Title: Human foraging for primary rewards is guided by fronto-hippocampal dynamics of anticipation

Authors: Reiko Shintaki,[1,*] Daiki Tanaka,[1,*] Shinsuke Suzuki,[2] Takaaki Yoshimoto,[3] Norihiro Sadato,[3] Junichi Chikazoe,[3,4] Koji Jimura[1,5]

Affiliations:
[1] Department of Biosciences and Informatics, Keio University, Yokohama, Japan
[2] Brain, Mind & Markets Laboratory, Department of Finance, Faculty of Business and Economics, The University of Melbourne, Victoria, Australia
[3] Division of Cerebral Integration, National Institute of Physiological Sciences, Okazaki, Japan
[4] Araya, Inc., Tokyo, Japan
[5] Research Center for Brain Communication, Kochi University of Technology, Kami, Japan

*: These authors contributed equally to the study.

The authors declare no conflict of interest.

Correspondence should be addressed to:
Koji Jimura, Ph.D.
Department of Biosciences and Informatics,
Keio University,
3-14-1 Hiyoshi Kohoku-ku,
Yokohama, 223-0061, Japan
Email: jimura@bio.keio.ac.jp
Phone: +81-45-566-1625

Foraging is a fundamental food-seeking behavior in a wide range of species that enables survival in an uncertain world. During foraging, behavioral agents constantly face a trade-off between staying in their current location or exploring another. Despite ethological generality and importance of foraging, it remains unclear how the human brain guides continuous decision in such situations. Here we show that anticipatory activity dynamics in the anterior prefrontal cortex (aPFC) and hippocampus underpin foraging for primary rewards. While functional MRI was performed, humans foraged for real liquid rewards available after tens of seconds, and continuous decision during foraging was tracked by a dynamic pattern of brain activity that reflected anticipation of a future reward. When the dynamic anticipatory activity in the aPFC was enhanced, humans remained in their current environment, but when this activity diminished, they explored a new environment. Moreover, the anticipatory activity in the aPFC and hippocampus was associated with distinct decision strategies: aPFC activity was enhanced in humans adopting an exploratory strategy, whereas those remaining stationary showed enhanced activity in the hippocampus. Our results suggest that anticipatory dynamics in the fronto-hippocampal mechanisms underlie continuous decision-making during human foraging.

Introduction

Foraging is an ethologically important behavior for acquiring nutrition in an uncertain world, and has been observed in a wide range of species from nematodes to primates[1-3]. During foraging, animals constantly face a trade-off between staying in a current food source or exploring to find another. However, such continuous decision in foraging has rarely been examined in humans.

Expectations of food acquisition and temporal resources are fundamental factors involved in resolving the trade-off inherent in foraging[4-6]. For example, consider an angler who is fishing in a particular spot (patch) (Fig. 1a). After catching a fish in that spot, the angler must continuously decide, based on the evaluation of the environment, whether to stay in the current spot or explore in order to find another one. The angler will remain if they anticipate catching another fish in the current spot; if not, they will abandon the spot and explore in order to find another. As this example illustrates, anticipation of future prey may play a key role in the foraging behavior[7-9].

Foraging has been examined in various species using particular paradigms and environments optimized for individual species, which have provided spot pieces of evidence[5,6,10-12]. One key issue in understanding foraging across species is the variability in what behavioral agents seek. In non-human animals, foraging has been examined in real situations using primary rewards involving nutrition (e.g., food or liquid)[5,10-12]. For humans, on the other hand, most paradigms have used secondary rewards (e.g., money or tokens)[12] in hypothetical situations[6]. Critically, human choice behaviors differ between real primary rewards and hypothetical secondary rewards[13].

Another issue is choice situations. In non-human animal studies, choices were continuous: animals were always able to abandon a current source and explore to find a new one[5,10-12,14,15]. In human studies, humans did not face such a constant trade-off, but made a series of discrete stay-explore choices after every outcome delivery[6,16]. These discrepancies of reward type and choice situations make it hard to compare foraging behavior across species. Thus, the use of real nutritious rewards in a continuous trade-off for humans should provide significant insight to extend our understanding of foraging behavior. Yet it remains elusive how humans forage in this context.

The current study was designed to bridge the critical gap between human and animal studies. Human participants performed a foraging task for real liquid rewards during functional MRI scanning (Fig. 1b). The current task made it possible to examine

the links between anticipation and foraging, and to identify underlying neural mechanisms. Participants first received a real liquid reward delayed by tens of seconds in a particular environment (patch) (experience trial). Then, they foraged for another liquid reward in the same environment (foraging trial). In this foraging trial, they were not exactly sure when they would receive a reward, but could stop waiting and move on to a novel environment whenever they preferred.

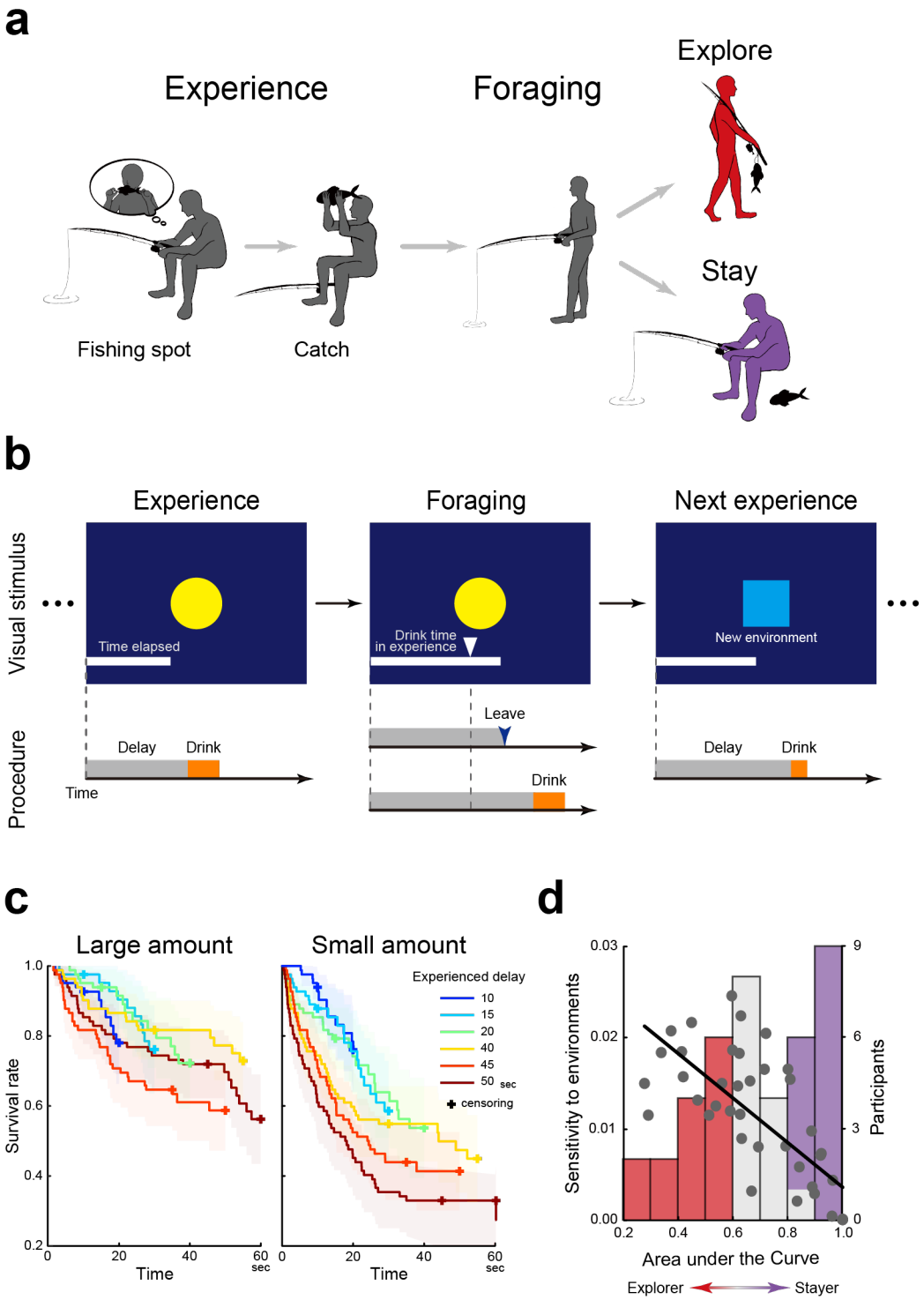


Figure 1. Humans forage for consumable rewards.
a, An illustration of foraging behavior. An angler catches a fish, and then decides whether to stay in the current spot or explore in order to find a new one. **b**, Behavioral procedure. Human participants were given an environment indicated by a picture (yellow circle). In this environment, they first waited for tens of seconds and consumed a real liquid reward (experience trial; *left*). Then, they waited for another liquid reward in the same environment (foraging trial; *middle*). During the experience and foraging trials, the elapsed time from the start of the trial was indicated by the length of a white horizontal bar extending to the right side of the screen. In the foraging trial, the reward delivery time in the experience trial was indicated by a white triangle. In both the experience and foraging trials, participants were unsure when

they would receive the reward. In the foraging trials, however, they were free to leave the current environment at any time and move on to the next experience trial with a new environment. The next environment was indicated by a different picture (blue square; *right*). Note that the picture indicating the environment was unique for each pair of experience and foraging trials. **c**, Survival rates in the foraging trial as a function of delay latency. Colors indicate delay durations in the experience trials. *Left*: large reward amount; *right*: small reward amount. Shaded areas indicate 95% CI. Crosses indicate censored events where participants continued to wait until reward delivery. **d**, The distribution of the area under the curve (AuC) of the survival functions of individual participants. The horizontal axis indicates the AuC of the survival rate, and the frequency of the participants is shown as a bar graph with the vertical axis on the right. Participants were labeled based on the AuC values as explorers (lowest tertile; red) or stayers (highest tertile; purple). The overlying scatter plot shows the correlation between the AuC and the sensitivity to the environmental parameters (reward amount and delay duration). Each plot denotes one participant.

Results

Foraging strategies reflects past experiences.

Participants stopped waiting more frequently and sooner after they experienced smaller amounts of a reward and a longer delay, as shown by survival functions reflecting the rate of trials in which participants continued to wait (Fig. 1c; reward amount: $\chi^2 = 49.0$, $P < .001$; delay duration: $\chi^2 = 26.1$, $P < .001$, log rank test). These results suggest that when the current environment was not promising in terms of acquiring another reward, they abandoned the environment and explored to identify a new one, as predicted by the optimal foraging theory[3,5,12,17,18]. Using laboratory setups where behavioral agents directly experience reward attainment, such foraging behavior has been observed in a wide range of species from non-mammals (e.g., nematodes, insects, and birds)[10,11] to mammals (e.g., rodents and primates)[5,12]. To our knowledge, our study is the first to demonstrate that like non-human animals, humans forage for primary rewards that are directly experienced in a laboratory setup.

The decision strategy in the foraging trial was quantified by the area under the curve (AuC) of the survival function for each trial condition across participants. A smaller AuC reflects the strategy of exploring to find a new environment, whereas a larger AuC indicates the strategy of staying in the current environment due to anticipation of a future reward. The AuC was smaller after delivery of a smaller amount of reward [$F(1, 40) = 16.6$, $P < .001$] and if there was a longer delay before the reward was delivered [$F(1, 40) = 37.9$, $P < .001$]. The adoption of an exploratory strategy was

enhanced by the joint experience of a smaller reward and longer delay [interaction of reward amount and delay duration: $F(1, 40) = 11.8$, $P < .01$].

For each participant across trials, the AuC value of the survival function was recalculated, and participants with the highest and lowest AuC tertiles were labeled as stayers (i.e., non-explorers) and explorers, respectively (Fig. 1d). Importantly, the explorers were more sensitive to environmental factors (reward amount and delay duration) than the stayers, in accordance with the optimal foraging theory ($r = -.73$, $P < .01$ with a permutation test; Fig. 1d and Supplementary Fig. 1). These results suggest that in the experience trial, explorers more frequently decided to abandon the current environment and move on to a novel one when they received a smaller reward after a longer delay, because the environment was not promising for them.

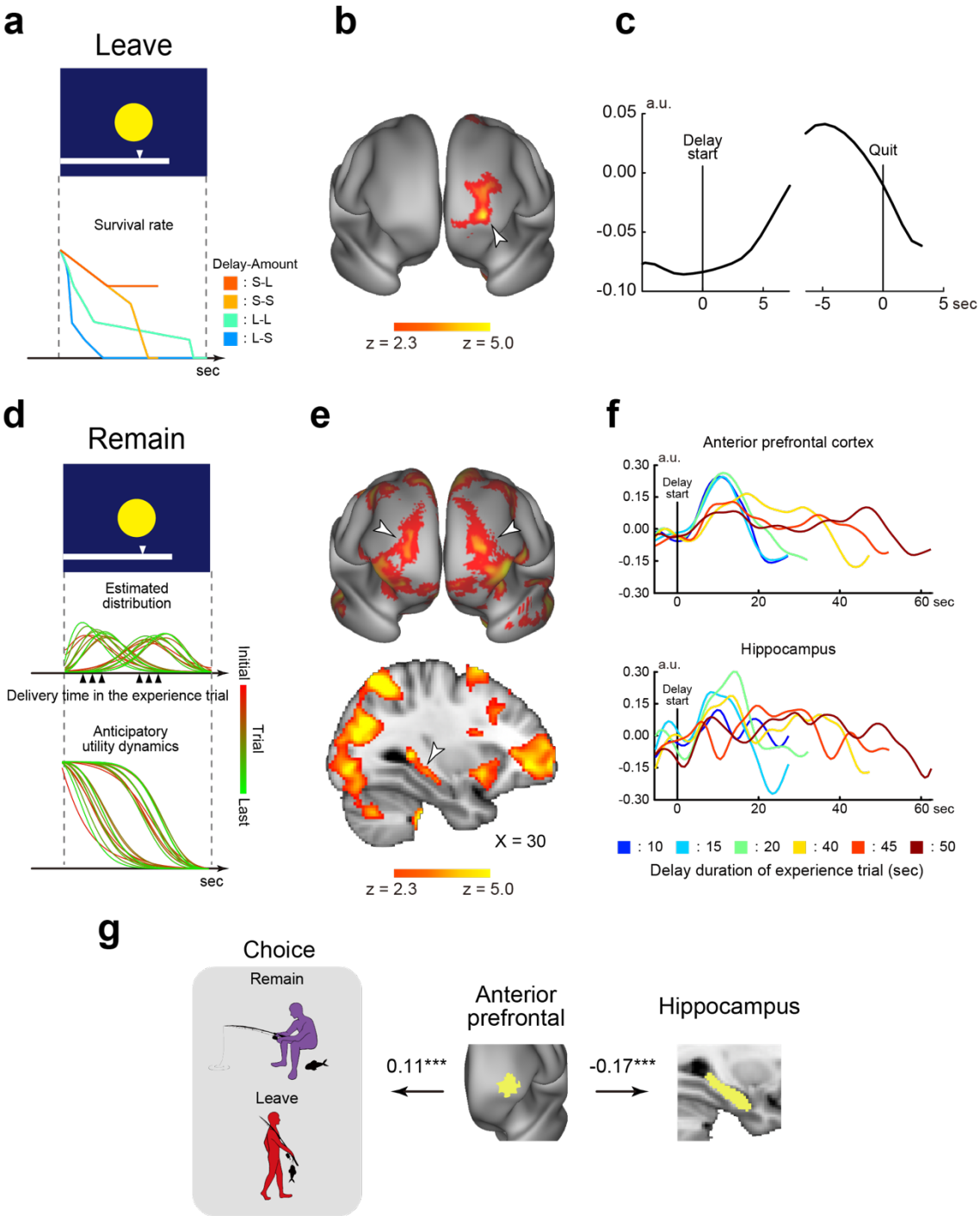


Figure 2. The anterior prefrontal cortex (aPFC) and hippocampus (HPC) showed anticipatory activity during foraging.

a, Anticipatory dynamics during the leave trials were modeled by an individual's survival function for each condition. Delay: Short (S)/Long (L); Reward amount: Small (S)/Large (L). **b**, Brain regions showing activation dynamics of the survival function during the delay period in the leave trials. Maps are overlaid onto the three-dimensional brain surface. The color bar indicates significance levels. The white arrowhead indicates the aPFC. **c**, The timecourse of brain activity in the aPFC. **d**, The probability density distribution of reward delivery expectation was updated throughout the experience and foraging trials (*middle*). Anticipatory utility (AU) dynamics during the remain trials were modeled as inverse dynamics of the cumulative probability of the density distribution (*bottom*). The colors of the lines indicate the trials in the color bar on the right. **e**, Brain regions showing AU dynamics during the delay period of the remain trials. Anterior view (*top*); sagittal section (*bottom*; X = 30). The arrowheads indicate the aPFC (*top*) and HPC (*bottom*). The formats are similar to those in Fig. 2b. **f**, The timecourses of activity in the aPFC (*top*)

and HPC (*bottom*). **g**, Choice behavior in the foraging trials was predicted by the aPFC activity, and the HPC activity was negatively coupled with the aPFC activity. Yellow areas indicate ROIs in the aPFC, defined based on prior studies, and in the HPC, which was defined anatomically. Values indicate regression coefficients. ***: P < .001.

Foraging involves anticipatory brain activity regarding a future reward.

In the foraging trial, participants continuously made stay-explore decisions, and stopped waiting when their anticipation of a future reward was reduced. Thus, the temporal characteristics (dynamics) of anticipation during the foraging is reflected in the survival functions coding for when participants stopped waiting (Fig. 1c). We then hypothesized that participants would leave the current environment when brain activity representing the anticipated reward was attenuated. This hypothesis was tested by exploring brain region showing dynamic brain activity reflecting survival functions from the start of the foraging trial until the onset of leaving the environment (leave trial; Fig. 2a; Supplementary Fig. 2a). Exploration of the whole brain revealed that the anterior prefrontal cortex (aPFC) showed dynamic anticipatory brain activity prominently (Fig. 2b; Supplementary Table 2). After its peak, aPFC activity gradually decreased toward leaving the current environment (Fig. 2c).

We next examined anticipatory brain activity in trials where participants continued to wait until a reward was delivered (remain trial). In theory, the survival function is not suitable to examine anticipatory activity in the remain trials because this function encodes when to stop waiting. Alternatively, the dynamics of anticipatory activity were modeled based on the theory of anticipatory utility (AU), based on the assumption that anticipation of a future reward itself provides pleasure and confers current utility[8,19,20]. Specifically, AU was modeled based on participants' expectation of reward delivery formulated by a probability density function, which was updated every time after participants performed foraging trial[21] (Fig. 2d; see Methods and Supplementary Fig. 2b/c).

Whole-brain exploratory analysis revealed a strong AU effect in the aPFC (Fig. 2e *top*). Interestingly, this region also showed anticipatory activity in the leave trials (Fig. 2b). The timecourse of aPFC activity revealed the temporal dynamics of AU (Fig. 2f *top*). The hippocampus (HPC) also showed a strong AU effect (Figs. 2e/f *bottom*; Supplementary Table 3), though such anticipatory activity was absent in the leave trials (Supplementary Fig. 3). The HPC activity during the remain trials may

reflect reproduction of the pleasure of anticipation encoded while participants waited for a reward in the experience trial[22-28], which may not be the case in the leave trials.

As suggested by previous studies, the value of a future reward was increased monotonically while the reward was anticipated[29,30]. The dynamics of the reward value are complementary to AU dynamics[9,19,21,31], and were modeled as the upcoming future reward (UFR; Supplementary Fig. 2d). The UFR effect was observed in multiple frontal and parietal regions (Supplementary Fig. 4a), but medially to the brain regions showing the AU effect (Figs. 2b/e), consistent with previous studies[19,21].

To compare the fit of AU and survival functions to the fMRI data, we performed a supplementary analysis by reversing the relationships between trials and models. Specifically, activity dynamics were modeled by the AU model for the leave trial, and by the survival model for the remain trial. In the aPFC, both the AU and survival models showed a significant effect (ts > 3.4; Ps < .005; Supplementary Fig. 5), but the survival model better fit to the leave trials than the remain trials [$t(35) = 2.5$; $P < .05$]. This suggests that aPFC activity is linked to leaving the current environment because the survival model directly codes the probability of leaving. In the HPC, on the other hand, in both models the effects were significant in the remain trials (ts > 2.1; Ps < .05) but were absent in the leave trials (ts < 0.6; Ps > .6), suggesting that HPC activity leads to remaining in the current environment.

<u>Foraging behavior is governed by anticipatory aPFC activity.</u>

Our results demonstrated that 1) the aPFC and HPC showed anticipatory activity in the remain trials (Fig. 2e), and 2) leaving an environment was associated with attenuated aPFC activity (Fig. 2c). We then hypothesized that 1) the continuous decision-making in the foraging trials would be predicted by aPFC activity, and 2) the HPC would functionally coordinated with the aPFC when the environment was not abandoned before reward delivery.

To test the former hypothesis, we performed a regression analysis in which anticipatory activity predicted whether there was departure until reward delivery or when a new environment was sought out during the foraging trials. The analysis was implemented by a multi-level mixed-effects general linear model (GLM) involving the aPFC or HPC activity in individual trials as predictors. The aPFC activity successfully predicted the choice behavior ($z = 3.8$; $P < .001$; Fig. 2g), indicating that weaker aPFC

activity is associated with leaving the current environment. The HPC did not predict the choice behavior (z = 0.77; P > .44).

The latter hypothesis was then tested by examining the functional connectivity between the HPC and aPFC based on a mixed-effects GLM. aPFC activity was negatively coupled with HPC activity in individual trials (z = 3.4; P < .001), suggesting that the aPFC and HPC play complementary roles in the choice to remain in the environment during foraging.

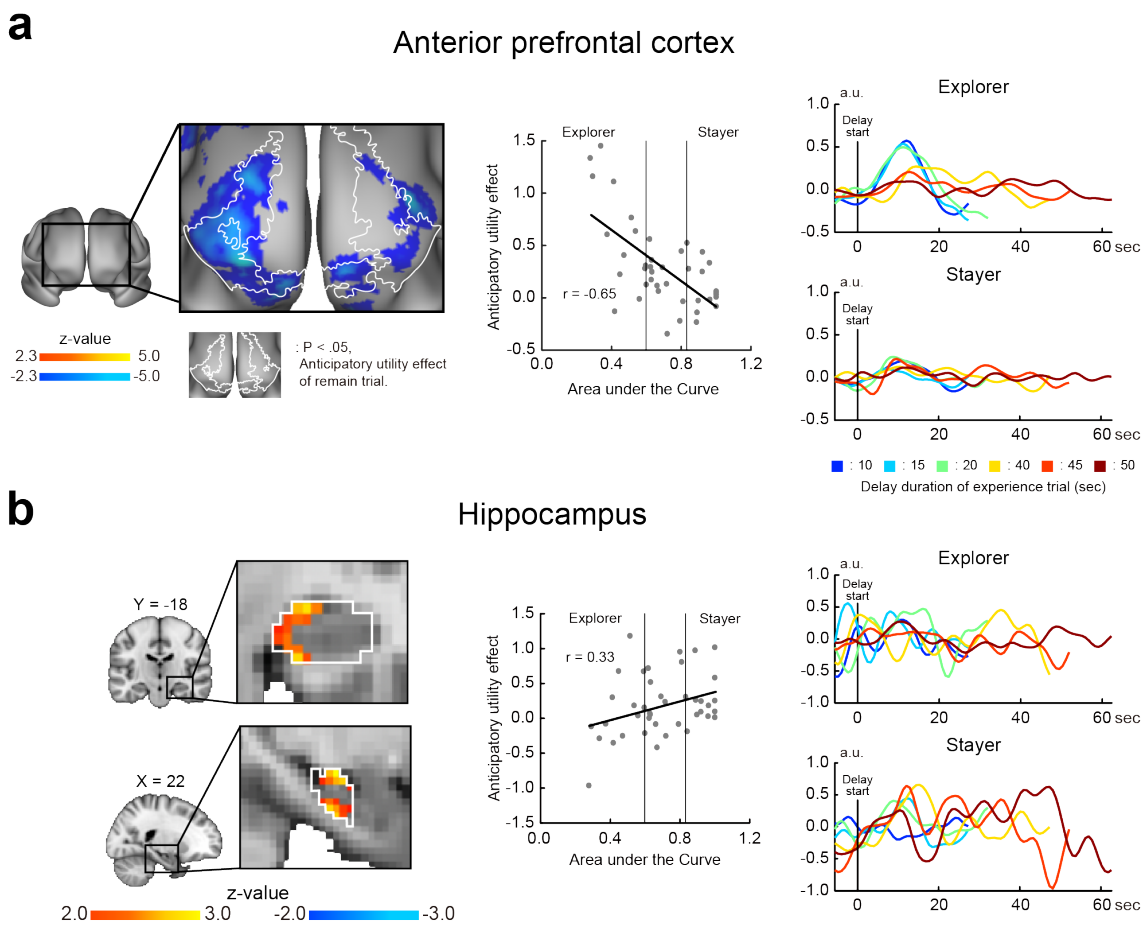


Figure 3. Stayers and explorers exhibit differential involvement of the aPFC and HPC.
a, Statistical maps of correlations between AU effects and AuCs were overlaid onto a three-dimensional brain surface (*left*). Cool and hot colors represent greater AU effects in explorers (smaller AuCs) and stayers (greater AuCs), respectively. The color bars indicate significance levels. The black rectangle on the map (*left*) represents the area magnified on the right. White closed lines indicate the aPFC regions showing AU effects in Fig. 2e *top*. *Middle*: A scatter diagram of AuC values and AU effects in the aPFC. Each plot denotes one participant. Timecourses of brain activity in the remain trials for explorers (*right top*) and stayers (*right bottom*). **b**, Statistical correlation maps in the HPC (*left*). White closed lines indicate anatomical borders of the HPC. Formats are similar to those in Fig. 3a.

aPFC and HPC differentially code foraging strategies.

Given the exploratory and stay (non-exploratory) strategies (Fig. 1d) and the anticipatory activity in the aPFC and HPC in the foraging trials (Figs. 2b-f), we next asked whether participants with distinct foraging strategies show differential involvement of these brain regions. To test this possibility, we explored brain regions showing cross-subject correlations between AU effects in the remain trials and the AuC values of individuals' survival functions.

The aPFC showed a strong negative correlation between the AU effect and AuC, (Fig. 3a *left*; Supplementary Table 4) indicating that explorers showed strong AU effects. In contrast, the HPC showed a positive correlation (Fig. 3b *left*; Supplementary Table 4), indicating that the stayers showed even stronger AU effects. Scatter plots of the AU effect against AuCs showed clear negative and positive correlations in the aPFC and HPC, respectively (Figs. 3a/b *middle*). Notably, these regions also showed strong AU effects (Figs. 2b/e). To show this correlation more specifically, timecourses of MRI signals in these regions were extracted for stayers and explorers as defined in Fig. 1d. Explorers showed strong anticipatory activity in the aPFC but not in the HPC, whereas stayer showed strong anticipatory activity in the HPC but not in the aPFC (Fig. 3a/b *right*). These results demonstrate that distinct foraging strategies involve the aPFC and HPC differentially in the remain trials.

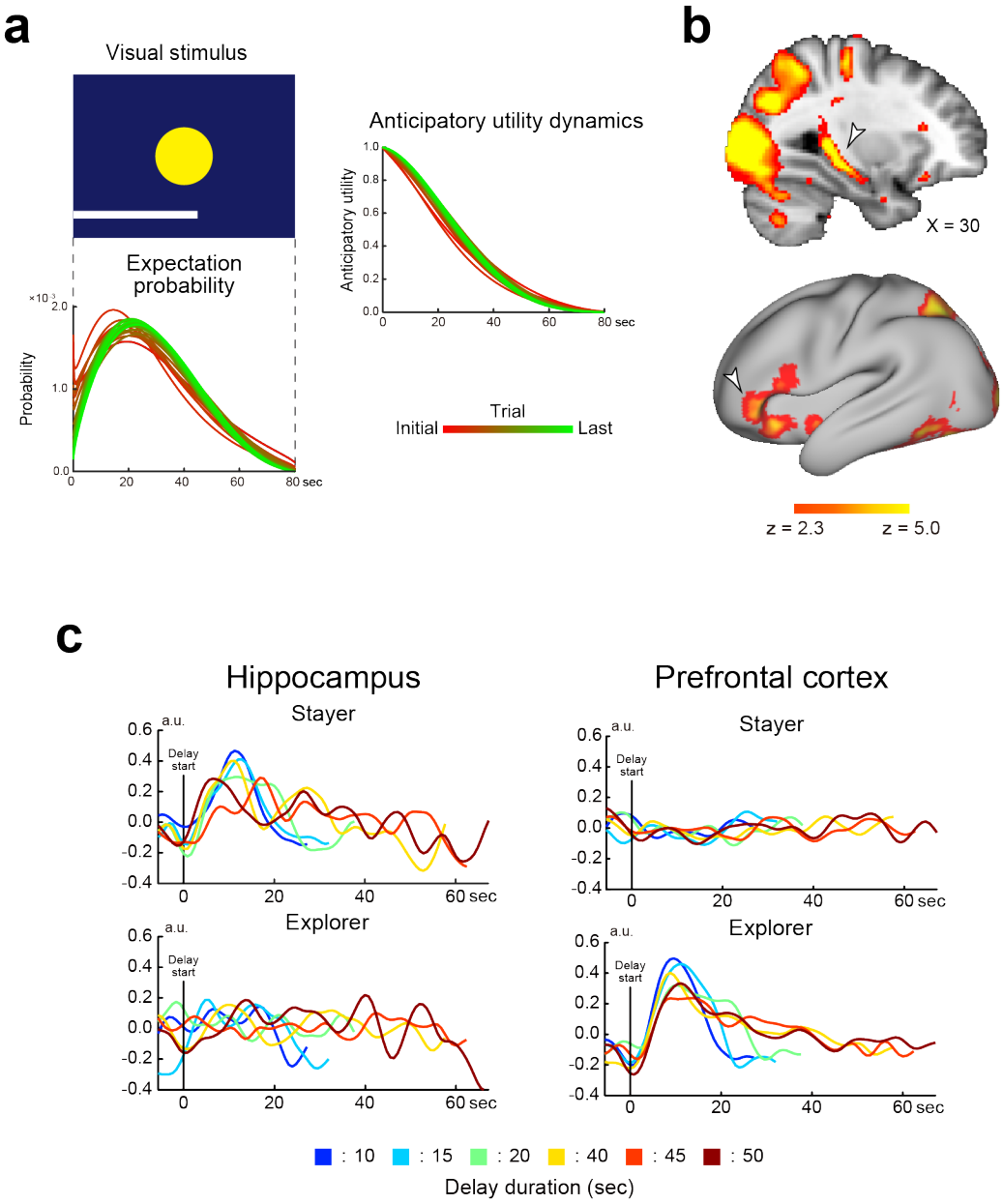


Figure 4. The experience of anticipation is encoded in the HPC.
a, The AU dynamics in the experience trials (*right*) were modeled similarly to those in the remain foraging trials. Formats are similar to those in Fig. 2d. **b**, Brain regions showing AU dynamics during the delay period of the experience trials. The arrowheads indicate the effects in the HPC (*top*) and aPFC (*bottom*). Formats are similar to those in Fig. 2e. **c**, Timecourses of brain activity during the delay period. Timecourses are shown for the HPC (*left*), aPFC (*right*), stayers (*top*), and explorers (*bottom*).

<u>The HPC encodes anticipatory experience in a novel environment.</u>

Before a foraging trial, participants received a delayed reward in an experience trial with a novel environment. We then asked whether anticipatory brain activity in the experience trials was associated with choice behavior in the foraging trials. We explored anticipatory brain activity in the experience trials by modeling AU and UFR dynamics (Fig. 4a; Supplementary Fig. 6a), as in the remain trials (Fig. 2d).

Whole-brain exploratory analysis revealed that the HPC and ventrolateral prefrontal cortex (vlPFC) showed a strong AU effect (Fig. 4b; Supplementary Fig. 6b/c). The vlPFC region was spatially segregated from the aPFC region that showed an AU effect in the foraging trial (23.7 mm apart).

Importantly, the stayers showed a stronger AU effect in the HPC, whereas in the explores this effect was stronger in the vlPFC (Fig. 4c; Supplementary Fig. 6d). It is interesting that in the foraging trials, a stronger AU effect in the HPC was also observed in the stayers, suggesting that the stay strategy may be guided by HPC-dependent encoding and reproduction of anticipation of a future reward[32-34].

Discussion

In the current study, humans foraged for real liquid rewards available after tens of seconds, and continuous decision was tracked by temporal dynamics of brain activity reflecting anticipation of a future reward. When the anticipatory activity in the aPFC diminished, they abandoned the current environment and explored a new one. The hippocampal activity was enhanced in humans remaining stationary, whereas those adopting exploratory strategy showed strong aPFC activity. As such, anticipatory dynamics in the fronto-hippocampal mechanisms underlie continuous stay-explore decision-making.

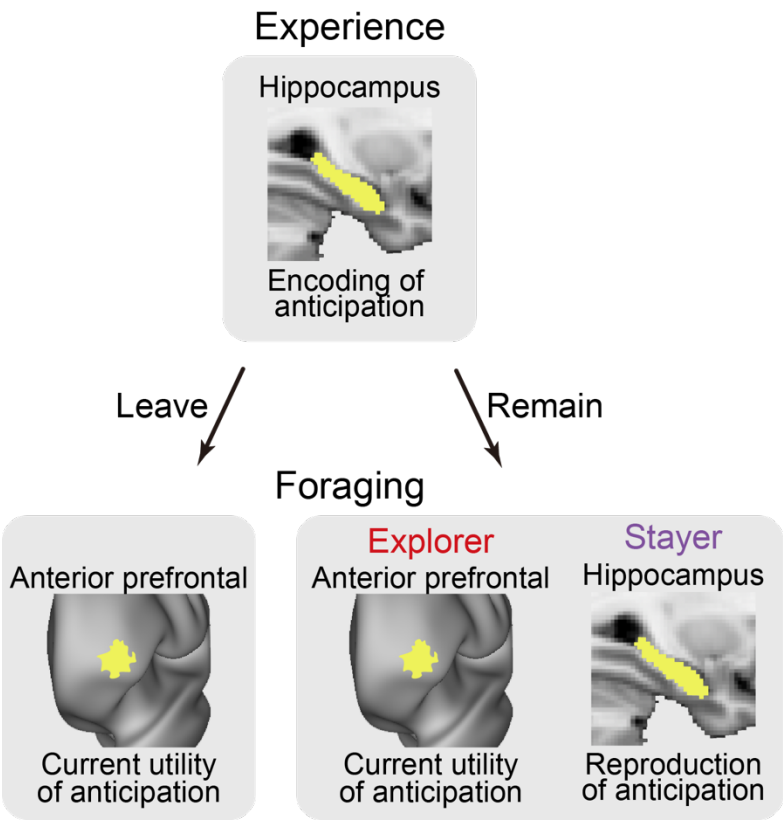


Figure 5. Schematic diagram of human prefrontal-hippocampal mechanisms involved in foraging for primary rewards.

In the literature on foraging, there are discrepancies between human and animal studies, despite the ethological generality and importance of foraging among a wide range of species. In non-human animal studies, behavioral agents directly foraged for primary rewards[5,10-12,14,15], whereas previous human studies used secondary rewards

mainly in hypothetical situations[6,35,36] (cf. entertaining videos as rewards in ref. 12). Due to these discrepancies, it was challenging to compare foraging and underlying biological mechanisms across species. The current study bridged the gap between human and animal studies, providing important consistency across species.

Our study highlights the distinctive roles of the aPFC and HPC in foraging (Fig. 5). The aPFC encodes choice behavior and is associated with exploratory strategy, but is not engaged in the experience trials, suggesting that the aPFC is sensitive to the current pleasure (utility) from moment to moment during foraging[7,19,21,37]. Thus, the aPFC role may reflect dopaminergic activity in the prefrontal cortex[38]. On the other hand, the HPC is involved in the experience trials and is associated with the stay strategy, but it is unrelated to leaving the current environment. The greater HPC effect in the stayers may reflect enhanced reproduction of the experience of reward anticipation[32,33,39-41], as suggested by prior non-human mammal studies demonstrating that the HPC encodes and reproduces temporal series of events[22-28].

The current finding that gradual decrease of aPFC anticipatory activity toward leaving the current environment (Fig. 2c) contrasts with previous rodent, primate, and human studies showing that increased neuronal activity in the medial frontal area triggers exploration of a new environment[5,15] and predicts when the decision to leave the environment is made[36,42-46]. In *C. elegans*, on the other hand, decreased activity of serotonergic neurons drives exploration[10], which better fit our finding in terms of exploration associated with activity reduction. It has also been demonstrated that stay-explore strategies during foraging of *C. elegans* are affected by genetic variations in adrenergic receptors[11] and serotonergic neuronal activity[10]. Our findings of individual differences in foraging strategies (Figs. 3a/b) and the underlying prefrontal-hippocampal mechanisms in humans may be reflected in molecular neurobiological mechanisms shared across species.

The current aPFC region is located in a polar region in the prefrontal cortex within Brodmann area 10, which is disproportionately developed in humans[47]. This region shows dynamic AU representation when a delayed reward is anticipated[19,21]. The region is also related to thinking about both future events[7,39,48,49] and exploration of a novel environment involving anticipatory evaluation of the current and novel environments[9]. Thus, our results highlight that a prefrontal region unique to humans represents a desirable future event while foraging for primary rewards. It is worth

noting that our results were obtained with a foraging model that has been used for non-human animals such as nematodes and primates.

It has been argued that given the complexity of the real world, foraging behavior is governed by several environmental factors[2]. Such factors include success and failure of prey acquisition[15], duration since previous prey acquisitions, costs of traveling between feeding sites[5], cost of engagement in the current site[6], and resources available in the environments[35]. These factors are well explained by the marginal value theorem and other formal models[17]. In the real world, predators, competitors, and nutrition can also be important[3]. In the current study, we aimed to examine the effect of anticipating future reward attainment, and then manipulated the amount of liquid rewards and the duration to the reward outcome. Consideration of exploration costs of humas is associated with the anterior cingulate cortex[6]. By contrast, we focused on anticipation of a future reward, and did not observe involvements in the anterior cingulate cortex.

The continuous decision required during the foraging trials involves a trade-off between staying in the current environment and beginning exploration to locate a novel environment. A similar trade-off is a choice between exploration and exploitation, where behavioral agents repeatedly make choices among options to maximize long-term reward attainment (bandit task). In such situations, the agent's valuation of options was updated based on their history of choice outcomes[16,50,51]. Interestingly, human studies using abstract rewards in hypothetical situations showed that exploration of an uncertain option involved a region in the anterior prefrontal cortex[16]. The current study suggests that the involvement of the aPFC reflects stronger anticipation for another choice.

While foraging, behavioral agents attempt to maximize long-term prey acquisition based on continuous consideration of the stay-explore trade-off. The maximization of future reward attainment has been examined in a decision-making situations involving a choice between a larger delayed reward and a smaller reward available sooner, formulated as intertemporal choice[8,14,19,21]. The preference in intertemporal choice reflects impulsivity and self-control in reward-seeking behavior[19]. Importantly, in intertemporal choice, the optimal choice to maximize long-term reward attainment, i.e., by choosing the larger delayed reward, is always obvious to the agents. In contrast, during foraging the optimal choice is not always obvious to agents because there is uncertainty about future prey acquisition in the surrounding environment. Thus,

the continuous stay-explore choice in foraging may not be compatible with the impulsivity-self-control spectrum in intertemporal choice, and choice preference may be associated with other trait constructs reflecting anticipation of future desirable events[8,20].

Author contribution

D.T., N.S., and K.J. designed research; R.S., D.T., T.Y., J.C., and K.J. performed research; R.S., D.T., S.S. and K.J. contributed analytic tools; R.S., D.T. and K.J. analyzed data; R.S. and K.J. wrote the first draft of the paper; R.S., S.S., N.S., J.C., and K.J. edited the paper; R.S. and K.J. wrote the paper.

Acknowledgments

This work was supported by JSPS Kakenhi (20K07727, 17K01989, and 26350986 to K.J.); NIPS Cooperative Study Program (20-639, 19-635, 18-633, 17-6229, and 21-544 to K.J.); ABiS (16A-073-M02 to K.J.); AMED (JP21dm0207086 to J.C.) We thank Drs. Kotaro Oka, Teppei Matsui, and Kentaro Miyamoto for their scientific comments on the manuscript. We thank Mimu Yabuta for illustration assistance.

Conflict of interest

The authors declare no conflict of interests.

Methods

Participants.

Human participants (N = 41; age range, 20-28 years; 20 females) were right handed and had no history of psychiatric or neurological disorders. Written informed consent was obtained from all participants. All experimental procedures were approved by the institutional review boards of Keio University and the National Institute of Physiological Sciences. Participants were instructed not to drink any liquid for 4 hours before the experiment and received 8,000 yen for participation. Five participants did not leave the current environment in any foraging trial, and these participants were excluded from the analyses examining trials where participants stopped waiting.

Reward.

The current study used commercially available drinks as a reward that could be immediately consumed. Before the experiment, participants were provided with a list of the following beverages: apple, orange, grape, grapefruit, lychee, pear, and mixed fruit juices; probiotic drinks; barley tea; and water. Each participant was then asked to choose the drink that would serve as their reward.

Apparatus.

E-prime programs (Psychology Software Tools) controlled the task as well as the delivery of liquid rewards via a syringe pump (SP210iw; World Precision Instruments). Liquids from two 60-ml plastic syringes mounted on the pump were merged into one tube and then delivered to the participant's mouth through a silicon tube. The flow rate of each syringe was set to 0.75 ml/s, and thus the reward flowed continuously at a rate of 1.5 ml/s. Participants were able to control the liquid flow. Reward delivery continued as long as they pressed a button on a box that they held in their right hand; delivery paused if they released the button, and resumed when they pressed the button again.

Behavioral procedures.

During functional MRI scanning, human participants performed a foraging task for real liquid rewards delayed by tens of seconds (Fig. 1b). Participants first experienced a delayed reward in a novel environment indicated by a picture presented on the center of the screen (experience trial), and then waited for another reward in the same

environment (foraging trial). Importantly, in the foraging trial, participants were unsure when they would receive the reward; however, they were able to stop waiting and move on to a novel environment whenever they preferred. They alternately performed the experience and foraging trials, and an environment was used only for one pair of the experience and foraging trials. Each environment was indicated by a unique picture.

Prior to each experience trial, participants were presented with a visual message to wait until delivery of a reward, and to press the button to start the trial. The visual message was presented until the participants pressed the button with their right thumb, at which point the visual message disappeared and a fixation cross was presented for 3 s. Then, a picture indicating an environment was presented, and the delay period started. During the delay period, the elapsed time from the start of the trial was indicated by a white horizontal bar that extended from the left to the right side of the screen every 250 ms. A full bar extending to the right end of the screen corresponded to 80 s, which was longer than the maximum delay duration in the foraging and experience trials (60 s). At the end of the delay, a visual message was presented indicating that the reward was ready, and participants consumed the liquid reward. After they consumed all of the liquid, a fixation cross was presented for 13 s.

Then, participants foraged for another reward in the same environment (foraging trial). Before starting of the foraging trials, they were presented with a visual message telling them to wait for another delayed reward, to stop waiting by pressing the button at any time, and to press the button to start the trial. When they pressed the button, the visual message disappeared and a fixation cross was presented for 3 s. Then, the picture indicating the environment in the experience trial was presented again, and participants started waiting. While they were waiting, the elapsed time was indicated by a white elongating bar as in the experience trials. Additionally, a white triangle was shown above the white bar, indicating the reward delivery time in the experience trial.

If participants waited until the delivery of a reward (remain trial), a visual message indicating that the reward was ready was presented, and participants consumed the liquid reward. If they stopped waiting before the reward delivery (leave trial), the environmental information (center picture, white bar, and triangle) disappeared, and a fixation cross was presented. After an inter-trial interval (ITI), the next experience trial started with a new environment. The duration from the onset of the foraging trial to the end of the ITI was 75-165 s, which varied by reward amount and delay duration (see

below), but was independent of participants' choices in the foraging trials. This meant that the ITI was longer in the leave trials, but no participants reported that they noticed a longer ITI when they stopped waiting. This ensured that participants did not consider the duration of the ITI when making decisions during the foraging trials.

Participants performed a total of 24 pairs of experience and foraging trials during fMRI scanning, and each scanning run consisted of four pairs of the trials. Supplementary Table 1 lists the delay durations of all trials. Specifically, there were six delay durations in the experience trials (10, 15, 20, 40, 45, and 50 s). The short-delay condition comprised of the 10-, 15-, and 20-s delays, and the long-delay condition consisted of the 40-, 45-, and 50-s delays. In the foraging trials, the delay duration was shorter or longer than that of the experience trials (Supplementary Table 1). The amount of the delayed reward was either 3 ml (small amount) or 12 ml (large amount). Thus the current experiment was based on a 2 x 2 factorial design (delay duration: short/long; reward amount: large/small). These conditions appeared pseudorandomly within one scanning run and across scanning runs.

Before fMRI scanning, participants received instructions for the task using a computer display. Participants were told that reward delivery times in foraging trials could be shorter or longer than those in experience trials (the latter indicated by a white triangle). In order to familiarize participants with the task, two pairs of experience and foraging trials were performed as practice trials in the MRI scanner [(delay duration in the experience trial, reward amount, delay duration in the foraging trial): (5 s, 6 ml, 15 s), (25 s, 6 ml, 15 s)]. The total liquid consumption in an entire experimental session ranged from 192 ml to 384 ml.

Each environment consisted of one of 24 unique pictures (six shapes x four colors). One color was used only once in each scanning run. Participants were told that the picture used in one trial pair was unrelated to that used in other trial pairs. Pictures used in practice trials were not used in scanning trial pairs.

Survival rate analysis and assessment of foraging strategies.

To assess participants' behavior in the foraging trials, we used survival analysis based on the Kaplan-Meier method[52]. This analysis examined how long participants remained in an environment by estimating the frequency with which they continued to wait until time t during the delay period of the foraging trials (survival rate). Specifically, the

rates were calculated based on observations of leaving the environment (event occurrence) and completion of the delay period of the foraging trial (censoring). The survival rate was formulated as

$$S(t) = \prod_{t(i) \leq t} \left(\frac{n-i}{n-i+1} \right)^{\delta(i)} \quad (1),$$

where $S(t)$ represents the estimation of the survival rate at the time t, n is the sample size, and $t_{(i)}$ denotes the time of the ith event occurrence. $\delta_{(i)}$ is an indicator function that takes a value of 0 if the ith observed event is a censoring observation and 1 if it is a non-censoring observation. The survival rate was calculated for each reward amount and delay duration of the experience trials across participants using *ecdf* implemented in MATLAB ver. 2017a.

The decision strategy in the foraging trial was quantified by the area under the curve (AuC) of the survival function for each trial condition across participants. A smaller AuC reflects the strategy to explore a new environment, whereas a greater AuC reflects the stay (non-exploratory) strategy in the current environment due to anticipation of a future reward. The AuC values were calculated for each delay and amount condition and compared across trial conditions.

The AuC values were also calculated for each participant across conditions to quantify individuals' choice strategies in the foraging trials. Individuals with greater AuCs waited longer (i.e., adopting a stay strategy) whereas those with smaller AuCs waited for less time (i.e., adopting an exploratory strategy). Given this characterization of the individuals' AuCs, we classified participants into three groups based on AuC values, and labeled the highest tertile as stayers and the lowest tertile as explorers.

<u>Sensitivity to environmental condition.</u>

To examine how environmental parameters (reward amount and delay duration in the experience trials) affected choice behaviors during foraging trials, we quantified individuals' environmental sensitivity. We first hypothesized that the duration that participants would wait in a foraging trial would be predicted by the reward amount and delay duration in the experience trials. This hypothesis was formulated by the following regression model for each participant:

$$T = \beta_a x_a + \beta_d x_d + \beta_c \quad (2),$$

where T is (waiting duration) / (time until reward delivery during foraging trials) (= 1, if participants waited until reward delivery); x_a and x_d are the reward amount and delay duration, respectively; β_a and β_d are the regression coefficients for the reward amount and delay duration, respectively; and β_c is a constant term.

Data for the estimated coefficients of β_a and β_d were collected from all participants and averaged across participants. The significance of the averaged β_a and β_d was tested by permutation tests. Specifically, within participants, the parameters of reward amount and delay duration were randomly shuffled and relabeled, and then the same regression analysis was performed and coefficients were collected and averaged across participants. Crucially, this random relabeling was performed within participants so that participants' individuality was preserved. This procedure was repeated 5000 times. Then, group-level averages from 5000 randomizations were collected, which provided for null distributions of β_a and β_d.

β_a and β_d were 0.0056 ± 0.0075 (mean ± SD) ($P < .001$) and -0.0072 ± 0.0059 ($P < .001$), respectively, suggesting that participants waited longer if they received a larger reward and/or if they received a reward after a shorter delay. Thus, a more positive β_a indicates that the wait duration was longer when a participant received a larger reward. Likewise, a more negative β_d indicates that the wait duration was longer when a participant received a reward after shorter delay.

Then, β_a, β_d, and β_c were estimated for each participant. The sensitivity to environmental parameters se was defined as

$$se = \sqrt{\beta_a^2 + \beta_d^2} \qquad (3).$$

To test whether the se value of individual participants was associated with their stay-explore strategy in the foraging trials, a correlation coefficient was calculated between the se values and the AuC values of the survival function across participants. Notably, AuC represents individuals' overall choice strategies across trials, whereas se represents trial-by-trial variability in choice behavior within individuals. Thus, se and AuC represent conceptually independent constructs. However, they were not quantitatively independent, because both AuC and se were calculated based on the behavioral data in the foraging trials. Due to the quantitative dependence of these values, they could be pseudo-correlated, which requires careful consideration when calculating a null distribution of the correlation. Because the pseudo-correlation is derived from the individuality of the behavioral data, when calculating a null

distribution, we preserved the individuality by permuting experimental conditions within individuals, similarly to the testing for the mean β_a and β_d values across participants as stated above. As expected, the null distribution shows a shift toward the negative direction (Supplementary Fig. 1). Then, significance of the correlation was tested based on the null distribution.

Imaging procedure.

fMRI scanning was conducted on a whole-body 3T MRI system (Siemens Verio, Germany). Functional images were acquired using multi-band accelerated gradient-echo echo-planar imaging [repetition time (TR) = 800 ms; echo time (TE) = 30 ms; flip angle (FA) = 45 degrees; slice thickness, 2 mm; in-plane resolution, 2 x 2 mm; multi-band factor (MBF) = 8; 80 slices]. Whole brain scanning with high temporal resolution allowed us to perform exploratory analyses of temporal dynamics during the delay period with sufficient scanning frames. Each run involved 635 volume acquisitions (508 s), and six functional runs were performed (total 3990 volumes). The initial 10 volumes of each run were excluded from imaging analysis to take into account the equilibrium of longitudinal magnetization. High-resolution anatomical images were acquired using an MP-RAGE T1-weighted sequence (TR = 2500 ms; TE = 4.32 ms; FA: 8 deg; 208 slices; slice thickness, 0.8 mm; in-plane resolution, 0.8 x 0.8 mm^2).

Imaging analysis procedures.

Preprocessing.

Functional images were preprocessed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/). All functional images were first temporally aligned across the brain volume, corrected for movement using correction for rigid-body rotation and translation correction, and then registered to the participant's anatomical images to correct for movement between the anatomical and function scans. Participants' anatomical images were transformed into a standardized MNI template. The functional images were then registered to the reference brain using the alignment parameters derived for the anatomical scans. The data were resampled into 2-mm isotropic voxels, and spatially smoothed with a 6-mm full-width at half-maximum Gaussian kernel.

In order to minimize motion-derived artifacts due to consumption of liquid rewards[19,21,37], functional images were further preprocessed by general linear model

(GLM) estimations with motion parameters and MRI signal time courses (cerebrospinal fluid, white matter, and whole brain), and their derivatives and quadratics as nuisance regressors[21,53,54] based on *fsl_regfilt* implemented in the FSL suite (http://www.fmrib.ox.ac.uk/; ver. 5.0.9). The residual of the nuisance GLM was used for standard GLM estimations to extract event-related brain activity as described below.

The mean magnitudes of absolute head motion were less than 0.23 mm for translation and 3.8 x 10^{-3} radian for rotation, which is consistent with our prior studies[19,21,37] (see ref. [37] for details). Critical image distortions or signal drop-outs were not observed, but robust brain activity was extracted in the primary motor and gustatory cortices during liquid consumptions, which assured us that as in our prior studies[19,21,37], motion-derived artifacts did not critically contaminate the current analyses.

General linear model.

Single level analysis.

A GLM approach was used to estimate trial event effects of brain activity. Parameter estimates were performed by *feat* implemented in the FSL suite. Events of particular interest were defined as periods during which participants waited for a future reward, i.e., the delay periods of the foraging and experience trials. The current analysis focused on the temporal dynamics of brain activity during these events[19,21].

Foraging trial: leave before reward delivery.

During the foraging trials, participants anticipated a future reward and continuously decided whether to stay in the current environment or to explore a new one. When the anticipation of a future reward diminished, they left the current environment. Thus, we hypothesized that this departure occurred when brain activity representing the anticipation of a future reward in the current environment was attenuated. To model temporal changes in anticipatory activity, we used the survival function of the foraging trials because continuous decision and anticipation are reflected in temporal changes in survival rates. This hypothesis was then tested by exploring brain regions showing dynamic brain activity associated with the survival functions.

Specifically, the anticipatory dynamics of brain activity were modeled using the survival functions calculated in behavioral analysis (Fig. 2a). For each participant, we estimated survival functions for each of the four trial conditions (reward amount:

large/small; delay duration: short/long; see Supplementary Table 1). Then, for each leave trial, the delay period was coded by the survival function from the onset of the foraging trial until the button press indicating that the participant was leaving the current environment. Alternative temporal dynamics were modeled by the inverse of the survival function (1 - survival function = death function), which is simultaneously coded in the GLM. Then, the survival and death models were convolved with a canonical hemodynamic response function (HRF) (Supplementary Fig. 2a).

By definition, the survival and death functions are anticorrelated, but when convolved with the HRF, they produced dissociable BOLD regression functions (Supplementary Figs. 2a/4b). Details regarding the orthogonality of the regressors and the control analyses for inspecting multicollinearity are described in the section on GLM estimations and control analyses below.

Foraging trial: remain until reward delivery.

We examined anticipatory dynamics of brain activity during the delay period of the foraging trial where participants remained in an environment until reward delivery. The survival function of the foraging trials is not suitable for modeling these dynamics because the survival function reflects when participants stopped waiting, whereas in the remain trials, they continued waiting until reward delivery.

One useful theorization of the anticipation of desirable events occurring in the future is the current utility of anticipation (anticipatory utility: AU) reflecting the pleasure of future anticipation[20]. In accordance with AU, our previous studies modeled brain activity dynamics while participants were awaiting a future reward[19,21]. Notably, in the current foraging trials, participants waiting for a reward were unsure about when this reward would be delivered. In our previous study[21], AU in these uncertain situations was modeled based on expectation of future reward delivery, which was updated following every experience of a delayed reward using a Bayesian learning inference approach[50]. In the current study, this framework was applied to modeling the anticipation of a future reward in the remain trials.

In the foraging trials, although participants were unsure when the reward would be delivered, it was possible to estimate when a reward would become available based on the previous reward experience in the same environment (i.e., the experience trial immediately before the foraging trial). The expectation should also depend on past

foraging trial experiences where participants waited until rewards were delivered or stopped waiting before delivery. Thus, the current modelling assumed that prior to the foraging trial, participants approximated the delivery time based on the immediately preceding experience trial, and the approximation was updated every time after participants performed a foraging trial.

The current analysis modeled the expectation of reward delivery and its updates using a Bayesian inference learning approach. We assumed that the expectation was expressed by a probability density distribution of the reward outcome, which was formulated based on a beta distribution as a function of the expected delivery time. A beta distribution was used because of its finite ends, which is consistent with participants' understanding of the maximum possible duration in the foraging trials. It should be noted that the amount of time that had passed in an individual foraging trial was shown by the length of the white bar, which did not extend beyond the right end of the screen (corresponding to 80 s; see Behavioral procedures).

The updates were implemented differently depending on whether participants waited until the reward delivery (remain) or left before it was delivered (leave) in the previous environment. For the remain trials, the update was based on 1) the reward delivery time in the immediately preceding experience trial, and 2) the disparity of reward delivery times between the experience and foraging trials in the previous environment (Fig. 2d and Supplementary Fig. 2b *top*). Specifically, Bayesian inference was performed as

$$P(\theta_r|D) \propto P(\theta_r)\, P(D|\theta_r) \qquad (4),$$

where D represents the delivery time of the foraging trial and θ_r is a parameter that determines the probability density defined by 1) the delivery time of the experience trial indicated by the white triangle and 2) the disparity between the triangle and the right end of the white bar when the reward was delivered.

For the leave trials, on the other hand, the update was based on 1) the reward delivery time in the immediately preceding experience trial, as in the remain trials, but also on 2) the time when participants stopped waiting (Supplementary Fig. 2b *bottom*). Specifically, Bayesian inference was performed as

$$P(\theta_r|D) \propto P(\theta_r)\, P(t > \tau|\theta_r) = P(\theta_r) \int_{\tau}^{T} P(t|\theta_r)\ dt \qquad (5),$$

where τ and T indicate the time participants left the current environment and the maximum possible delay (80 s), respectively. This formulation reflects the fact that the

reward would have been delivered between times τ and T because leaving at time τ indicates that the reward had not yet been delivered by this time.

Then, the probability density function (PDF) in the remain and leave trials was calculated as

$$PDF = P(D) = \int P(D|\theta_r)P(\theta_r)\, d\theta_r \qquad (6),$$

where $P(D|\theta_r)$ takes a beta distribution. The average of the initial distribution was 15 s (i.e., average of delay durations in the practice foraging trials).

To model anticipatory utility dynamics during the foraging trials, the PDF was first integrated from the start of the delay to time t, defined as cumulative probability (CP),

$$CP(t) = \int_0^t PDF(t)dt \qquad (7).$$

The idea of this integration is that as the delay period elapsed, participants' expectations of receiving a reward continued to increase, although they did not exactly know when the reward would become available. $CP(t)$ is similar to a hazard rate in that it reflects an estimation of the probability that the reward is likely to become available[31,55]. As time t approaches its maximum value (i.e., 80 s), $CP(t)$ reaches the upper limit,

$$\lim_{t \to T} CP(t) = 1 \qquad (8).$$

Then, AU was defined as

$$AU = 1 - CP(t) \qquad (0 < t < D)$$
$$AU = 0 \qquad (t > D) \qquad (9),$$

where D denotes the duration of the delay duration in the foraging trials. Immediately after the delay finished and the reward became available, AU was set to 0. AU shows temporal dynamics that peaks at the beginning of the foraging trial and monotonically decrease as the foraging trial period elapses (Fig. 2d *bottom* and Supplementary Fig. 2c), as in prior models[19,21].

$CP(t)$ is also associated with the value of a future reward that would eventually become available in the foraging trial, as in prior models[19,21,31,55,56]. The current study defined the future reward value as the upcoming future reward (UFR), formulated as $CP(t)$[31,55] (Supplementary Fig. 2c),

$$UFR(t) = CP(t) \qquad (0 < t < D)$$
$$UFR(t) = 0 \qquad (t > D) \qquad (10).$$

Like the AU value, the value of UFR was set to 0 immediately after the delay finished and the reward became available to drink.

The AU and UFR models were calculated for each remain trial for each participant, and then convolved with a canonical HRF (Supplementary Fig. 2c *right*).

Like the survival and death functions for the leave trials, AU and UFR are anticorrelated by definition in the economic models, but when convolved with the HRF, they produced dissociable BOLD regression functions (Supplementary Fig. 2d). Details about the orthogonality of the regressors and the control analyses for inspecting multicollinearity are described in the section on GLM estimations and control analyses below.

Experience trial.

For the experience trials, brain activity dynamics of anticipating a future reward were modeled similarly to the remain trials based on Bayesian inference learning. We assumed that participants 1) were unsure about when the reward would be delivered, 2) approximated the delivery time based on the previous experience trials, and 3) updated this approximation every time after participants performed an experience trial. Like the foraging remain trials, the approximation was formulated by a PDF of the reward outcome using a beta distribution as a function of the expected delivery time (for details, see the "Foraging trial: remain until reward delivery" section above). Unlike the remain trials, in every experience trial the participants were unable to stop waiting and had to continue to wait until reward delivery, and therefore the update was only dependent on the reward delivery time in previous experience trials. Thus, Bayesian inference was performed as

$$P(\theta_e|D) \propto P(\theta_e)\, P(D|\theta_e) \qquad (11),$$

where D represents the delivery time of the experience trial and θ_e is a parameter that determines probability density. Then, PDF representing the approximation of the reward outcome was calculated, as in eq. (6). The average of the initial PDF was 15 s (i.e., average of delay durations in the practice experience trials). The CP, AU, and UFR were calculated as eqs. (7)-(10). The AU and UFR models were calculated for each experience trial for each participant, and then convolved with a canonical HRF (Fig. 4a and Supplementary Fig. 6a).

As with the remain trials, AU and UFR are anticorrelated by definition in the economic models, but when convolved with the HRF, they produced dissociable BOLD regression functions (Supplementary Fig. 2d). Details about the orthogonality of

the regressors and the control analyses for inspecting multicollinearity are described in the section on GLM estimations and control analyses below.

GLM estimations and control analyses.

The regressors of the survival and death models for the leave trials, and also the AU and UFR models for the remain and experience trials, were simultaneously coded in the GLM model for each participant. Other events consisted of the presentation of visual messages notifying the participant of the start of the trial start and of reward delivery, and the consumption of liquid rewards, but were not of interest to the current study. These nuisance events were coded simultaneously and separately, and convolved with a canonical HRF. Then, parameters were estimated for each voxel across the whole brain.

As stated above, the survival and death models were anticorrelated, and likewise, AU and UFR models were anticorrelated in the remain and experience trials. However, when convolved with the HRF, they produced dissociable BOLD regression functions (Supplementary Figs. 2a/d and 6a). Specifically, the correlations of the regressors were 0.58 ± 0.18 between the survival and death models for the leave trials, 0.31 ± 0.06 between the AU and UFR models for the remain trials, and 0.08 ± 0.11 between the AU and UFR models for the experience trials (mean ± SD), allowing sufficient dissociation[57].

Additionally, we performed separate control GLM analyses, where only one of the two anticorrelated models was coded in GLM. In those control analyses, the effects of the temporal dynamic models were reasonably reproduced, confirming that multicollinearity of these regressors was not an issue, consistent with our previous studies[19,21].

In the leave and remain trials of the foraging trials, the survival and AU dynamics to model temporal changes in brain activity show similar temporal characteristics: they peak at the beginning of the foraging trials and monotonically decrease as the foraging trial period elapses (Figs. 2a/d and Supplementary Fig. 2). Similarly, the death and UFR dynamics show similar temporal characteristics that are the inverse of the survival and AU models. We then asked which model better fit the empirical data. To address this issue, we performed supplementary GLM analyses, where the relationships between trials and GLM models were interchanged.

Specifically, anticipatory dynamics during the remain and leave trials were modeled by the survival and AU models, respectively. Then, parameters were estimated based on a standard GLM approach.

Group-level analysis.

Maps of parameter estimations for delay period effects in the experience and foraging trials were collected from all participants and subjected to group-mean one-sample t-tests based on permutation methods (5000 permutations) implemented by *randomise* in the FSL suite. Voxel clusters were identified using a voxel-wise uncorrected threshold of $P < .01$, and the voxel clusters were tested for significance with a threshold of $P < .05$ corrected by the family-wise error rate. This non-parametric permutation procedure was validated to appropriately control the false-positive rate[58]. The peaks of significant clusters were then identified and listed in tables. If multiple peaks were identified within 12 mm, the most significant peak was retained.

To examine the relationships between choice strategy and anticipatory brain activity, correlations between the AuCs of survival functions and AU effects in the foraging (remain) and experience trials were explored across the brain. For the remain trials of the foraging trials, statistical significance was tested similarly to the group-mean tests across the whole brain. For the HPC, given the group-mean effect of AU in a middle part of the HPC (Fig. 2e), the exploration of the correlation was restricted within an anatomically defined middle HPC region using Harvard-Oxford cortical and subcortical structural atlases.

For the experience trials, as we asked whether the aPFC and HPC regions showing significant AU effects in the remain trials also showed the AU effect in the experience trials, the exploration of correlations was restricted within regions of interest (ROIs). Specifically, the HPC ROI was defined anatomically and the aPFC ROI was defined based on the group-mean effect of anticipatory activity in the foraging trials (Supplementary Fig. 6d).

To evaluate the fit of dynamic models to the empirical data, parameter estimates of aPFC and HPC ROIs were extracted (Supplementary Fig. 5). The aPFC ROIs were defined based on the coordinates in the aPFC that showed AU effects identified in our previous studies[19,21]. The exact coordinates were (-30, 58, -8), and

ROIs were created as spheres with radii of 6 mm centered on the peaks. The HPC ROIs were defined anatomically.

Mixed-effects GLM.

The aPFC showed the AU effect in the remain trials (Fig. 1e), and decreased aPFC activity preceded leaving the environment (Figs. 1b/c). We then hypothesized that the aPFC anticipatory activity would predict continuous choice behavior during the foraging trials. To test this hypothesis, we performed a trial-based regression analysis based on a hierarchical mixed-effects GLM[59]. The model included two levels, one within-subject trial-by-trial effects, and the other between-subject effects. The lower-level within-subject effects was modeled as:

$$X_{cho} = \beta_{aPFC} X_{aPFC} + \beta_{constC} + \varepsilon_{cho} \quad (12),$$

where X_{cho} indicates (waiting duration) / (time until reward delivery during foraging trials) ($X_{cho} = 1$, if participants waited until reward delivery), and X_{aPFC} indicates the MRI signal during the foraging trials in the aPFC. β-values represent regression coefficients, and ε_{cho} is an error term. The higher-level between-subject effect was modeled as:

$$\beta_{aPFC} = \gamma_{aPFC} + \varepsilon_{aPFC} \quad (13),$$

where γ-values indicate regression coefficients and ε_{aPFC} is an error term. In this model, participants were treated as a random effect.

The aPFC ROIs were defined independently of the tested effects in order to avoid circular analysis. Specifically, the ROIs were created as spheres with a radii of 6 mm centered on the peaks based on the coordinates in our prior studies (-30, 58, -8)[19,21].

In the ROIs, MRI signal timecourses relative to the fixation baseline were then extracted during the delay period of the foraging trials. Because the anticipatory models (i.e., survival function and AU) show monotonic signal decreases toward the end of the delay period and was almost absent in the late delay period (Fig. 2a/d), the last 25% of the scanning frames of each foraging trial was discarded. To eliminate a bleeding-over effect of the BOLD signal from the pre-trial period, the first four scanning frames were also discarded. Then, remaining scanning frames were averaged over the time frames for each trial. These trial-by-trial signal values were submitted to the model and all parameters were simultaneously estimated using the *lmer* procedure in R (http://www.r-project.org/). In supplementary analyses, we confirmed that

fundamental results were retained when adjusting the data discard period from 0% to 50%.

A separate analysis was performed by replacing the aPFC activity with HPC activity, where HPC ROI was defined anatomically.

Functional connectivity analysis.

In the remain trials, anticipatory activity was observed in the aPFC and HPC (Fig. 2e), and attenuation of aPFC activity was associated with leaving the environment (Fig. 2c). Then, we hypothesized that the aPFC and HPC would be functionally coordinated during the remain trials where leaving did not occurred . To test this hypothesis, trial-by-trial based regression analysis was performed. Specifically, the anticipatory activity in the HPC was predicted by the anticipatory activity in the aPFC; this is formulated as

$$X_{HPC} = \beta_{aPFC} X_{aPFC} + \beta_{constH} + \varepsilon_{HPC} \quad (14),$$

where X_{aPFC} and X_{HPC} indicate MRI signals during the foraging trials in the aPFC and HPC, respectively. The HPC ROI was defined anatomically. MRI signals in the aPFC and HPC ROIs were calculated similarly to those used in the mixed-effects analysis above. Then, regression coefficients were estimated by the *lmer* procedure in R.

Data and code availability.

The datasets and code supporting the current study are available from the corresponding author (Koji Jimura, jimura@bio.keio.ac.jp) on reasonable request.

References

1 Rushworth, M. F., Noonan, M. P., Boorman, E. D., Walton, M. E. & Behrens, T. E. Frontal cortex and reward-guided learning and decision-making. *Neuron* **70**, 1054-1069, doi:10.1016/j.neuron.2011.05.014 (2011).

2 Pearson, J. M., Watson, K. K. & Platt, M. L. Decision making: the neuroethological turn. *Neuron* **82**, 950-965, doi:10.1016/j.neuron.2014.04.037 (2014).

3 Mobbs, D., Trimmer, P. C., Blumstein, D. T. & Dayan, P. Foraging for foundations in decision neuroscience: insights from ethology. *Nat Rev Neurosci* **19**, 419-427, doi:10.1038/s41583-018-0010-7 (2018).

4 Wikenheiser, A. M., Stephens, D. W. & Redish, A. D. Subjective costs drive overly patient foraging strategies in rats on an intertemporal foraging task. *Proc Natl Acad Sci U S A* **110**, 8308-8313, doi:10.1073/pnas.1220738110 (2013).

5 Hayden, B. Y., Pearson, J. M. & Platt, M. L. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat Neurosci* **14**, 933-939, doi:10.1038/nn.2856 (2011).

6 Kolling, N., Behrens, T. E., Mars, R. B. & Rushworth, M. F. Neural mechanisms of foraging. *Science* **336**, 95-98, doi:10.1126/science.1216930 (2012).

7 Atance, C. M. & O'Neill, D. K. Episodic future thinking. *Trends Cogn Sci* **5**, 533-539, doi:10.1016/s1364-6613(00)01804-0 (2001).

8 Berns, G. S., Laibson, D. & Loewenstein, G. Intertemporal choice--toward an integrative framework. *Trends Cogn Sci* **11**, 482-488, doi:10.1016/j.tics.2007.08.011 (2007).

9 Iigaya, K. *et al.* The value of what's to come: Neural mechanisms coupling prediction error and the utility of anticipation. *Sci Adv* **6**, eaba3828, doi:10.1126/sciadv.aba3828 (2020).

10 Rhoades, J. L. *et al.* ASICs Mediate Food Responses in an Enteric Serotonergic Neuron that Controls Foraging Behaviors. *Cell* **176**, 85-97.e14, doi:10.1016/j.cell.2018.11.023 (2019).

11 Bendesky, A., Tsunozaki, M., Rockman, M. V., Kruglyak, L. & Bargmann, C. I. Catecholamine receptor polymorphisms affect decision-making in C. elegans. *Nature* **472**, 313-318, doi:10.1038/nature09821 (2011).

12 Sweis, B. M. *et al.* Sensitivity to "sunk costs" in mice, rats, and humans. *Science* **361**, 178-181, doi:10.1126/science.aar8644 (2018).

13 Jimura, K. *et al.* Domain independence and stability in young and older adults' discounting of delayed rewards. *Behav Processes* **87**, 253-259, doi:10.1016/j.beproc.2011.04.006 (2011).

14 Carter, E. C. & Redish, A. D. Rats value time differently on equivalent foraging and delay-discounting tasks. *J Exp Psychol Gen* **145**, 1093-1101, doi:10.1037/xge0000196 (2016).

15 Tervo, D. G. R. *et al.* The anterior cingulate cortex directs exploration of alternative strategies. *Neuron* **109**, 1876-1887.e1876, doi:10.1016/j.neuron.2021.03.028 (2021).

16 Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876-879, doi:10.1038/nature04766 (2006).

17 Charnov, E. L. Optimal foraging, the marginal value theorem. *Theor Popul Biol* **9**, 129-136, doi:10.1016/0040-5809(76)90040-x (1976).

18 Stephens, D. & Krebs, J. *Foraging Theory.*, (Princeton University Press, 1986).

19 Jimura, K., Chushak, M. S. & Braver, T. S. Impulsivity and self-control during intertemporal decision making linked to the neural dynamics of reward value representation. *J Neurosci* **33**, 344-357, doi:10.1523/JNEUROSCI.0919-12.2013 (2013).

20 Loewenstein, G. Anticipation and the Valuation of Delayed Consumption. *Oxford Journals* **97**, 666-684 (1987).

21 Tanaka, D. *et al.* Self-Controlled Choice Arises from Dynamic Prefrontal Signals That Enable Future Anticipation. *J Neurosci* **40**, 9736-9750, doi:10.1523/JNEUROSCI.1702-20.2020 (2020).

22 Shikano, Y., Ikegaya, Y. & Sasaki, T. Minute-encoding neurons in hippocampal-striatal circuits. *Curr Biol* **31**, 1438-1449.e1436, doi:10.1016/j.cub.2021.01.032 (2021).

23 Pastalkova, E., Itskov, V., Amarasingham, A. & Buzsáki, G. Internally generated cell assembly sequences in the rat hippocampus. *Science* **321**, 1322-1327, doi:10.1126/science.1159775 (2008).

24 MacDonald, C. J., Carrow, S., Place, R. & Eichenbaum, H. Distinct hippocampal time cell sequences represent odor memories in immobilized rats. *J Neurosci* **33**, 14607-14616, doi:10.1523/JNEUROSCI.1537-13.2013 (2013).

25 Thakral, P. P., Madore, K. P., Kalinowski, S. E. & Schacter, D. L. Modulation of hippocampal brain networks produces changes in episodic simulation and divergent thinking. *Proc Natl Acad Sci U S A* **117**, 12729-12740, doi:10.1073/pnas.2003535117 (2020).

26 MacDonald, C. J., Lepage, K. Q., Eden, U. T. & Eichenbaum, H. Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron* **71**, 737-749, doi:10.1016/j.neuron.2011.07.012 (2011).

27 Kraus, B. J., Robinson, R. J., White, J. A., Eichenbaum, H. & Hasselmo, M. E. Hippocampal "time cells": time versus path integration. *Neuron* **78**, 1090-1101, doi:10.1016/j.neuron.2013.04.015 (2013).

28 Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y. & Engel, S. A. Remembering episodes: a selective role for the hippocampus during retrieval. *Nat Neurosci* **3**, 1149-1152, doi:10.1038/80671 (2000).

29 Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* **9**, 545-556, doi:10.1038/nrn2357 (2008).

30 Green, L. & Myerson, J. A discounting framework for choice with delayed and probabilistic rewards. *Psychol Bull* **130**, 769-792, doi:10.1037/0033-2909.130.5.769 (2004).

31 McGuire, J. T. & Kable, J. W. Medial prefrontal cortical activity reflects dynamic re-evaluation during voluntary persistence. *Nat Neurosci* **18**, 760-766, doi:10.1038/nn.3994 (2015).

32 de Hoz, L. & Wood, E. R. Dissociating the past from the present in the activity of place cells. *Hippocampus* **16**, 704-715, doi:10.1002/hipo.20207 (2006).

33 Carr, M. F., Jadhav, S. P. & Frank, L. M. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nat Neurosci* **14**, 147-153, doi:10.1038/nn.2732 (2011).

34 Bulley, A. & Schacter, D. L. Deliberating trade-offs with the future. *Nat Hum Behav* **4**, 238-247, doi:10.1038/s41562-020-0834-9 (2020).

35 Constantino, S. M. & Daw, N. D. Learning the opportunity cost of time in a patch-foraging task. *Cogn Affect Behav Neurosci* **15**, 837-853, doi:10.3758/s13415-015-0350-y (2015).

36 Shenhav, A., Straccia, M. A., Cohen, J. D. & Botvinick, M. M. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nat Neurosci* **17**, 1249-1254, doi:10.1038/nn.3771 (2014).

37 Misonou, A. & Jimura, K. Prefrontal-striatal mechanisms of behavioral impulsivity during consumption of delayed real liquid rewards. *Frontiers in Behavioral Neuroscience* (2021).

38 Schultz, W., Stauffer, W. R. & Lak, A. The phasic dopamine signal maturing: from reward via behavioural activation to formal economic utility. *Curr Opin Neurobiol* **43**, 139-148, doi:10.1016/j.conb.2017.03.013 (2017).

39 Schacter, D. L., Addis, D. R. & Buckner, R. L. Remembering the past to imagine the future: the prospective brain. *Nat Rev Neurosci* **8**, 657-661, doi:10.1038/nrn2213 (2007).

40 Eichenbaum, H. Time cells in the hippocampus: a new dimension for mapping memories. *Nat Rev Neurosci* **15**, 732-744, doi:10.1038/nrn3827 (2014).

41 Eichenbaum, H. Prefrontal-hippocampal interactions in episodic memory. *Nat Rev Neurosci* **18**, 547-558, doi:10.1038/nrn.2017.74 (2017).

42 Murakami, M., Shteingart, H., Loewenstein, Y. & Mainen, Z. F. Distinct Sources of Deterministic and Stochastic Components of Action Timing Decisions in Rodent Frontal Cortex. *Neuron* **94**, 908-919.e907, doi:10.1016/j.neuron.2017.04.040 (2017).

43 Russo, E. *et al.* Coordinated Prefrontal State Transition Leads Extinction of Reward-Seeking Behaviors. *J Neurosci* **41**, 2406-2419, doi:10.1523/JNEUROSCI.2588-20.2021 (2021).

44 Lak, A. *et al.* Orbitofrontal cortex is required for optimal waiting based on decision confidence. *Neuron* **84**, 190-201, doi:10.1016/j.neuron.2014.08.039 (2014).

45 Ballesta, S., Shi, W., Conen, K. E. & Padoa-Schioppa, C. Values encoded in orbitofrontal cortex are causally related to economic choices. *Nature* **588**, 450-453, doi:10.1038/s41586-020-2880-x (2020).

46 Soutschek, A., Moisa, M., Ruff, C. C. & Tobler, P. N. Frontopolar theta oscillations link metacognition with prospective decision making. *Nat Commun* **12**, 3943, doi:10.1038/s41467-021-24197-3 (2021).

47 Petrides, M. Lateral prefrontal cortex: architectonic and functional organization. *Philos Trans R Soc Lond B Biol Sci* **360**, 781-795, doi:10.1098/rstb.2005.1631 (2005).

48 Peters, J. & Büchel, C. Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-mediotemporal interactions. *Neuron* **66**, 138-148, doi:10.1016/j.neuron.2010.03.026 (2010).

49 Benoit, R. G., Gilbert, S. J. & Burgess, P. W. A neural mechanism mediating the impact of episodic prospection on farsighted decisions. *J Neurosci* **31**, 6771-6779, doi:10.1523/JNEUROSCI.6559-10.2011 (2011).

50 Behrens, T. E., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. Learning the value of information in an uncertain world. *Nat Neurosci* **10**, 1214-1221, doi:10.1038/nn1954 (2007).

51 Racey, D., Young, M. E., Garlick, D., Pham, J. N. & Blaisdell, A. P. Pigeon and human performance in a multi-armed bandit task in response to changes in variable interval schedules. *Learn Behav* **39**, 245-258, doi:10.3758/s13420-011-0025-7 (2011).

52 Kaplan, E. L. & Meier, P. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association* **53**, 457-481 (1958).

53 Ciric, R. *et al.* Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *Neuroimage* **154**, 174-187, doi:10.1016/j.neuroimage.2017.03.020 (2017).

54 Keerativittayayut, R., Aoki, R., Sarabi, M. T., Jimura, K. & Nakahara, K. Large-scale network integration in the human brain tracks temporal fluctuations in memory encoding performance. *Elife* **7**, doi:10.7554/eLife.32696 (2018).

55 Janssen, P. & Shadlen, M. N. A representation of the hazard rate of elapsed time in macaque area LIP. *Nat Neurosci* **8**, 234-241, doi:10.1038/nn1386 (2005).

56 Berns, G. S. *et al.* Neurobiological substrates of dread. *Science* **312**, 754-758, doi:10.1126/science.1123721 (2006).

57 Otten, L. J., Henson, R. N. & Rugg, M. D. State-related and item-related neural correlates of successful memory encoding. *Nat Neurosci* **5**, 1339-1344, doi:10.1038/nn967 (2002).

58 Eklund, A., Nichols, T. E. & Knutsson, H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci U S A* **113**, 7900-7905, doi:10.1073/pnas.1602413113 (2016).

59 Raudenbush, S. W. & Bryk, A. S. *Hierarchical linear models*. 2 edn, (2002).

Supplementary information

Title: Human foraging for primary rewards is guided by fronto-hippocampal dynamics of anticipation

Authors: Reiko Shintaki,[1,*] Daiki Tanaka,[1,*] Shinsuke Suzuki,[2] Takaaki Yoshimoto,[3] Norihiro Sadato,[3] Junichi Chikazoe,[3,4] Koji Jimura[1,5]

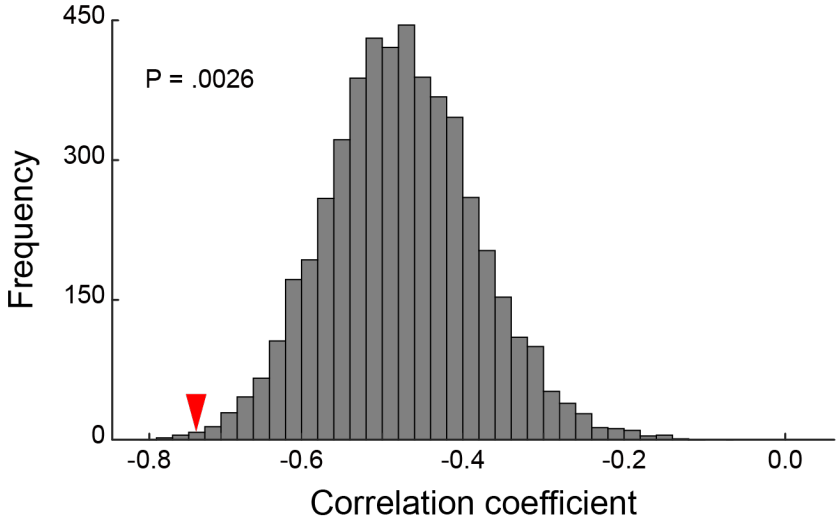
Affiliations:
[1] Department of Biosciences and Informatics, Keio University, Yokohama, Japan
[2] Brain, Mind & Markets Laboratory, Department of Finance, Faculty of Business and Economics, The University of Melbourne, Victoria, Australia
[3] Division of Cerebral Integration, National Institute of Physiological Sciences, Okazaki, Japan
[4] Araya, Inc., Tokyo, Japan
[5] Research Center for Brain Communication, Kochi University of Technology, Kami, Japan

*: These authors contributed equally to the study.

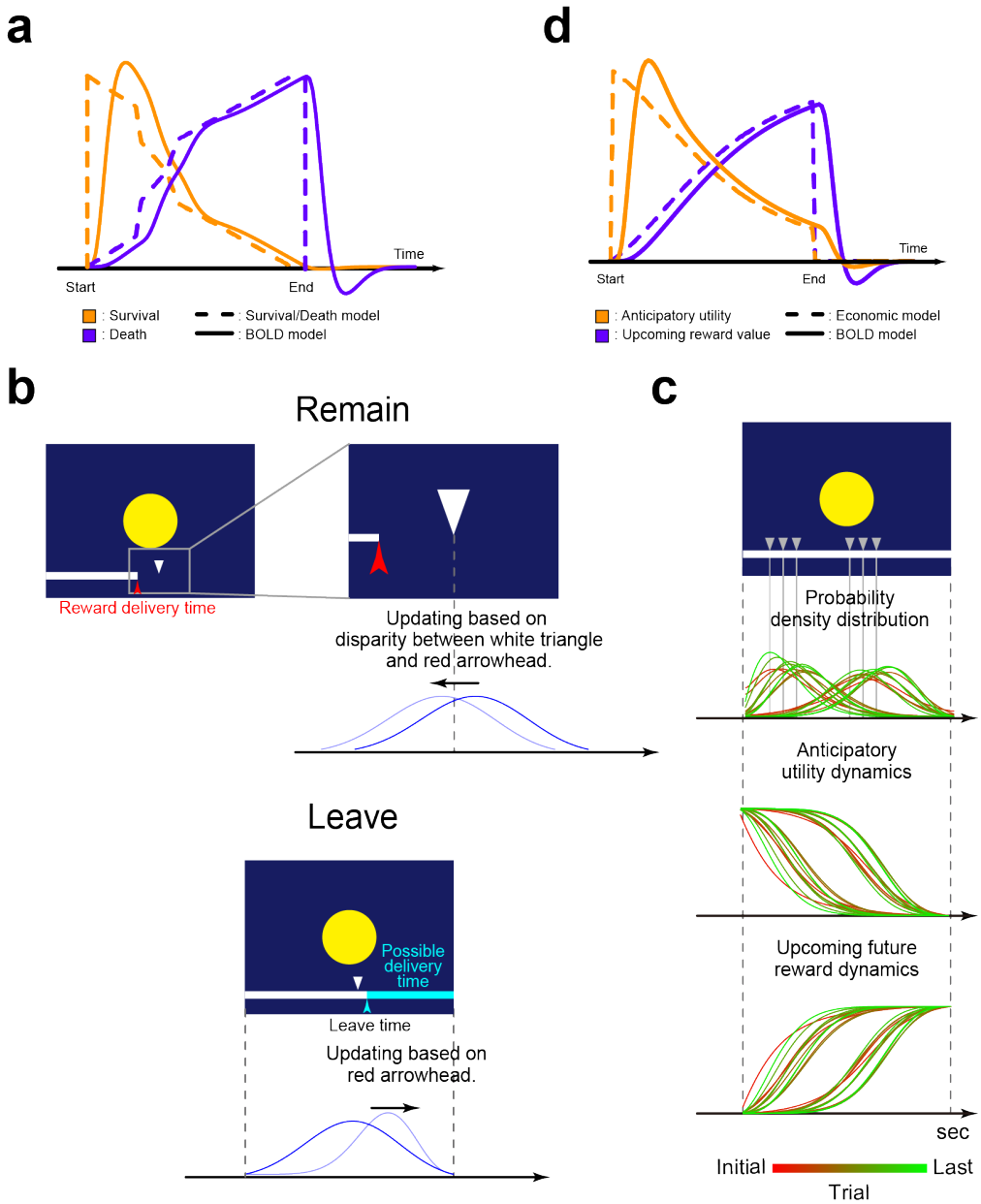
The authors declare no conflict of interest.

Correspondence should be addressed to:
Koji Jimura, Ph.D.
Department of Biosciences and Informatics,
Keio University,
3-14-1 Hiyoshi Kohoku-ku,
Yokohama, 223-0061, Japan
Email: jimura@bio.keio.ac.jp
Phone: +81-45-566-1625

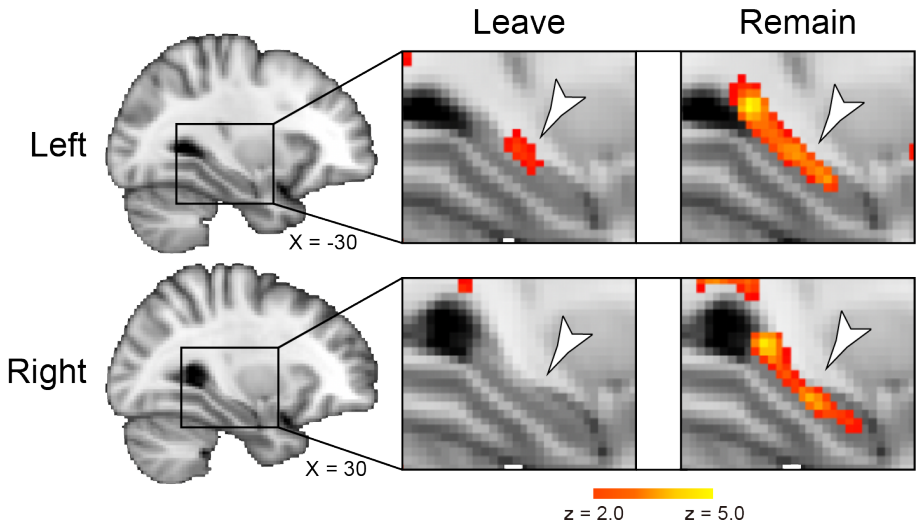
Supplementary figures



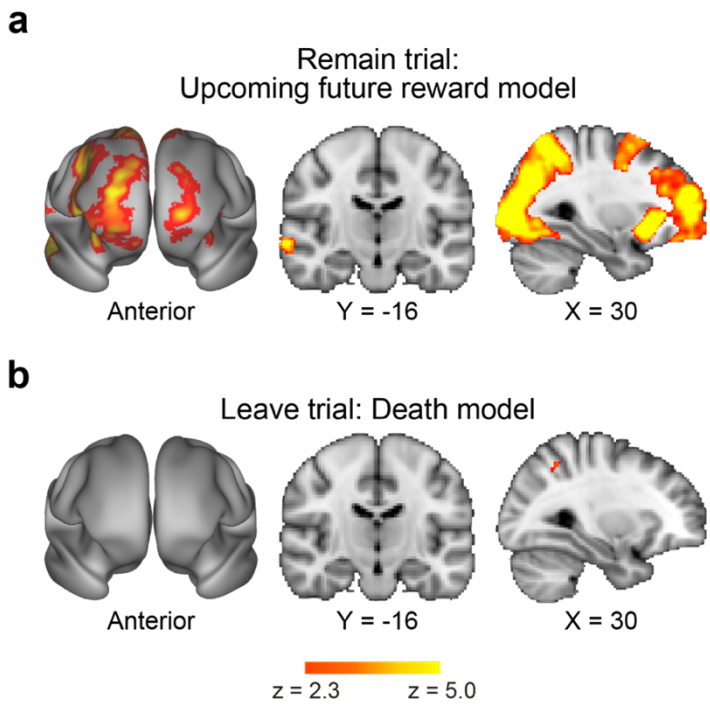
Supplementary figure 1. A null distribution of correlation between AuCs of survival functions and sensitivity to the environmental parameters across participants. The distribution was estimated by randomly shuffling reward amount and delay duration parameters and re-labeling within individual participants. Thus, individuals' choices were preserved when the distribution was estimated. Horizontal and vertical axes indicate the correlation coefficient, and the frequency of the permuted data, respectively. The red triangle shows the observed correlation (Fig. 1).



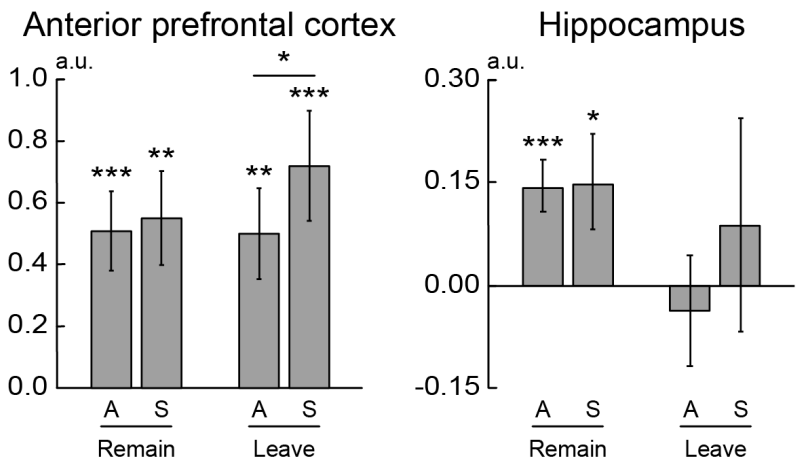
Supplementary figure 2. Dynamic activity models for the foraging trials. **a**, Survival and death models for the leave trials were convolved with the canonical HRF. **b**, For the remain trials (*top*), the expectation of a reward outcome was updated based on 1) the reward delivery time in the experience trial (white triangle) in the environment (yellow circle) (*top left*), and 2) the disparity of the reward delivery times between the experience trial (white triangle) and the foraging trial in the environment (red arrowhead) (*top right*). For the leave trials (*bottom*), the updates were based on 1) the reward delivery time in the experience trial in the environment (*bottom left*), and 2) the leave time (cyan arrowhead) that determined possible reward delivery time (i.e., somewhere between leave onset and the right end of the screen, as indicated by a horizontal cyan bar) (*bottom right*). **c**, Expectation of reward delivery and dynamic brain activity models. Probability density distribution for expectation of reward delivery was updated throughout the experience and foraging trials (*top*). The colors of the lines indicate trial experiences as shown in the color bar at the bottom. Gray triangles indicate delivery times in the experience trials. AU dynamics as an inverse function of cumulative probability (*middle*). Upcoming reward value dynamics as a function of cumulative probability (*bottom*). **d**, AU and UFR models for the remain trials were convolved with the canonical HRF.



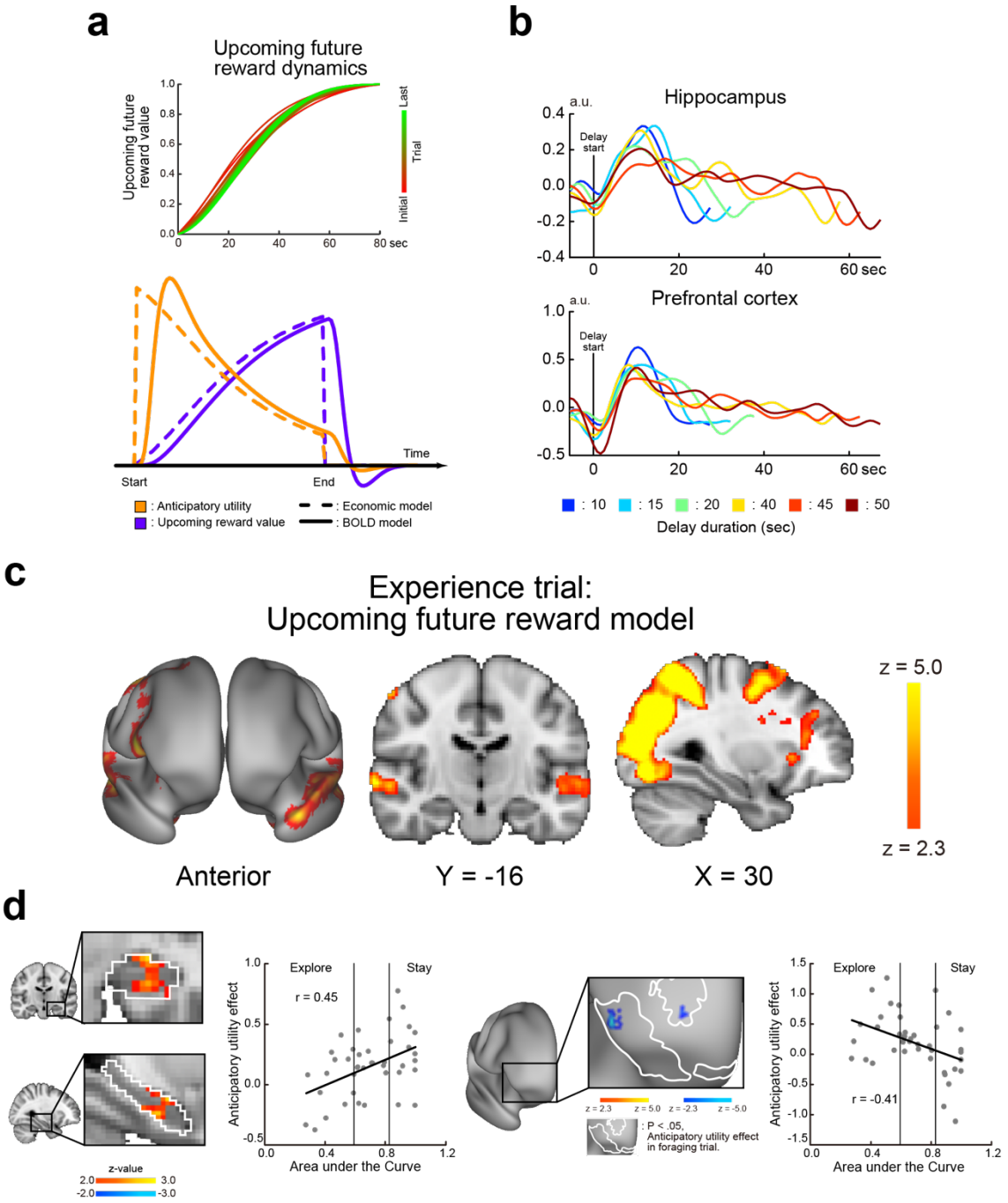
Supplementary figure 3. Anticipatory brain activity in the HPC (*left*) during the leave trials (*middle*) and remain trials (*right*). Statistical maps are thresholded using a voxel-wise statistical threshold of P < .05 (uncorrected) for display purposes. Participants who did not leave the current environment in any foraging trial (N = 5) were excluded in the analyses. The white arrowheads indicate anatomical locations of the HPC. *Top*: left HPC; *bottom*: right HPC. Formats are similar to those in Fig. 3a.



Supplementary figure 4. a, Brain regions showing UFR dynamics during the delay period of the remain trials. Anterior view (*left*); coronal section (*middle*); sagittal section (*right*). Levels of sections are shown below. **b**, Brain regions showing death model dynamics during the delay period of the leave trials. The formats are similar to the those in Fig. 2b.



Supplementary figure 5. To evaluate the fit of the dynamic models for the remain and leave trials to empirical data, the relationships between trials and models (AU and survival) were interchanged, and parameters were estimated. Then, estimated parameters were extracted within ROIs in the aPFC and HPC. The aPFC ROI was defined by previous studies, and the HPC ROI was defined anatomically. The horizontal axis indicates trials (remain and leave) and models (A: AU; S: survival), and the vertical axis indicates parameter estimates. Error bars indicate standard errors of the mean across participants. *: P < .05; **: P < .01; ***: P < .001.



Supplementary figure 6. Brain activity dynamics in the experience trials. **a**, Dynamic models of UFR during the delay period of the experience trials (*top*). The colors of the lines indicate trial experiences as shown in the color bar on the right. AU and UFR models were convolved with the canonical HRF (*bottom*). **b**, The timecourses of activity during the delay period of the experience trials in the HPC (*top*) and prefrontal cortex (*bottom*). **c**, Brain regions showing UFR dynamics during the delay period of the experience trials. The formats are similar to the those in Supplementary Fig. 3a. **d**, Statistical maps of correlation between the AuC and the AU effect in the HPC (*left*) and prefrontal cortex (*right*). White closed lines indicate anatomical borders of the HPC and the prefrontal regions show a significant AU effect in foraging trials in which effects in the remain and leave trials underwent weighted averaging. Formats are similar to those in Fig. 3a.

Supplementary tables

Supplementary table 1. Delay duration in the experience and foraging trials (s).

Experience		Foraging	
Short	10	10	20
	15	10	30
	20	15	40
Long	40	30	55
	45	35	50
	50	45	60

Supplementary table 2. Brain regions showing significant effects of survival and death models during delay period of the leave foraging trials. Coordinates are listed in MNI space. BA indicate Brodmann areas and is approximate.

Effect	Area	X	Y	Z	t-value	BA
Survival effect	Frontal	-36	62	-10	5.1	10
		-22	46	-12	4.8	11
		-26	56	6	4.4	10
		-38	6	62	3.9	6
		-22	2	60	3.8	6
		-28	-4	48	3.7	6
		-18	22	66	3.6	6
		-24	6	72	3.5	6
		-18	58	-12	3.5	11
	Parietal	-18	-72	58	7.8	7
		-4	-72	58	6.7	7
		-40	-58	44	5.9	39
		-26	-74	32	5.6	7
		10	-64	64	5.4	7
		24	-56	54	5.0	7
		-6	-58	50	4.3	7
		-36	-66	-40	4.2	39
		-30	-50	52	4.1	7
		-30	-44	40	3.9	7
		-52	-58	40	3.7	39
		0	-50	70	3.7	7
		10	-72	48	3.7	7
		-40	-44	32	3.5	39
		-32	-62	52	3.5	7
		-42	-60	60	3.2	39
	Occipital	16	-96	0	6.0	17
		-24	-82	14	5.1	19
		-10	-84	-8	4.8	18
		-18	-80	-18	4.6	18
		-44	-78	2	4.4	19
		28	-94	-8	4.4	18
		-26	-88	26	4.3	19
		-36	-68	-6	4.2	19
		-20	-78	-4	4.2	18
		26	-84	0	4.0	0
		-32	-72	-20	3.9	19
		40	-86	-10	3.3	18
		42	-90	2	2.8	18
	Others	-26	-72	-50	5.1	Cerebellum
		-28	-64	-30	4.9	Cerebellum
		44	-62	-42	4.3	Cerebellum
		2	-82	-22	4.3	Cerebellum
		-8	-74	-24	4.1	Cerebellum
		34	-56	-38	3.9	Cerebellum
		38	-70	-52	3.8	Cerebellum
		46	-72	-32	3.7	Cerebellum
		-36	-66	-54	3.4	Cerebellum
		38	-46	-44	3.0	Cerebellum
Death effect	Frontal	-30	26	4	4.5	13/45
	Parietal	24	-50	48	5.8	7

	38	-54	56	4.8	7/39
	18	-72	56	4.4	7
	-36	16	-4	3.7	13
	24	-58	58	3.1	7
	22	-72	42	2.9	7
Temporal	-36	16	14	3.3	44/45
	42	-56	4	3.3	37/19
Occipital	44	-66	-4	6.2	19/37

Supplementary table 3. Brain regions showing significant effects of AU and UFR models during delay period of the remain foraging trials. Formats are similar to those in Supplementary table 2.

Effect	Area	X	Y	Z	t-value	BA
AU	Frontal	-54	36	-6	7.6	47
		-2	34	38	7.3	8
		-38	24	48	6.8	8
		22	42	-2	6.7	10
		-20	54	26	6.6	47
		36	44	-10	6.4	47
		-58	26	4	6.4	45
		26	40	12	6.3	10
		22	56	26	6.1	10
		52	42	-12	6.1	47
		22	60	6	6.0	10
		-14	26	58	5.9	6
		-16	52	-16	5.8	11
		16	28	-12	5.8	11
		-44	46	-6	5.7	10
		6	36	-24	5.7	11
		-42	26	34	5.6	9
		-26	50	-4	5.5	10
		6	10	-26	5.5	25
		-28	40	4	5.2	46
		44	18	52	5.1	8
		14	38	52	5.0	8
		40	32	-16	5.0	47
		-44	42	-20	4.9	47
		-32	56	10	4.8	10
		-14	36	50	4.8	8
		16	48	-12	4.8	11
		-6	14	-26	4.7	11
		36	34	42	4.7	9
		-50	20	12	4.7	45
		4	48	-24	4.6	11
		12	20	58	4.6	6
		24	2	54	4.6	6
		48	28	-6	4.5	47
		6	50	40	4.5	9
		16	32	30	4.5	8
		-22	30	20	4.5	24
		38	12	62	4.3	8
		26	24	14	4.3	45
		-40	56	-12	4.2	10
		46	50	-4	4.1	10
		-28	6	24	4.1	44
		6	24	-28	4.1	11
		28	22	32	3.8	9
		30	44	36	2.9	9
		-16	36	0	2.9	47
		-30	32	52	2.9	8
	Parietal	24	-64	60	11.1	7
		-18	-72	56	10.8	7
		32	-56	54	9.4	7

	-54	-58	32	9.2	39
	-24	-54	52	8.9	7
	-44	-52	38	8.5	39
	8	-70	64	8.5	7
	28	-70	32	8.0	7
	-28	-70	28	7.5	39
	20	-78	54	6.8	7
	-38	-54	56	6.5	39
	32	-22	54	6.2	4
	36	-42	28	5.7	7
	16	-22	26	5.6	23
	54	-56	42	5.6	39
	42	-46	42	5.3	40
	-2	-44	70	5.1	6
	-4	-28	66	4.8	4
	34	-52	18	4.7	39
	24	-34	38	4.4	7
	4	-22	74	4.2	6
	-4	-58	70	4.1	7
	8	-50	76	3.6	7
	14	-34	72	3.5	1
	-2	-54	42	2.9	31
	-34	6	66	5.3	6
Temporal	48	-68	2	11.1	37
	44	-58	-8	8.5	37
	60	-40	-18	6.9	37
	-70	-22	-10	6.8	21
	68	-10	-18	6.4	21
	-20	-18	-30	6.3	36
	72	-30	-6	6.2	21
	-58	-32	-12	5.9	21
	-70	-32	0	5.9	21
	-56	-44	-20	5.7	37
	-68	-12	-18	5.6	21
	-40	-48	0	5.5	37
	-42	28	-10	5.5	47
	-62	0	-26	5.4	38
	60	26	8	5.3	9
	48	-32	-16	5.3	20
	-26	18	30	5.3	8
	-48	-40	-6	5.3	21
	60	-32	-6	5.1	21
	-56	-14	2	5.0	41
	-46	-32	-16	4.4	20
	-56	8	-34	4.4	38
	-48	16	-38	4.1	38
	-68	-44	-10	4.1	21
	70	-42	4	4.0	21
	-50	20	-24	3.5	38
	-68	-14	0	3.3	21
	22	-14	-20	3.1	36
Occipital	22	-92	-4	10.9	18
	18	-96	10	10.6	18
	38	-82	-2	9.6	19
	-44	-70	2	7.1	19
	-16	-94	-6	6.7	18
	-24	-86	14	6.6	19

		-42	-84	-4	6.2	18
		-26	-86	-6	5.9	18
		-12	-84	-18	5.6	18
		-32	-78	-16	4.4	19
		26	-38	22	3.8	18
	Others	-18	-34	10	9.2	Thalamus
		-4	2	-28	7.7	Pituitary
		-46	-42	-42	6.8	Cerebellum
		-24	-32	-2	6.2	Hippocampus
		-32	-40	4	5.8	Hippocampus
		-26	-32	-40	5.4	Cerebellum
		30	-36	2	5.4	Caudate
		-32	12	-8	5.3	Insula
		38	-26	-12	5.2	Hippocampus
		32	12	-6	4.8	Putamen
		2	-18	-46	4.8	Pons
		-38	-62	-34	4.7	Cerebellum
		16	-32	6	4.5	Thalamus
		-16	-40	-58	4.4	Cerebellum
		10	-78	-24	4.1	Cerebellum
		22	-82	-24	3.8	Cerebellum
		-22	18	-30	3.6	Pons
		34	-82	-22	3.4	Cerebellum
		50	-68	-22	3.4	Cerebellum
		-38	-78	-28	3.2	Cerebellum
		-22	24	8	3.1	Caudate
UFR	Frontal	32	28	-4	8.7	13/47
		48	22	2	8.2	45/44/47
		24	48	20	8.2	10
		42	26	42	8.0	8/9
		6	38	38	7.4	8
		42	6	34	6.9	8/6
		-32	16	-10	6.6	13
		12	14	64	6.4	6
		38	4	48	6.4	8/6
		8	28	64	6.4	6
		34	56	8	6.3	10
		4	22	52	6.1	8
		36	30	8	6.0	45
		36	36	22	5.9	9
		-28	50	20	5.4	10
		48	14	16	5.4	44
		-28	26	-2	5.3	13
		26	44	-16	5.1	47/11
		34	58	-4	5.1	10
		26	-2	44	5.0	6
		-44	22	-2	4.9	47/45
		-12	22	64	4.9	6
		14	32	28	4.8	9/8
		-28	64	-6	4.7	10
		44	38	0	4.6	46/47
		26	48	32	4.5	9
		26	4	56	4.4	6
		30	14	64	4.4	6
		50	46	-12	4.3	47/10
		14	38	10	4.1	32
		42	20	56	4.0	8

	40	30	-18	3.7	47
	-26	54	-14	3.4	10
	-32	60	6	3.2	10
	40	42	-20	3.1	10
	24	68	-4	3.0	10
	4	58	30	2.9	9
	10	46	-26	2.8	11
Parietal	52	-46	36	8.8	40/39
	26	-52	46	8.4	7
	28	-62	58	7.8	7
	14	-68	60	7.0	7
	54	-44	48	6.6	40
	26	-64	44	6.5	39
	60	-46	14	6.3	39/22
	48	-32	46	5.8	40
	-16	-66	60	5.4	7
	-26	-52	50	5.4	7
	40	-54	62	4.9	7
	-48	-56	54	4.8	39
	-50	-52	36	4.6	39
	42	-62	48	3.8	39
	10	-54	74	3.4	7
	-52	-40	58	3.4	40
	-50	-40	44	3.2	40
Temporal	46	-62	-6	12.2	37
	62	-26	-8	8.2	21/22
	50	-30	-10	7.7	21
	50	-58	-26	5.0	37
	62	-40	-14	3.9	37/21
	46	-46	-28	3.4	20
Occipital	26	-94	2	10.6	18
	30	-84	-4	10.2	18
	-48	-72	2	8.6	19
	28	-68	28	7.8	19
	30	-80	18	7.3	19
	-24	-94	6	6.6	18
	-32	-92	-4	5.2	18
	-6	-	-12	4.9	18
	-32	-66	-10	4.0	19

Supplementary table 4. Brain regions showing significant correlations between AuCs and AU effects during delay period of the remain foraging trials. Formats are similar to those in Supplementary table 2.

Effect	Area	X	Y	Z	t-value	BA
negative correlation	Frontal	-2	34	38	7.2	8
		-4	24	46	5.8	8
		32	54	-10	5.8	10
		36	34	42	5.6	9
		-36	44	8	5.5	46
		42	18	52	5.0	8
		-18	54	-16	4.9	11
		32	58	16	4.7	10
		18	40	-24	4.6	11
		-24	44	14	4.6	10
		24	62	6	4.3	10
		24	50	2	4.3	10
		-24	44	26	4.0	10
		16	32	18	4.0	10
		4	40	52	4.0	8
		34	6	58	3.9	6
		50	48	-4	3.9	10
		40	20	40	3.9	8
		-24	56	24	3.9	10
		26	36	18	3.8	10
		18	16	58	3.6	6
		18	40	-8	3.5	10
		40	34	30	3.5	9
		14	34	8	3.2	10
		-40	52	-16	3.1	10
		20	28	54	3.1	8
		-22	32	20	3.1	10
		-14	34	-28	3.0	11
		30	44	38	2.8	9
		12	32	28	2.8	9
	Parietal	30	-66	56	5.3	7
		-46	-56	46	5.2	39
		46	-48	52	4.9	40
		56	-44	34	4.8	40
		44	-46	32	4.7	39
		-36	-62	52	4.7	7
		32	-54	52	4.7	7
		-54	-58	32	4.6	39
		46	-62	52	4.5	39
		-24	-76	48	4.4	7
		-30	-80	38	4.3	39
		-28	-68	60	4.1	7
		-6	-72	58	3.9	7
		-46	-40	36	3.6	40
		36	-46	22	3.6	7
		-28	-82	26	3.2	39
		48	-34	48	3.1	40
	Others	-44	-74	-34	4.5	Cerebellum
		-44	-60	-38	3.6	Cerebellum
		-52	-54	-30	3.2	Cerebellum

		-50	-56	-38	3.0	Cerebellum
		-40	-74	-24	2.8	Cerebellum
positive correlation	Others	24	-16	-12	3.5	Hippocampus

Supplementary table 5. Brain regions showing significant effects of AU and UFR models during delay period of the experience trials. Formats are similar to those in Supplementary table 2.

Effect	Area	X	Y	Z	t-value	BA
AU	Frontal	-36	30	-14	4.7	47
		-52	28	-2	4.6	45
		-34	24	16	4.4	44/45
		-48	42	-8	4.2	47
		-32	4	20	3.7	44
		-26	16	28	3.5	44
		-54	28	10	3.4	45
		-44	42	-20	2.9	47
	Parietal	30	-70	32	6.9	7
		26	-50	48	6.2	7
		-28	-62	50	5.6	7
		-26	-72	28	5.5	39
		28	-60	56	5.1	7
		32	-60	42	4.3	7/39
		-18	-72	56	4.3	7
		-50	-62	-28	4.2	39
		38	-34	28	3.9	39/40
		-38	-52	32	3.8	39
		0	-68	62	3.7	7
	Temporal	46	-62	-6	9.2	37
		-48	-54	-14	6.8	37
		-46	-42	-12	6.2	37
		48	-44	-10	5.6	37
		-34	-50	-16	4.7	37
		34	-52	-16	4.6	37
		-46	10	-12	4.1	38/22
		-54	-50	-28	4.1	37
		-36	-8	-24	3.5	36
	Occipital	20	-98	10	13.0	18
		22	-92	-2	11.5	18
		-8	-84	-4	10.2	18
		36	-84	2	9.2	18
		-4	-86	-16	7.2	18
		-18	-	2	6.6	18
		-6	-66	6	6.4	18
		-32	-80	-16	5.8	19
		30	-76	-12	5.5	19
		-24	-86	12	5.4	19
		-10	-48	-2	5.4	19
		-26	-80	-2	5.3	18
		-32	-68	-10	5.3	19
		-44	-68	2	4.8	19
		-26	-84	26	4.5	19
		24	-68	8	4.2	17
		-28	-66	12	4.1	18
	Others	-26	-30	-4	10.2	Hippocampus
		26	-32	4	7.3	Hippocampus
		-18	-36	4	6.3	Hippocampus
		20	-26	24	6.2	Thalamus
		-28	-20	-14	6.0	Hippocampus

		14	-16	28	5.6	Caudate
		30	-24	-10	5.3	Hippocampus
		-16	-24	26	5.1	Caudate
		-40	-72	-22	5.8	Cerebellum
		38	-70	-22	5.2	Cerebellum
		-46	-42	-34	5.0	Cerebellum
		32	-66	-46	4.3	Cerebellum
		48	-58	-28	4.2	Cerebellum
		40	-38	-42	4.2	Cerebellum
		50	-44	-32	4.2	Cerebellum
		-30	-32	-36	4.1	Cerebellum
		6	-76	-24	3.9	Cerebellum
		22	-80	-24	3.7	Cerebellum
		-34	-64	-36	3.2	Cerebellum
UFR	Frontal	24	2	52	7.0	6
		46	8	28	6.3	6/44
		56	26	10	5.3	44/45
		28	14	66	4.7	6
		22	22	26	4.5	9
		34	10	32	4.4	8/44
		32	24	2	4.2	13
		28	40	32	4.2	9
		30	34	18	3.9	9
		34	16	20	3.9	44
		42	32	30	3.4	9
		48	8	56	3.3	6/8
		56	20	26	3.1	44
		48	28	-10	2.8	47
	Parietal	-24	-52	54	9.5	7
		-18	-68	56	9.5	7
		34	-70	28	8.7	39
		34	-54	56	8.2	7
		18	-64	56	8.2	7
		-24	-68	30	7.5	7
		-14	-60	68	6.7	7
		-20	-72	42	6.6	7
		50	-30	50	6.2	40
		14	-60	70	5.9	7
		-28	-46	44	5.7	7
		50	-20	62	5.5	1
		10	-46	44	4.2	31
		8	-56	60	3.9	7
		38	-44	24	3.9	39
		-54	-60	14	3.6	39
		38	-40	70	3.1	1
	Temporal	50	-52	-14	6.6	37
		60	-34	4	6.3	21/22
		68	-18	-2	6.1	22
		50	-42	16	6.0	22
		-46	-66	-12	5.5	37
		-40	-48	-16	5.3	37
		48	-40	-18	5.2	37
		-50	-2	-16	5.1	22
		34	-68	-12	5.1	37
		-56	-40	0	5.0	21
		-46	16	-32	4.9	38
		-58	8	-16	4.7	38

	-52	-12	-6	4.4	22
	-36	-44	2	4.3	21
	-58	-44	12	4.3	22
	48	-32	6	4.2	22
	-48	-40	-12	4.2	21
	-62	-18	-2	4.2	21
	62	-32	16	4.2	22
	-44	-52	8	4.1	37
	46	-26	-6	4.0	21/22
	-34	22	-34	3.6	38
	-46	-24	-4	3.1	22
	-32	-50	20	3.0	39
Occipital	46	-66	-2	11.1	19
	38	-76	2	10.8	18/19
	-24	-84	10	8.8	18
	-24	-80	26	8.5	19
	-42	-70	2	8.5	19
	26	-86	14	7.7	18
	-40	-84	2	7.6	18
	30	-80	-8	7.2	18/19
	12	-90	-4	7.2	17
	38	-86	10	6.4	19
	-34	-78	-14	5.7	19
	-24	-82	-8	5.6	18
	-8	-92	-12	4.3	18
Others	48	-42	-32	5.6	Cerebellum
	-40	-38	-34	4.8	Cerebellum
	40	-38	-44	3.4	Cerebellum
	-20	-38	6	3.2	Hippocampus

Table 6. Brain regions showing significant correlation between AuCs and AU effects during delay period of the experience trials. Formats are similar to those in Supplementary table 2.

Effect	Area	X	Y	Z	t-value	BA
Positive correlation	Others	28	-14	-14	3.3	Hippocampus
Negative correlation	Frontal	-24	44	18	6.0	10
		-30	34	10	4.1	45/46
		26	56	-6	3.6	10
		40	48	-4	3.0	10