ($r^2 = 0.54$, $P = 1.4 \times 10^{-8}$, $n = 40$, **Fig. 2c**) than to sequence length ($r^2 = 0.21$, $P = 0.003$) or final juice amount ($r^2 = 0.07$, $P = 0.09$). Multiple regression confirmed a relationship between neuronal activity and sequence value ($P = 3.8 \times 10^{-6}$, Eq. 6) and factored out other variables, including subjective values related to single trial

choices ($P > 0.05$, **Fig. 2d**, **Supplementary Fig. 3**). The relationship between activity and sequence value disappeared in externally cued trials when saving was instructed (**Fig. 2e**, $P > 0.05$, multiple regression), despite comparable behavioral outcome anticipation (regression of sequence length on reaction times: $P < 0.05$; **Supplementary Fig. 2d**). Thus, during internally controlled step-by-step saving, the neuron showed prospective activity related to the subjective value of the animal's saving plan.

We found different forms of planning activity as illustrated in **Fig. 3** for four single amygdala neurons. The neuron in **Fig. 3a** resembled the one in **Fig. 2** as it signaled sequence value across all trials ($P = 0.04$, $n = 53$, multiple regression). In addition, it encoded spend values on single trials (**Fig. 3a right**). To be engaged in planned saving, amygdala neurons should also encode the initial setting of a plan, which may occur as early as the first trial of a sequence. This is exactly what we observed for the neuron in **Fig. 3b**. This neuron encoded sequence value specifically on the first trial of each sequence (**Fig. 3a,b**, bold colors) but not on subsequent trials (light colors). In this neuron, planning activity occurred early on at trial start before fixation. Multiple regression confirmed a parametric value signal ($P = 5.9 \times 10^{-4}$, $n = 42$, **Fig. 3b** right), which differed distinctly from categorical coding of sequence onset previously found in amygdala neurons during an instructed task[33]. Accordingly, this neuron encoded the prospective valuation of an internal saving plan, well before the animal implemented the plan.

Sequence value neurons signaled the value of the animal's plan but not the required steps for its implementation. By contrast, the neurons in **Fig. 3c,d** encoded the planned number of choice steps for a given sequence, i.e. the planned sequence length. The neuron in **Fig. 3c** showed planning activity that predicted the sequence length the animal would eventually produce. It encoded sequence length throughout all trials in a sequence ($P = 0.0032$, multiple regression, $n = 41$) with higher activity for shorter sequences. The neuron in **Fig. 3d** also encoded sequence length but it did so specifically on initial trials ($P = 0.0049$, multiple regression, $n = 40$), with higher activity predicting longer sequences. Thus, sequence length neurons encoded the animal's internal plan in terms of the required number of saving steps.

Taken together, prospective activity in amygdala neurons encoded crucial components of the animal's saving plan, including subjective value and objective length of the planned sequence. Such planning activities occurred either on initial trials or throughout whole saving sequences.

Planning activity in amygdala neurons: population data

Among 329 task-related neurons, 123 (37%, 66/57 from animal A/B) showed planning activity related to sequence value or sequence length, either throughout saving sequences or specifically on initial trials (**Fig. 4a–d**, **Table 1**, **Supplementary Tables 1–3, Supplementary Fig. 4**). The average activity of sequence value neurons followed closely the average subjective value profile, which was a non-monotonic function of sequence length ($r^2 = 0.91$, P<0.0001, linear regression; compare magenta curve and black bars in **Fig. 4b**). Analysis of trial-by-trial activity in these neurons confirmed this effect ($P = 2.2 \times 10^{-15}$, partial correlation factored out sequence length). By contrast, activity of sequence length neurons increased linearly with sequence length ($r^2 = 0.85$, $P = 0.0035$, linear regression, **Fig. 4d**). Analysis of trial-by-trial activity in these neurons confirmed this effect ($P = 7.4 \times 10^{-5}$, partial correlation factored out sequence value). Supplementary analysis confirmed a graded, parametric representation of sequence value or sequence length, rather than sharp tuning to specific sequences (**Supplementary Fig. 5**). A subset of neurons with planning activity was tested in the imperative task. In most of them (53/57, 93%), planning activity was not found when saving was externally instructed (**Fig. 4e,f**). Thus, planning activity appeared to be largely specific for internally controlled saving behavior.

Although planning activity often occurred without coding of other variables, some planning activities reflected additional task-related variables for guiding behavior on single trials (**Supplementary Table 2**), including previously reported trial-by-trial save-spend choices[34].

Additional tests confirmed the statistical significance of planning activity. Compared to randomly shuffled data, the distribution of regression coefficients for planning activity was shifted towards higher positive and negative values (**Fig. 4g**, $P = 1.8 \times 10^{-27}$, Kolmogorov-Smirnov test). The observed proportion of planning activities exceeded that expected by chance ($P < 10^{-62}$, binomial probability test); less than five percent of coefficients from shuffled data were significant. Using alternative regression models (see Methods, **Supplementary Table 1**), we found that the number of identified neurons with planning activity depended little on the specific model used and on the inclusion of different control covariates: the number of neurons with planning activity over several alternative models varied by less than 5% compared to our main models, with percentages ranging from 35% to 41% (our original analysis resulted in 37%).

Thus, planning variables explained unique variance components in neuronal activity relative to other variables.

Histological reconstructions verified that the recording sites were restricted to the amygdala and covered basolateral and centromedial regions (**Fig. 4h**, **Supplementary Fig. 6**). Although neurons with planning activity were found in both basolateral and centromedial amygdala, they occurred significantly more frequently in the basolateral complex ($\chi^2_1 = 7.86$, $P = 0.005$, χ^2-test **Fig. 4i**). A similar clustering was not found for other types of activities, including those reflecting current-trial save-spend choices ($P > 0.05$, χ^2-test). This anatomical trend could indicate relatively greater importance of the basolateral amygdala for planned reward-saving.

Adaptation dynamics of planning activity

If a neuron encoded components of the animals' saving plan, its activity should update once a sequence is completed and begin to reflect properties of the subsequent sequence. Accordingly, we examined sequence transitions by comparing activity on spend trials and subsequent save trials (the last and first trials of two successive sequences). **Figure 5a** illustrates such transitions in a single neuron with planning activity related to sequence length. Transitions (dashed vertical lines) were marked by activity changes that scaled with changes in planned sequence length (compare thick gray and green lines). The neuron's activity reflected planned sequence length within sequences (**Fig. 5b** left) and changes in planned sequence lengths at transitions (**Fig. 5b** middle). Activity was unrelated to within-sequence reward proximity (trials until reward, **Fig. 5b** right). Sequence-by-sequence adaptation was also evident in population activity (**Fig. 5c**, left and middle). Thus, planning activity adapted sequence-by-sequence to reflect changes in the animals' internal plan.

The observed sequence-by-sequence updating differs substantially from sustained activity increases typically associated with reward expectation[20]. In control analyses, population activity was unrelated to within-sequence reward proximity (**Fig. 5c** right, $r = 0.06$, $P = 0.1$, linear regression) and few individual planning activities reflected reward proximity (12/123, 10%; supplementary regression with reward proximity covariate, **Fig. 5d**) or reward expectation indexed by reaction times (11/123 responses, 9%; supplementary regression with reaction times covariate). Thus, most planning activities were insensitive to trial-by-trial reward proximity and reward expectation.

Sequence value neurons were of particular interest, for they allowed us to test whether planning activity at the start of a testing session was in-step with behavior as the animal adapted to current interest rate. We defined criteria for behavioral and neuronal adaptation, and plotted the number of steps to criterion (**Fig. 5e**, upper panel). On average, as soon as the animal chose its preferred sequence, activity began to accurately reflect the sequence's current value (**Fig. 5e**, compare black and magenta curves). In most cases, planning activity reflected sequence value accurately on the first time the preferred sequence was chosen (**Fig. 5e**, black). Thus, amygdala sequence value neurons adapted in-step with the animals' behavior.

Planning activity predicts performance, including errors

If planning activity in the amygdala participated in guiding the animals' behavior, it should fluctuate with behavioral performance. We tested this hypothesis by regressing a measure of the animals' reward-saving efficiency on the standardized neuronal regression coefficients for sequence value and sequence length. We measured reward-saving efficiency as the accumulated sequence value per unit time, which indicated the extent to which the animals maximized subjective value. Across neuronal responses, stronger planning activity in a given testing session predicted more efficient reward-saving ($r = 0.39$, P $= 2.9 \times 10^{-5}$, linear regression, **Fig. 6a**). This relationship remained highly significant after factoring out alternative variables, including interest rate, juice amount, error rate, number of trials, and reward range ($P < 0.001$, partial correlation). Thus, the strength of planning activity in amygdala neurons explained variation in the animals' saving efficiency.

We also tested whether planning activity tracked fluctuations in behavior as indexed by errors in trial-by-trial performance. In a population analysis, we regressed neuronal activity on sequence value and sequence length separately for trials on which the animals committed errors—which implied failure to progress towards rewards—and for the immediately preceding and following non-error trials. Just before error trials, population activity exhibited a significant relationship to planning variables (**Fig. 6b**, 'Pre-error). However, this relationship declined when the animals committed an error ('Error'), and subsequently reappeared when they resumed saving towards their current goal ('Post-error'). Thus, planning activity transiently declined on error trials, thereby reflecting performance fluctuations within a testing session.

DISCUSSION

We found prospective activity in amygdala neurons that reflected the animal's plan to save rewards towards specific goals several trials ahead. This activity predicted behavior not for individual trials but for whole choice sequences. In different neurons, it coded the subjective value of the planned choice sequence (sequence value) or the objective number of planned saving steps (sequence length). Crucially, saving plans were not signaled by the environment but were self-defined and existed only internally. Accordingly, such activities constitute the neuronal building blocks of an internal behavioral plan. The occurrence of planning activity on initial trials and throughout saving sequences matches the timing of key cognitive processes thought to underlie goal-directed behavior[1-3]: the formation of a plan and its subsequent pursuit. In many neurons, the disappearance of prospective activity during instructed trials, activity updating in-step with the animals' behavior, absence of reward proximity coding, and relationship to performance provided further evidence for the encoding of an internal plan. By encoding the central components of a plan to obtain a future reward, prospective amygdala neurons may participate in guiding self-controlled behavior over several steps towards distant goals.

Functional significance of planning activity

Neurons in different brain structures encode reward values based on external cues and reinforcement history[20,21,35-42]. Although important components of value coding, such activities by themselves could not guide sequential behavior towards internal, distant goals. By contrast, the presently observed sequence value signals seem ideally suited for this purpose—they reflected the value of the animal's current plan, appeared several trials before a reward goal was obtained, reoccurred at each choice step until reward receipt, and fluctuated with performance. Such value-related planning activity in the amygdala could serve in the guidance of behavior towards an internal goal, and in the ongoing regulation of affective and cognitive processes during goal pursuit. The separate coding of sequence value specifically on first saving trials could reflect the initial formation of a plan, or a decision process that selects among alternative plans.

Prospective activity in a different category of neurons coded the planned sequence length, thereby reflecting the means by which a distant reward would be obtained. These neurons did not

specify a movement plan—which was precluded by the experimental design using randomized cue positions—but an abstract, movement-independent plan based on the number of choice steps. Encoding behavioral plans in such abstract form seems advantageous for goal-directed behavior, as specific movement requirements are often not known in advance. These amygdala sequence length signals observed during economic, free choices may complement frontal lobe signals related to final target positions[8] and categories of action sequences[9] found in instructed tasks. The observed encoding of sequence length specifically on first trials is consistent with the updating of an internal behavioral plan, analogous to updating of externally cued motor plans seen in frontal cortex[10].

Planning activity and relation to other brain systems

Consistent with classical concepts[43], we suggest that amygdala planning activity provides directive inputs to frontal lobe-basal ganglia structures involved in sequential, self-initiated behavior[7-14]. Amygdala sequence value neurons could send a value or goal signal to striatal and frontal areas to influence the initial selection of a plan and guide ongoing behavior towards an internal, distant reward goal. Via the same routes, amygdala sequence length neurons may participate in transforming abstract, value-based plans into concrete action. Amygdala reward-planning activity may also influence multi-step learning processes involving frontal-striatal[44] and parietal areas[45], and may complement prospective activity observed in rodent hippocampus during spatial navigation[46,47].

Existing evidence supports our interpretation that amygdala planning activity informs frontal-striatal systems during goal-directed behavior[16,19,26,27]. For example, in a recent study, value coding in primate orbitofrontal cortex during reward-based choice was diminished following amygdala lesions[48]. Notably, our data cannot determine whether planning activity originated locally within amygdala circuits or elsewhere; resolving this important issue will require further experimentation.

Planning activity and amygdala function in affective state

Some amygdala neurons combined planning activity with additional task-relevant variables, including trial-specific values and reward expectation (**Supplementary Table 3**, **Supplementary Fig. 7**). Via known amygdala outputs to basal forebrain, hypothalamus and

brain stem[23], such hybrid neurons could be involved in regulating motivation, attention and affective state[21,33,40,49], based on the animal's current plan. By contrast, the 'pure' planning activity is unlikely to reflect these processes: First, planning activity often disappeared during instructed behavior, despite comparable reward timing and anticipation. Second, most planning activities were unrelated to reward proximity and expectation, which seems incompatible with general functions in motivation or attention. Finally, the functionally different profiles of planning activity—coding sequence value or sequence length, either throughout sequences or on initial trials—seem inconsistent with simple roles in reward expectation or arousal.

Thus, most planning activities failed to show standard measures of reward expectation and related state value, attention and arousal[20,21,40], and therefore appear to reflect the animal's internal saving plan. Although pure planning activity seems unrelated to attention, its combined coding with single-trial values and reward expectation in hybrid neurons could serve to focus processing onto current plans, which may be important in reward-saving behavior as suggested by psychological and economic theories[1].

Amygdala planning activity at sequence start

In a previous study, amygdala neurons signaled the start of behavioral sequences during forced, multi-step reward schedules[33]. Although we also found sequence onset responses in some neurons (17%), our typical planning activity failed to occur in forced, imperative trials and was largely restricted to free choices. Crucially, planning activity was based entirely on internally generated goals and associated saving plans, and the animals made own choices rather than follow cue instructions. Planning activity also reflected parametrically the key variables sequence value or sequence length; a generalized, non-parametric response to sequence onset was factored out by multiple regression. Further, in supplementary analyses, only few neurons (< 5%) showed systematic trial-order dependent activity beyond first trials.

Thus, although visually similar to activity in forced, multi-step reward schedules, our planning activity critically reflected the internal nature of the task.

Planning activity and theories of amygdala function

Current theories emphasize the amygdala's capacity as a valuation structure to signal behavioral goals based on external cues and past experience[15-21]. Our data significantly extend

these accounts by demonstrating that amygdala goal representations can reflect internally generated goals for distant, future rewards. This finding is difficult to reconcile with the influential view of the amygdala as an impulsive, stimulus-bound system that signals immediate rewards in response to external cues[50]—a view often used to interpret amygdala dysfunction in addiction and other disorders. By contrast, the presently described amygdala neurons signaled the prospect of internally generated, future rewards that became available only after multi-step planning. Such prospective activity typically disappeared during externally cued behavior, and its timecourse did not resemble simple reward expectation.

We propose an updated view of the amygdala that incorporates a planning function for internally generated, distant reward goals. This conceptual advance may open up new avenues for understanding amygdala function in health and disease, including in addiction and other states with dysfunctional reward pursuit.

Conclusion

A basic principle underlying goal-directed behavior is the formation of an internal plan and its pursuit over successive steps. Our findings, experimentally focused on shorter timescales, demonstrate neuronal building blocks for these fundamental processes in the amygdala, although additional mechanisms are likely required for planned behavior over longer periods.

As a valuation system, the amygdala seems predisposed to provide the goals for internally planned behavior. However, we do not believe the amygdala is unique in encoding reward-based plans. Our experimental approach—combining neurophysiology with an internally controlled, sequential reward-planning task—may help uncover reward-based planning activity in other brain structures and perhaps other species.
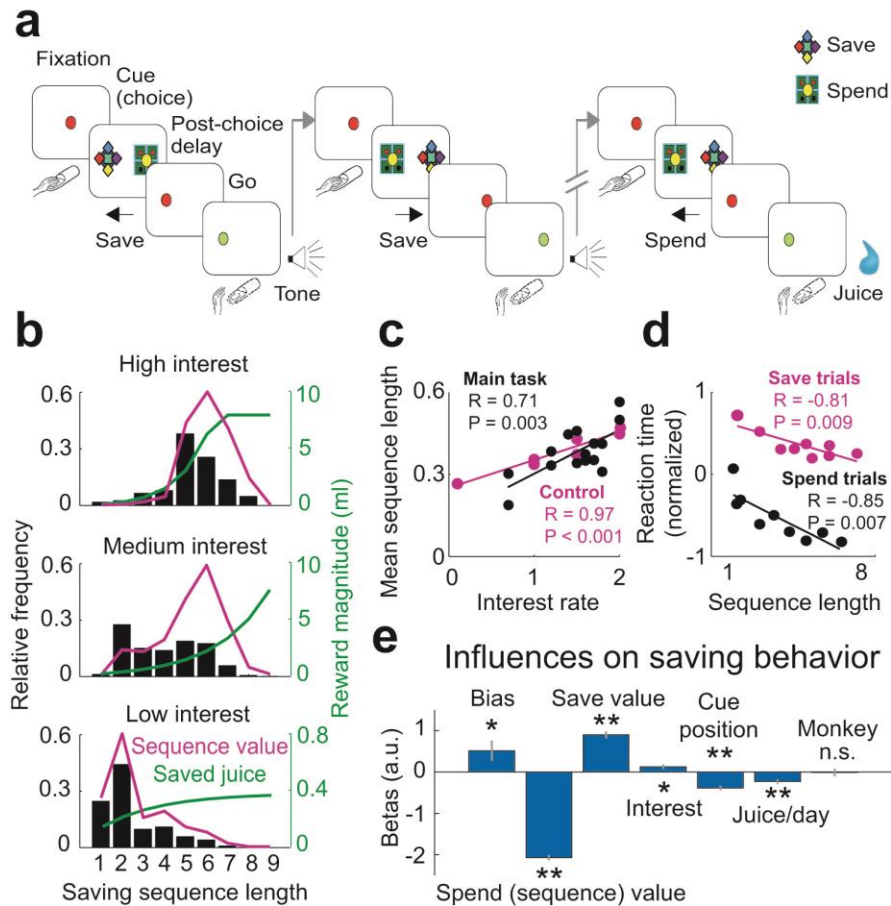
FIGURE LEGENDS

Figure 1 Reward-saving behavior in monkeys. (**a**) Sequential saving task. Animals chose freely to save or spend reward and determined internally the length of each saving sequence. Consecutive save choices increased reward amounts (determined by interest rate); spend choice resulted in reward delivery. Sequences lasted up to 9 consecutive trials (~12 s cycle time/trial). (**b**) Saving behavior, reward increases, and subjective value functions for different interest rates. Bars: relative frequencies with which animals produced different sequences, combined across animals. Green curves: reward amounts for different sequences. Magenta: subjective values (normalized), combining choice frequencies with reward magnitudes. With highest interest rate, reward stagnated after seven trials; most neuronal recordings involved intermediate interest rates. (**c**) Monkeys adapted their saving behavior to interest rate. Linear regression of weighted mean sequence length on interest for main task (black, *n* = 17) and control test with uncued changes in interest (magenta, *n* = 9). Data combined across animals. (**d**) Linear regression of reaction time on final sequence length. Reaction times (equally populated bins pooled over animals and

interest rates, z-normalized within sessions) on spend trials (black, averaged over $n = 3,033$ trials) and save trials (magenta, averaged over $n = 8,500$ trials) were shorter for longer sequences (i.e. higher rewards). **(e)** Logistic regression of trial-by-trial choices. Spend/save value: subjective value associated with spending/saving on current trial; sequence value: subjective sequence value (spend value on final trial). Bias: constant; Cue position: left/right save cue position; Juice/day: consumed juice; Monkey: animal identity. $**P < 0.005$, $*P < 0.05$; n.s. not significant. Error bars: s.e.m.
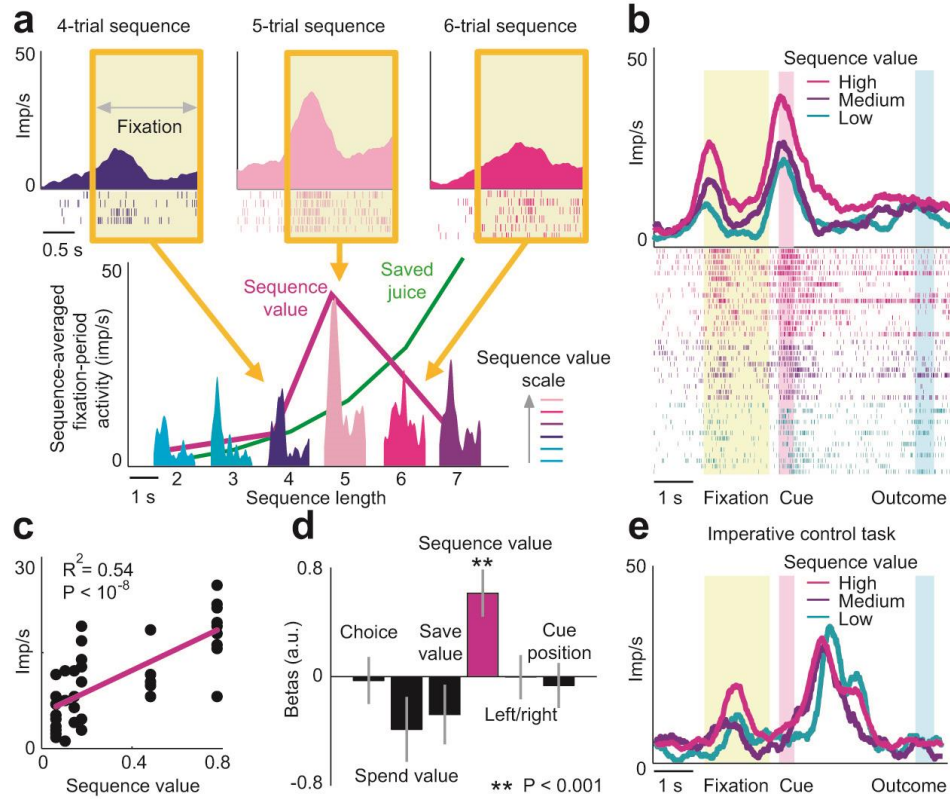
Figure 2 A single amygdala neuron with prospective activity that reflected the value of the monkey's internal saving plan. (**a**) Activity during step-by-step saving depended on the final saving sequence that the animal eventually produced. Specifically, activity depended on the subjective value of the current sequence ('sequence value'), which would only be achieved several trials ahead. Upper panels: activity (spike density functions) during three saving sequences of different lengths. Activity during fixation (yellow area) was highest for the sequence in which the monkey would eventually spend on the fifth trial, as this sequence had the highest subjective value (Imp/s: impulses per second; raster display: ticks indicate impulses, rows indicate trials). Lower panel: activity averages for all sequence lengths (e.g. light-pink activation indicates mean fixation activity for all five-trial sequences, averaged over trials one to five). Activity reflected sequence value (magenta curve, normalized), rather than linear sequence length or objective reward amount (green curve, normalized). Behaviorally derived sequence values reflected the animal's preferences for different combinations of sequence length and final reward—five-trial sequences had the highest value as the monkey chose them most frequently. Saving sequences were freely determined by the animal; visual stimulation was constant across sequences. (**b**) Within-trial activity sorted according to sequence value (terciles). (**c**) Linear

regression of activity on sequence value. Different value levels resulted from different sequence lengths as shown in (**a**). (**d**) Multiple regression coefficients (betas ± s.e.m., Eq. 6). (**e**) Activity in the imperative task, when saving was instructed, did not reflect sequence value.
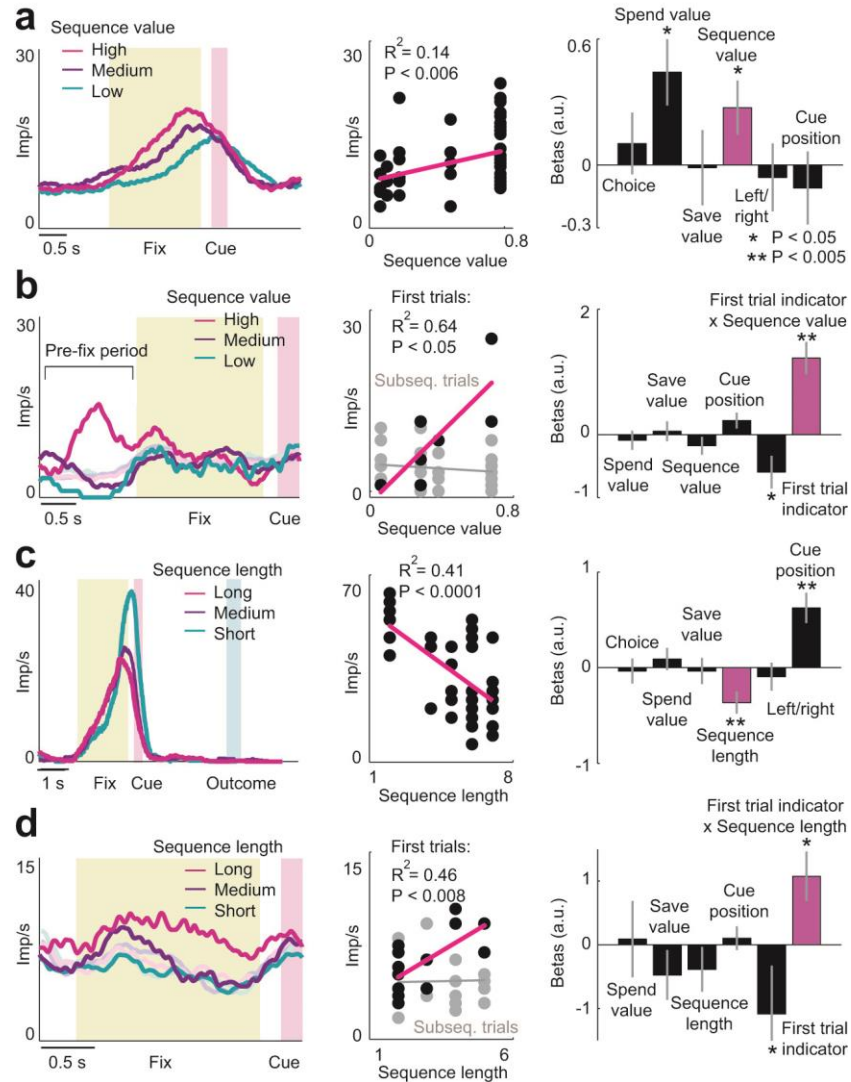
Figure 3 Different forms of planning activity in four single amygdala neurons. **(a)** Activity of this neuron, as in Fig. 2, reflected sequence value across all trials. Right panel: regression betas obtained by fitting Eq. 6 to neuronal activity. (**b**) Activity of this neuron at trial start before fixation ("Pre-fix period") reflected sequence value specifically on the first trial of each sequence (bold colors), but not on subsequent trials (light colors). Right panel: regression betas obtained by fitting Eq. 8 to neuronal activity. *First trial indicator*: indicator variable for the first trial in a saving sequence. *First trial indicator × sequence value*: regressor for testing sequence value coding specifically on first saving trials. (Statistics for first-trial effects were based on data from all save trials including bold and light colored data. The linear regression in the middle panel remained significant when the effect of outliers was reduced using robust regression.) **(c)** Activity of this neuron during the fixation and cue periods reflected final sequence length, rather

than sequence value, across all saving trials. Activity was higher for shorter sequences. Right panel: regression betas obtained by fitting Eq. 7 to neuronal activity. **(d)** Activity of this neuron in the fixation period reflected sequence length specifically on first saving trials. Right panel: regression betas obtained by fitting Eq. 9 to neuronal activity. *First trial indicator × sequence length*: regressor for testing sequence length coding specifically on first saving trials.
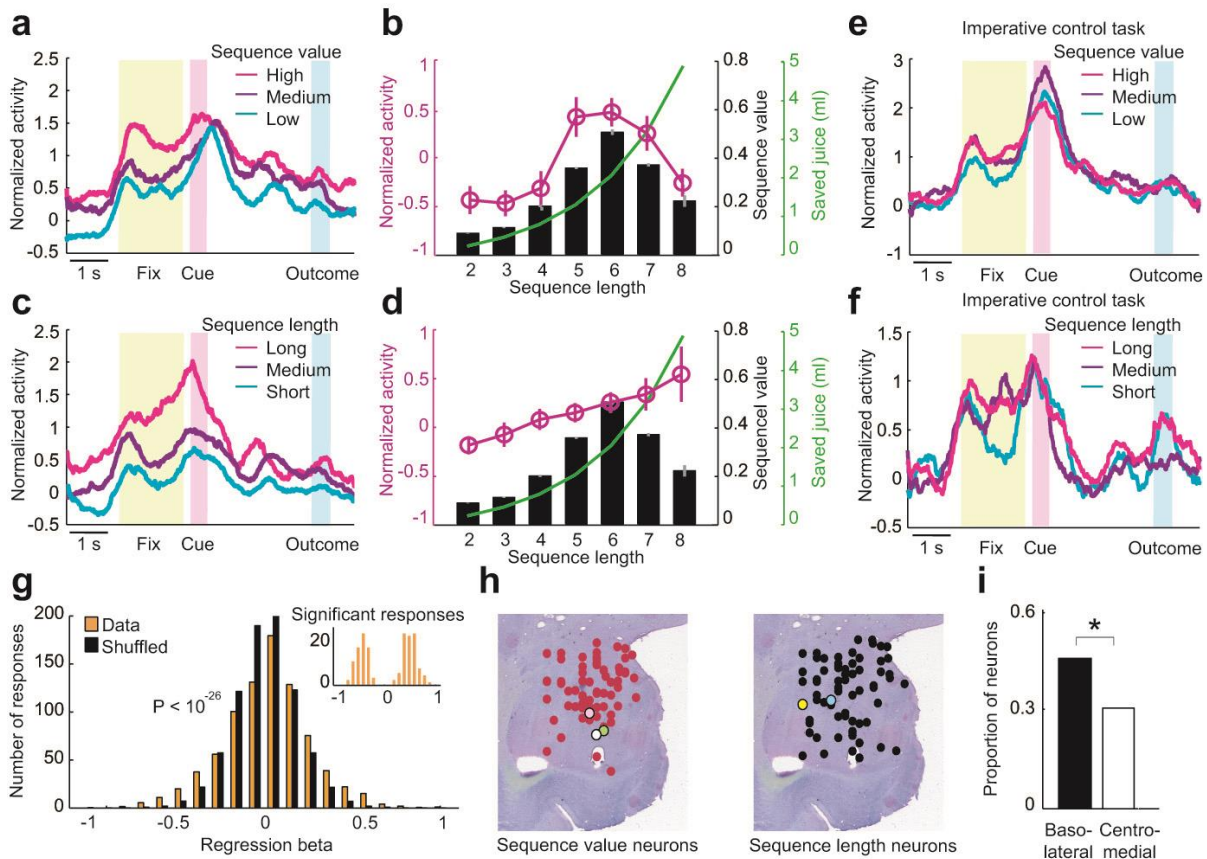
Figure 4 Planning activity in amygdala neurons: population data. **(a,b)** Planning activity (z-normalized) of 72 neurons encoding sequence value across all trials or specifically on first trials. (**b**) Population activity (magenta, $n = 93$ responses) reflected sequence value ($r^2 = 0.91$, $P < 0.0001$, linear regression, $n = 7$) rather than sequence length ($r^2 = 0.38$, $P > 0.1$). **(c,b)** Planning activity of 71 neurons encoding sequence length across all trials or specifically on first trials. **(d)** Population activity (magenta, $n = 92$ responses) reflected sequence length ($r^2 = 0.85$, $P = 0.0035$, $n = 7$) rather than sequence value ($r^2 = 0.14$, $P > 0.4$). **(e, f)** Activity of neurons tested in the imperative task failed to reflect sequence value or sequence length when saving was instructed (data from 30 neurons encoding sequence value and 29 neurons encoding sequence length). **(g)** Regression betas for observed data (orange, $n = 829$ responses from 329 neurons, collapsed across sequence value and sequence length) and trial-shuffled data (black, scaled down 1,000 times). The distribution of observed data was shifted towards higher positive and negative values (Kolmogorov-Smirnov test). **(h)** Histological reconstruction of 72 sequence value neurons and 71 sequence length neurons. Green, white, pink, yellow and blue symbols: example neurons in

Fig. 2 and Fig. 3a–d, respectively. Collapsing across anterior-posterior dimension resulted in symbol overlap. **(i)** Proportion of neurons with planning activity (n = 123 neurons, collapsed across sequence value and sequence length) in basolateral and centromedial amygdala (P = 0.005, χ^2-test) and corresponding recording depths (reference: bregma).
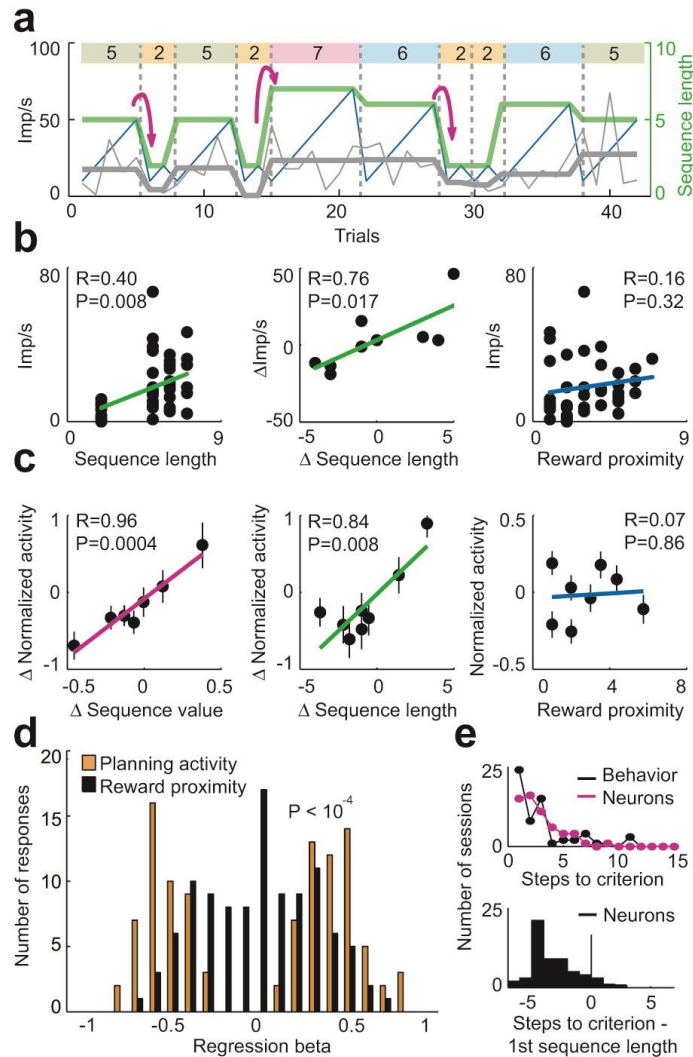
Figure 5 Adaptation dynamics of planning activity, reward proximity control. **(a)** Sequence-by-sequence adaptation in a single neuron encoding sequence length. Activity changes from spend to save trials (dashed lines) reflected changes in sequence length between successive sequences. Gray curves: sequence-averaged activity (thick line) and trial-by-trial activity (thin line). Green curve: sequence length. Blue curve: within-sequence reward proximity. Arrows: examples for activity changes scaling with sequence length changes. Colored boxes indicate sequences and corresponding lengths. **(b)** Linear regression of activity of the neuron in **(a)** on sequence length (left, $n = 41$), difference in length between subsequent sequences (*ΔSequence length*, middle, $n = 7$), and reward proximity (right, $n = 41$). **(c)** Population data. Left: sequence value responses ($n = 61$); activity changes at sequence transitions reflected changes in sequence value (linear regression). Middle: sequence length responses ($n = 55$); activity changes reflected changes in sequence length. Right: Population activity (sequence value and sequence length responses, $n =$

116) was unrelated to within-sequence reward proximity. **(d)** Regression betas for planning activity and reward proximity ($n = 116$ sequence value and sequence length responses, Kolmogorov-Smirnov test). **(e)** Behavioral-neuronal adaptation in sequence value neurons. Upper: With a new testing session, planning activity adapted readily to current interest rate, in-step with behavior ($r = 0.82$, $P = 1.7 \times 10^{-4}$; both Medians $= 1$, $n = 61$). Lower: Neurons typically reached adaptation criterion within the first sequence (Median $= -3$, implying adaption within 3 trials before end of first sequence, $t_{60} = -10.17$, $P = 1.0 \times 10^{-14}$, one-sample t-test).
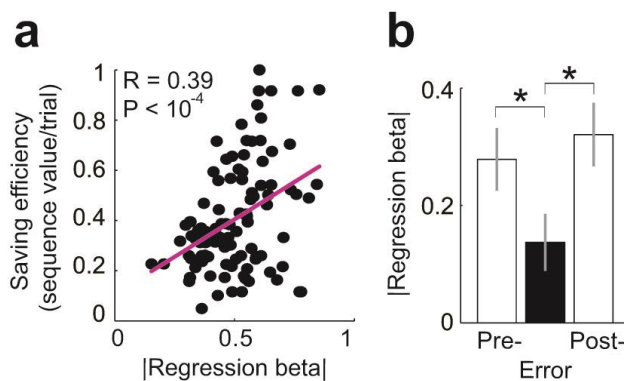
Figure 6 Relationship between amygdala planning activity and behavioral performance. **(a)** Relationship to saving efficiency. Stronger planning activity (sign-corrected regression betas, collapsed across responses encoding sequence value or sequence length across all trials, $n = 116$) predicted behavioral saving efficiency (accumulated sequence value per unit time, normalized, linear regression). This effect was confirmed in a partial correlation analysis ($P < 0.001$) that factored out potential confounding variables. **(b)** Relationship to performance errors. Bars show regression betas ($\pm$ s.e.m) from a population analysis (combining sequence value and sequence length responses, $n = 116$) for trials immediately preceding errors (Pre-), error trials (Error), and trials following errors (Post-). The relationship between activity and planning variables was significantly reduced on error trials, when the animals failed to progress towards their saving goal ($t_{1453} = -2.69$, $P < 0.01$, dependent-samples t-test comparing betas on pre-error and error trials), and subsequently reappeared after error correction ($t_{1453} = 3.47$, $P < 0.001$, dependent-samples t-test comparing betas on error and post-error trials).

Table 1. Number of neurons with planning activity.

	All trials[1]		First trials[2]		Combined[3]
	Sequence value	Sequence length	Sequence value	Sequence length	Sequence value/length
Animal A	27 (15%)	21 (12%)	20 (11%)	15 (8%)	66 (36%)
Animal B	19 (13%)	24 (16%)	10 (7%)	18 (12%)	57 (39%)
Total	46 (14%)	45 (14%)	30 (9%)	33 (10%)	123 (37%)

1: Neurons encoding the planning variables sequence value or sequence length across all trials in a saving sequence. Percentages calculated with respect to 181 neurons in animal A, 148 neurons in animal B, and 329 task-related neurons in both animals. Percentages are referenced to the number of neurons that were recorded because they were task-related, i.e. responsive to events in the saving task.

2: Neurons encoding planning variables specifically on first saving trials.

3: Neurons encoding planning variables either across all trials or specifically on first saving trials. The number of neurons in this column can be smaller than the row sum as some neurons showed multiple significant effects.

Author contributions: I.H., F.G. and W.S. designed the research. I.H. performed the experiments. F.G. analysed the data. F.G. and W.S wrote the manuscript.

METHODS

Neurophysiological recordings. All animal procedures conformed to US National Institutes of Health Guidelines and were approved by the Home Office of the United Kingdom. Experimental procedures for neurophysiological recordings from awake behaving macaque monkeys have previously been described[25,34,51]. Two adult male rhesus monkeys (Macaca mulatta) weighing 9.2 and 12.0 kg participated in the experiments. The number of animals used is typical for primate neurophysiology experiments. The animals had no history of participation in previous experiments. A head holder and a recording chamber were fixed to the skull under general anesthesia and aseptic conditions. We located the amygdala from bone marks on coronal and sagittal radiographs taken with a guide cannula and electrode inserted at a known coordinate in reference to the stereotaxically implanted chamber[52]. We recorded activity from single amygdala neurons from extracellular positions during task performance, using standard electrophysiological techniques including on-line visualization and threshold discrimination of neuronal impulses on oscilloscopes. We aimed to record representative neuronal samples from the dorsal, lateral, and basal amygdala.

We sampled activity from about 700 amygdala neurons in exploratory tests with the reward-saving task. We recorded and saved the activity of neurons that appeared to respond to at least one task event during online inspection of several trials. This procedure resulted in a database of 329 neurons with task-related responses which we analyzed statistically. The number of neurons is similar to those reported in previous studies on primate amygdala[21,25]. We aimed to identify neurons that were generally task-responsive but did not screen selectively for planning activity. Accordingly, statements about the proportion of amygdala neurons with planning activity refer to the proportion of neurons that we found to be related to the behavioral events in the saving task.

After completion of data collection, recording sites were marked with small electrolytic lesions (15–20 µA, 20–60 s). The animals received an overdose of pentobarbital sodium (90 mg/kg iv) and were perfused with 4% paraformaldehyde in 0.1 M phosphate buffer through the left ventricle of the heart. Recording positions were reconstructed from 50-µm-thick, stereotaxically oriented coronal brain sections stained with cresyl violet. The histological reconstructions validated also the previously radiographically assessed anatomical position of the

amygdala in agreement with earlier reports[25,52]. For **Fig. 4h,i**, we collapsed recording sites from both monkeys spanning 3 mm in the anterior-posterior dimension onto the same coronal section.

Behavioral task. On each trial (**Fig. 1a**) the monkey chose to either save the liquid reward that was available on that trial, which increased its magnitude by a variable 'interest rate', or spend the saved reward for immediate consumption. (The term 'interest rate' provides an intuitive description of the variable that governed increases in reward across save choices; this should not imply exact comparability with human economic saving.) The increase of reward magnitude over successive save choices was determined by a geometric series (Eq. 1)

$$x_n = b \sum_{i=0}^{n-1} q^i$$

with x_n as reward magnitude on trial n, b as base rate of reward magnitude, and q as interest rate, resulting in geometric increases for higher interest rates. Different liquid volumes were delivered using different opening durations of the solenoid valve. Monkeys were free to produce saving sequences of various lengths, i.e. saving behavior was self-determined (following one required save choice per sequence). We found that in early stages of task training the animals were unable to drink more than 8 ml on a single trial. Accordingly, for the high interest rate condition (**Fig. 1b**, upper panel) we adjusted the reward magnitude so that reward stagnated after 7 consecutive save trials at 8 ml. However, the animals were still free to produce longer saving sequences, i.e. we did not impose an upper limit on the sequence length. By the time of neuronal recordings, the animals only generated saving sequences that resulted in reward amounts that they could comfortably drink. The animals initiated trials by placing their hand on an immobile, touch-sensitive key. The trial then started with an ocular fixation spot of 1.3° of visual angle at the center of the computer monitor. Animals were required to keep their gaze on the fixation spot at stimulus center within 2–4°. Eye position was monitored using an infrared eye tracking system at 125 Hz (ETL200; ISCAN). At 1,500 ms plus mean of 500 ms (truncated exponential distribution) after fixation spot onset, the two save and spend visual stimuli of 7.0° appeared on the left and right side of the computer monitor (pseudorandomized). The cues were approximately similar in luminance. In different blocks of typically 40–100 consecutive trials,

different stimuli were used as save cues to indicate different interest rates. Animals indicated their choice with a saccade as soon as the visual cues appeared. The chosen stimulus was then replaced by a peripheral fixation spot of 7.0° of visual angle. After a delay period of 1,500 ms a color change of the peripheral fixation spot served as a 'Go' signal for the monkey to release the touch key. The release of the touch key was followed by the delivery of the reinforcer (an auditory or visual cue on save trials vs. a drop of juice reward on spend trials). For most recording sessions, we used an auditory cue as secondary reinforcer on save trials, which signaled successful trial completion without providing information about saved reward amount. Thus, animals had to track internally the accumulated reward amounts during saving behavior. Failures of key touch or fixation breaks were considered errors and resulted in trial cancellation. More than three sequential errors led to a pause in behavioral testing. Accumulated saved rewards were retained across error trials. The animals were overtrained by the time of neuronal recording and showed consistent, meaningful saving behavior for different interest rates without further signs of learning (**Fig. 1b**, **Supplementary Fig. 1a,b**).

To provide an example of how rewards were calculated, consider a series of two successive save choices by the monkey with a base rate of reward $b = 0.11$ and interest rate $q = 1.5$. On the second trial of the choice sequence, after the first save choice, reward $R = 0.11 \times (1 + 1.5) = 0.275$ ml. On the third trial, after two successive save choices, reward $R = 0.11 \times (1 + 1.5 + 1.5^2) = 0.523$ ml.

Each neuron was typically tested with one to two different interest rates. The duration required for testing neurons with statistically sufficient numbers of trials in both free choice and imperative tasks usually precluded using more than two interest rates.

Task training. We trained each animal during 3–4 months prior to neuronal recordings with the different visual stimuli and the different interest rates (300–400 trials/day, 5 days/week). Initially, the animals learned that responding to visual cues lead to reward delivery. We then introduced two different visual cues and taught the animals that choice of one of the cues lead to reward if the other cue had been chosen immediately before. This helped to train the animals to alternate choices between save and spend cues. We then introduced interest rates in the form of different save cues and the monkeys learned the underlying reward contingencies by sampling different sequence lengths. Thus, we did not shape the animals behavior towards producing

different sequence lengths at different interest rates. In parallel, we introduced imperative trials using the same cues with variable sequence lengths, with a small visual stimulus indicating the correct choice on each trial. We proceeded to neuronal recordings when performance in control tasks (see below) indicated that the animals adapted their choices to interest rate in a meaningful and flexible manner.

Rewards. A computer-controlled solenoid valve delivered juice reward from a spout in front of the animal's mouth (valve opening time of 100 ms corresponding to 0.38 ml). For monkey A the base rate of reward magnitude, b from equation 1, was set to 0.11 ml for all sessions, for monkey B the base rate was set to 0.11 ml for half of the sessions and 0.13 ml for the other half of the sessions. The animal's tongue interrupted an infrared light beam below the adequately positioned spout. An optosensor monitored licking behavior with 0.5-ms resolution (STM Sensor Technology).

Imperative Control Task. In this control task, saving behavior was not self-controlled by the animals but was externally determined. A small visual cue was presented next to either the save or the spend cue to indicate the correct choice on each trial that was otherwise identical to a free choice trial. We matched the ratio of save to spend trials between imperative and free choice task for a given monkey and interest rate. This made it possible for the monkeys to anticipate final saving outcomes, as confirmed by analysis of behavioral reaction times (**Supplementary Fig. 2d**).

Control task with uncued changes in interest rate. To test the extent to which the monkeys adapted their saving behavior to changes in the interest rate even when interest rates changed without notification, we performed, in behavioral testing sessions, a variant of the free choice saving task in which the interest rate varied without associated changes in the visual save cue (**Fig. 1c**, **Supplementary Fig. 1e**). In this control test, we introduced a new, unfamiliar save cue on each day and varied the interest rate without notification in blocks of 40–100 trials that were randomly interleaved. The save cue was fixed throughout a testing session and the animal had to keep track of the current interest rate.

Control task with fixed reward. To test whether the monkeys kept track of the amount of reward they had accumulated through consecutive save choices, we offered them, on randomly interspersed trials, a choice between the accumulated reward and fixed amounts indicated by pre-trained visual cues (**Supplementary Fig. 1f**).

Data analysis

Description of saving behavior. We constructed distributions of the relative frequencies with which each animal produced saving sequences of specified length, separately for different interest rates. **Figure 1b** shows these frequency distributions averaged over animals for low ($q = 0.7$), medium ($q = 1.5$) and high ($q = 2.0$) interest rates. **Figure 1c** shows weighted means of these distributions pooled over animals. For calculation of these weighted means, each relative choice frequency was weighted by its corresponding sequence length. **Supplementary Fig.1a** shows distributions separately for both animals and for various interest rates.

Definition of subjective values. To model the animals' saving behavior trial-by-trial, we derived estimates of the subjective values that the animals likely associated with saving sequences and save/spend choice options. For unbiased estimates, we used one half of the choice data within each monkey and interest rate to estimate subjective values and the other half for analysis. For each interest rate, we measured the relative spending frequency at each step in a saving sequence (**Fig. 1b** black), and multiplied it with the objective reward magnitude that would results from spending on that trial (**Fig. 1b** green curve), in order to account for differences in reward magnitude between interest rate conditions. This measure constituted the subjective value of spending on each trial ('spend value', **Fig. 1b** magenta curve). Thus, the subjective value for spending, *SVspend,* at a given point i in a saving sequence was defined as (Eq. 2)

$$SVspend_i = P_i M_i,$$

where P_i is the probability with which the monkey produced a saving sequence of length i, and M_i is the objective reward magnitude in ml of juice that would result from spending at point i of the sequence length given the current interest rate. The spend value actually realized in a saving sequence constituted the value of the current sequence, which we labeled 'sequence value'.

Magenta curves in **Fig. 1b** show examples of subjective value functions for different interest rates. We defined the 'save value' for each trial as the average spend value that the animal could obtain in all future trials of that sequence. Accordingly, the save value for a given trial i not only depended on the spend value of the immediately following trial, SV_{i+1} but also on the spend values of other future trials of the current sequence (SV_{i+2}, SV_{i+3} etc.) Thus, the subjective value $SVsave$ for saving at a given point n in a save sequence was defined as (Eq. 3)

$$SVsave_n = \frac{1}{m-n} \sum_{i=n+1}^{m} SVspend_i,$$

with m defining the upper limit of the save sequence (given by the maximal observed sequence length for the monkey). Thus, spend value and save value reflected the animals' trial-by-trial valuations, whereas sequence value constituted the value of the current saving sequence.

Logistic regression analysis of choice data. To model the monkeys' trial-by-trial choices we used a multiple logistic regression analysis with the following general linear model (GLM):

GLM-1 (Eq. 4):

$$y = \beta_0 + \beta_1 SVspend + \beta_2 SVsave + \beta_3 Interest + \beta_4 Cue\ position + \beta_5 Juice/day + \beta_6 Monkey + \varepsilon,$$

with y as trial-by-trial save-spend choice (1 indicating save choice and 0 indicating spend choice), $SVspend$ and $SVsave$ as the subjective value of spending or saving on the current trial, $Interest$ as the current interest rate, $Cue\ position$ as the left-right position of the save cue on the current trial, $Juice/day$ as the amount of liquid already consumed on that day, $Monkey$ as animal identity, β_1 to β_6 as the corresponding slope parameter estimates, β_0 as constant and ε as residual.

Linear regression analysis of reaction times and licking durations. As a measure of the animals' trial-by-trial reward expectation, we analyzed the latencies with which the monkeys released the touch key at the end of the trial to initiate reinforcer delivery. We adopted this approach based on previous findings[53] and preliminary analyses which indicated that touch key

release latencies rather than saccade latencies reflected upcoming reward magnitudes. Reaction times were z-normalized separately for each animal within each experimental session by subtracting the session mean and dividing by the session standard deviation. To test whether sequence value influenced the animals' reaction times during saving, we used the following multiple regression model:

GLM-2 (Eq. 5):

$$y = \beta_0 + \beta_1\,Choice + \beta_2\,SVspend + \beta_3\,SVsave + \beta_4\,SVfinal + \beta_5\,Interest + \beta_6\,Juice/day + \beta_7\,Monkey + \varepsilon,$$

with y as reaction time (key release latency) and *SVfinal* as sequence value (all other regressors as defined for Eq. 4).

Analysis of neuronal data. We counted neuronal impulses in each neuron on correct trials relative to different task events with time windows that were fixed across neurons: 1,000 ms before fixation spot (Pre-fixation), 1,775 ms after fixation spot but before cues (Fixation, starting 25 ms after fixation spot onset), 300 ms after cues (Cue, starting 20 ms after cue onset), 1,500 ms post-choice delay (Delay, starting 25 ms after the animal had indicated its choice), and 500 ms during the reward/outcome period (Outcome, starting 50 ms after reinforcer onset). Our analysis followed established approaches to analyze neuronal data in reward structures with heterogeneous populations of neurons[35-37,54], as follows.

We first identified task-related responses in individual neurons and then used multiple regression analysis to test for different forms of planning activity while controlling for the most important behaviorally relevant covariates. We identified task-related responses by comparing activity in the Fixation, Cue, Delay and Outcome periods to a control period (Pre-fixation) using the Wilcoxon test ($P < 0.01$, Bonferroni-corrected for multiple comparisons). A neuron was included as task-related if its activity in at least one task period was significantly different to that in the control period. Because the Pre-fixation period served as control period we did not select for task-relatedness in this period and included all neurons with observed impulses in the analysis. We chose the pre-fixation period as control period because it was the earliest period at the start of a trial in which no sensory stimuli were presented.

We next used multiple regression analysis to assess relationships between neuronal activity and planning variables. The use of multiple regression was considered appropriate for the present data after testing assumptions of randomness of residuals, constancy of variance, and normality of error terms. Statistical significance of regression coefficients was determined using t-test with $P < 0.05$ as criterion, and was supported by a bootstrap as described in the Results. All tests performed were two-sided. Each neuronal response was tested with the following multiple regression models:

GLM-4a (Eq. 6):

$$y = \beta_0 + \beta_1\,Choice + \beta_2\,SVspend + \beta_3\,SVsave + \beta_4\,SVfinal + \beta_5\,Left/right + \beta_6\,Cue\ position + \varepsilon,$$

GLM-4b (Eq. 7):

$$y = \beta_0 + \beta_1\,Choice + \beta_2\,SVspend + \beta_3\,SVsave + \beta_4\,SeqLength + \beta_5\,Left/right + \beta_6\,Cue\ position + \varepsilon,$$

with y as trial-by-trial neuronal impulse rate, *SVfinal* as sequence value, *SeqLength* as sequence length and *Left/right* as an indicator function denoting whether the monkey made a saccade to the left or to the right (all other variables as defined above for GLM-2). GLM-4 was used to identify neurons whose activity reflected sequence value or sequence length across all trials within saving sequences. Coefficients for all regressors within a model were estimated simultaneously. Thus, significant regressors for sequence value or sequence length would indicate that a significant portion of the variation in neuronal impulse rate can be uniquely attributed to these variables.

The following models were used to test specifically for relationships between neuronal activity and planning activity on first saving trials.

GLM-5a (Eq. 8):

$$y = \beta_0 + \beta_1\,SVspend + \beta_2\,SVsave + \beta_3\,SVfinal + \beta_4\,FirstSave + \beta_5\,SVfinal \times FirstSave + \beta_6\,Cue\ position + \varepsilon$$

GLM-5b (Eq. 9):

$$y = \beta_0 + \beta_1\,SVspend + \beta_2\,SVsave + \beta_3\,SeqLength + \beta_4\,FirstSave + \beta_5\,SeqLength \times FirstSave + \beta_6\,Cue\ position + \varepsilon,$$

with y as the trial-by-trial neuronal impulse rate on all save trials (excluding spend trials), *FirstSave* as an indicator function denoting the first trial within each saving sequence, *SVfinal×FirstSave* as an interaction term to model sequence value coding specifically on first trials, and *SeqLength×FirstSave* as an interaction term to model sequence length coding specifically on first trials. To limit the number of regressors in the model, we only considered save trials for this analysis; therefore, GLM-5 did not include regressors for the trial-specific save-spend choice (which was constant). We also did not include a regressor for the current-trial left-right action as few neurons (< 5%) showed effects related to action choice in initial exploratory analysis.

We analyzed all task-related responses with the GLMs described in Eq. 6-9 to test for significance of the regression coefficient related to planning activity in each model. A task-related response was categorized as planning activity if it had a significant regressor for sequence value or sequence length in GLM-4 or in GLM-5. In cases were both sequence value and sequence length regressors were significant, we calculated coefficients of partial determination (CPDs)—a measure of the variance explained by one regressor in a multiple regression model— and assigned the response to the category with the higher CPD. CPDs were calculated as $CPD(X_i) = [SSE(X_{-i}) - SSE(X_{-i}, X_i)]\ /\ SSE(X_{-i})$, with $SSE(X)$ indicating the sum of squared errors in a regression model that includes a set of regressors X_i, and X_{-i} indicating the set of regressors that includes all regressors except X_i. For most planning activities (94.6%), this approach allowed clear categorization as either sequence value-coding or sequence length-coding. Some remaining responses with equal CPDs (5.4%) were categorized as sequence value-coding as this was our *a priori* hypothesis for a reward structure. Exclusion of these few ambiguous responses did not alter any results or conclusions.

We followed standard procedures[55,56] to confirm that our regression approach was not compromised by multicollinearity. First, we confirmed that our results were robust to variations

in statistical modelling when predictor variables were added or deleted (see below). Second, inspection of correlation matrices revealed that correlations between variables were within acceptable ranges (e.g. the average correlation between SVsave and SVspend was –0.178). Third, we confirmed that variance inflation factors (VIFs) for the behavioral GLMs were < 3, and thus well below the cut-off recommended in statistical literature[56]. For all neuronal GLMs, VIFs were equally low (Mean = 2.48 ± 0.13 *s.e.m.*) and 95% of VIFs were < 3.44 (VIFs calculated separately within each neuronal testing sessions).

We evaluated the extent to which our key findings were robust to variations in statistical modeling using alternative analysis windows and regression models. Results for the fixation period were robust (< 5% change in number of significant responses) to changes in analysis window (200, 250, or 350 ms offset after fixation or restricting analysis window to 350 ms offset unril 1,500 ms post fixation). Further, compared to 123 neurons with planning activity obtained in our main regression models, we obtained the following numbers in a series of alternative models: 130 neurons when GLM-4 and GLM-5 were combined into one model, 125 neurons when including reward proximity as a covariate, 134 neurons when including reaction time as a covariate, 128 neurons when including an autoregressive term of neuronal impulse rate as covariate, 120 neurons when choice probability was included as covariate, and 115 neurons when using single linear regression models. Finally, a stepwise variable selection[55] procedure with all variables in Eq. 6-9 included in the starting set identified 135 neurons with planning activity (**Supplementary Table 1**).

Normalization of population activity. We subtracted from the measured impulse rate in a given task period the mean impulse rate of the control period and divided by the standard deviation of the control period (z-score normalization). Next, we distinguished neurons that showed a positive relationship to sequence value or sequence length and those with a negative relationship, based on the sign of the regression coefficient, and sign-corrected responses with a negative relationship.

Normalization of regression coefficients. Standardized regression coefficients were defined as $x_i(s_i/s_y)$, x_i being the raw slope coefficient for regressor i, and s_i and s_y the standard deviations of independent variable i and the dependent variable, respectively.

Analysis of neuronal adaptation dynamics. To examine behavioral and neuronal adaptation to the current interest rate at the start of a new testing session, we defined a criterion for behavioral adaptation as the number of sequences that the animal produced before it produced its 'preferred' sequence for the first time. The preferred sequence was the one with the highest sequence value given the current interest rate. Our rationale was that each interest rate condition was characterized by a subjective value function that depended on the animal's choice preferences. The animals would then adapt to current interest rate by changing their behavior according to this value function, and corresponding changes might be seen in sequence value neurons. To examine neuronal adaptation in sequence value neurons, a criterion for neuronal adaptation was defined as the number of trials before the neuronal response to the preferred sequence was within 0.5 s.d. of the neuron's mean response to that sequence. (Very similar results were obtained if this criterion was adjusted to 1 or 1.5 standard deviations.) For all sequence value responses, the distribution of this neuronal adaptation criterion over sessions is plotted in **Fig. 5e** (upper panel, magenta data points represent means over responses). The lower panel in **Fig. 5e** shows the distribution of the difference between this criterion and the length of the preferred sequence. Thus, negative values on the x-axis in **Fig. 5e**, lower panel, indicate that the neuronal adaptation criterion was achieved during the first preferred sequence that the animal produced in that session.

Analysis across neurons. For **Fig. 6a**, we plotted the sign-corrected, standardized regression betas for each neuronal response against a measure of saving efficiency, defined as the cumulative sequence value that the animal obtained in the session in which the neuron was recorded, normalized to the number of trials in that session. For the error analysis shown in **Fig. 6b**, we selected trials immediately before the animal committed an error within a saving sequence ('Pre-error'), the error trial itself ('Error') and the subsequent trial ('Post-error'). Errors occurred when the animal failed to complete a trial due to a fixation error or release of touch key. We included trials on which an error occurred following the trial period in which the neuron exhibited planning activity. As this analysis matched the number of error and non-error trials, any observed effect could not be explained by lower statistical power for error trials. The regression coefficients shown in **Fig. 6b** were obtained by performing single linear regressions of

normalized population activity on sequence value or sequence length, separately for pre-error, error, and post-error trials.

Analysis of neuronal tuning to sequence length. For the analysis shown in **Supplementary Fig. 5**, we calculated a breadth of tuning metric that has previously been used to examine sensory tuning functions[57]. We determined the relative magnitude of the neuronal response to a specific sequence length (defined as the mean response to that sequence length, expressed as the proportion of the summed mean responses to all sequences). Based on these relative magnitudes, the breadth of tuning metric was calculated as (Eq. 10)

$$H = -K \sum_{i=1}^{n} p_i \log p_i$$

with H as breadth of tuning, K as scaling constant (set so that $H = 1.0$ if the neuron had equal responses to all sequence lengths in the set of n sequence lengths), and p_i as the response to a given sequence length, expressed as the proportion of the total (summed) response to all sequences. The set of sequences considered for each neuron was determined by the range of sequences that the animal produced while the neuron was recorded. The metric ranges from 0 to 1.0, with 0 indicating total specificity to one sequence length and 1.0 indicating equal responses to all sequences.

A **Supplementary Methods Checklist** is available.

References

1. Benhabib, J. & Bisin, A. Modeling internal commitment mechanisms and self-control: A neuroeconomics approach to consumption-saving decisions. *Games Econ. Behav.* **52**, 460-492 (2005).
2. Dickinson, A. & Balleine, B. Motivational control of goal-directed action. *Anim. Learning & Behav.* **22**, 1-22 (1994).
3. Berkman, E.T. & Lieberman, M.D. The neuroscience of goal pursuit. in *The Psychology of Goals* (ed. G. Moskowitz & H. Grant) 98-126 (Guilford Press, New York, 2009).
4. Tanji, J. Sequential organization of multiple movements: involvement of cortical motor areas. *Annu. Rev. Neurosci.* **24**, 631-651 (2001).
5. Andersen, R.A. & Cui, H. Intention, action planning, and decision making in parietal-frontal circuits. *Neuron* **63**, 568-583 (2009).
6. Passingham, R.E. & Wise, S.P. *The Neurobiology of the Prefrontal Cortex* (Oxford University Press, Oxford, 2012).
7. Mushiake, H., Saito, N., Sakamoto, K., Itoyama, Y. & Tanji, J. Activity in the lateral prefrontal cortex reflects multiple steps of future events in action plans. *Neuron* **50**, 631-641 (2006).
8. Saito, N., Mushiake, H., Sakamoto, K., Itoyama, Y. & Tanji, J. Representation of immediate and final behavioral goals in the monkey prefrontal cortex during an instructed delay period. *Cereb. Cortex* **15**, 1535-1546 (2005).
9. Shima, K., Isoda, M., Mushiake, H. & Tanji, J. Categorization of behavioural sequences in the prefrontal cortex. *Nature* **445**, 315-318 (2007).
10. Shima, K., Mushiake, H., Saito, N. & Tanji, J. Role for cells in the presupplementary motor area in updating motor plans. *Proc. Natl. Acad. Sci. U S A* **93**, 8694-8698 (1996).
11. Lee, I.H., Seitz, A.R. & Assad, J.A. Activity of tonically active neurons in the monkey putamen during initiation and withholding of movement. *J. Neurophysiol.* **95**, 2391-2403 (2006).
12. Okano, K. & Tanji, J. Neuronal activities in the primate motor fields of the agranular frontal cortex preceding visually triggered and self-paced movement. *Exp. Brain Res.* **66**, 155-166 (1987).
13. Romo, R. & Schultz, W. Neuronal activity preceding self-initiated or externally timed arm movements in area 6 of monkey cortex. *Exp. Brain Res.* **67**, 656-662 (1987).
14. Kurata, K. & Wise, S.P. Premotor and supplementary motor cortex in rhesus monkeys: neuronal activity during externally- and internally-instructed motor tasks. *Exp. Brain Res.* **72**, 237-248 (1988).
15. Rolls, E.T. *Emotion and Decision-Making Explained* (Oxford University Press, Oxford, 2014).
16. Baxter, M.G. & Murray, E.A. The amygdala and reward. *Nat. Rev. Neurosci.* **3**, 563-573 (2002).
17. Cardinal, N., Parkinson, J.A., Hall, J. & Everitt, B.J. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci. Biobehav. Rev.* **26**, 321-352 (2002).
18. Seymour, B. & Dolan, R. Emotion, decision making, and the amygdala. *Neuron* **58**, 662-671 (2008).
19. Murray, E.A. & Rudebeck, P.H. The drive to strive: goal generation based on current needs. *Frontiers Neurosci.* **7**, 112 (2013).
20. Schultz, W. Multiple reward signals in the brain. *Nat. Rev. Neurosci.* **1**, 199-207 (2000).
21. Paton, J.J., Belova, M.A., Morrison, S.E. & Salzman, C.D. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* **439**, 865-870 (2006).
22. Schoenbaum, G., Chiba, A.A. & Gallagher, M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat. Neurosci.* **1**, 155-159 (1998).
23. Price, J.L. Comparative aspects of amygdala connectivity. *Ann. N. Y. Acad. Sci.* **985**, 50-58 (2003).
24. Belova, M.A., Paton, J.J., Morrison, S.E. & Salzman, C.D. Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron* **55**, 970-984 (2007).
25. Bermudez, M.A., Gobel, C. & Schultz, W. Sensitivity to temporal reward structure in amygdala neurons. *Curr. Biol.: CB* **22**, 1839-1844 (2012).
26. Wellman, L.L., Gale, K. & Malkova, L. GABAA-mediated inhibition of basolateral amygdala blocks reward devaluation in macaques. *J. Neurosci.* **25**, 4577-4586 (2005).
27. Rudebeck, P.H. & Murray, E.A. Amygdala and orbitofrontal cortex lesions differentially influence choices during object reversal learning. *J. Neurosci.* **28**, 8338-8343 (2008).
28. Grabenhorst, F., Rolls, E.T., Parris, B.A. & D'Souza, A. How the brain represents the reward value of fat in the mouth. *Cereb. Cortex* **20**, 1082-1091 (2010).
29. Grabenhorst, F., Schulte, F.P., Maderwald, S. & Brand, M. Food labels promote healthy choices by a decision bias in the amygdala. *Neuroimage* **74**, 152-163 (2013).

30. De Martino, B., Kumaran, D., Seymour, B. & Dolan, R.J. Frames, biases, and rational decision-making in the human brain. *Science* **313**, 684-687 (2006).

31. Bechara, A., Damasio, H., Damasio, A.R. & Lee, G.P. Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J. Neurosci.* **19**, 5473-5481 (1999).

32. Brand, M., Grabenhorst, F., Starcke, K., Vandekerckhove, M.M. & Markowitsch, H.J. Role of the amygdala in decisions under ambiguity and decisions under risk: evidence from patients with Urbach-Wiethe disease. *Neuropsychologia* **45**, 1305-1317 (2007).

33. Sugase-Miyamoto, Y. & Richmond, B.J. Neuronal signals in the monkey basolateral amygdala during reward schedules. *J. Neurosci.* **25**, 11071-11083 (2005).

34. Grabenhorst, F., Hernadi, I. & Schultz, W. Prediction of economic choice by primate amygdala neurons. *Proc. Natl. Acad. Sci. U S A* **109**, 18950-18955 (2012).

35. Padoa-Schioppa, C. Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* **34**, 333-359 (2011).

36. Lee, D., Seo, H. & Jung, M.W. Neural basis of reinforcement learning and decision making. *Annu. Rev. Neurosci.* **35**, 287-308 (2012).

37. Kennerley, S.W., Behrens, T.E. & Wallis, J.D. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat. Neurosci.* **14**, 1581-1589 (2011).

38. Hayden, B.Y., Pearson, J.M. & Platt, M.L. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat. Neurosci.* **14**, 933-939 (2011).

39. Sugrue, L.P., Corrado, G.S. & Newsome, W.T. Matching behavior and the representation of value in the parietal cortex. *Science* **304**, 1782-1787 (2004).

40. Belova, M.A., Paton, J.J. & Salzman, C.D. Moment-to-moment tracking of state value in the amygdala. *J. Neurosci.* **28**, 10023-10030 (2008).

41. Lak, A., Stauffer, W.R. & Schultz, W. Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proc. Natl. Acad. Sci. U S A* **111**, 2343-2348 (2014).

42. Kable, J.W. & Glimcher, P.W. The neurobiology of decision: consensus and controversy. *Neuron* **63**, 733-745 (2009).

43. Shima, K. & Tanji, J. Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* **282**, 1335-1338 (1998).

44. Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P. & Dolan, R.J. Model-based influences on humans' choices and striatal prediction errors. *Neuron* **69**, 1204-1215 (2011).

45. Gersch, T.M., Foley, N.C., Eisenberg, I. & Gottlieb, J. Neural correlates of temporal credit assignment in the parietal lobe. *PLoS One* **9**, e88725 (2014).

46. Dragoi, G. & Tonegawa, S. Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature* **469**, 397-401 (2011).

47. Pfeiffer, B.E. & Foster, D.J. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* **497**, 74-79 (2013).

48. Rudebeck, P.H., Mitz, A.R., Chacko, R.V. & Murray, E.A. Effects of amygdala lesions on reward-value coding in orbital and medial prefrontal cortex. *Neuron* **80**, 1519-1531 (2013).

49. Pessoa, L. & Adolphs, R. Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. *Nat. Rev. Neurosci.* **11**, 773-783 (2010).

50. Bechara, A. Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective. *Nat. Neurosci.* **8**, 1458-1463 (2005).

51. Bermudez, M.A. & Schultz, W. Reward magnitude coding in primate amygdala neurons. *J. Neurophysiol.* **104**, 3424 (2010).

52. Aggleton, J.P. & Passingham, R.E. Stereotaxic surgery under X-ray guidance in the rhesus monkey, with special reference to the amygdala. *Exp. Brain Res.* **44**, 271-276 (1981).

53. Watanabe, M.*, et al.* Behavioral reactions reflecting differential reward expectations in monkeys. *Exp. Brain Res.* **140**, 511-518 (2001).

54. O'Neill, M. & Schultz, W. Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron* **68**, 789-800 (2010).

55. Draper, N.R. & Smith, H. *Applied Regression Analysis* (Wiley-Interscience, 1998).

56. Kutner, M.H., Nachtsheim, C.J., Neter, J. & W., L. *Applied linear statistical models* (McGraw-Hill Irwin, New York, 2004).

57. Smith, D.V. & Travers, J.B. A metric for the breadth of tuning of gustatory neurons. *Chem. Sens.* **4**, 215-219 (1979).

58.   Paxinos, G., Huang, X.-F. & Toga, A.W. *The rhesus monkey brain in stereotaxic coordinates* (Academic Press, San Diego, CA, 2000).