



OPEN

Effects of fictive reward on rat's choice behavior

SUBJECT AREAS:

REWARD

DECISION

Ko-Un Kim^{1,3,4}, Namjung Huh^{1,3}, Yunsil Jang^{1,2}, Daeyeol Lee⁵ & Min Whan Jung^{1,2,3,4}

Received

14 August 2014

Accepted

29 December 2014

Published

27 January 2015

Correspondence and requests for materials should be addressed to M.W.J. (mwjung@kaist.ac.kr)

¹Center for Synaptic Brain Dysfunctions, Institute for Basic Science, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea, ²Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea, ³Neuroscience Laboratory, Institute for Medical Sciences, Ajou University School of Medicine, Suwon 443-721, Korea, ⁴Neuroscience Graduate Program, Ajou University School of Medicine, Suwon 443-721, Korea, ⁵Department of Neurobiology, Yale University School of Medicine, New Haven, CT 06510, USA.

Choices of humans and non-human primates are influenced by both actually experienced and fictive outcomes. To test whether this is also the case in rodents, we examined rat's choice behavior in a binary choice task in which variable magnitudes of actual and fictive rewards were delivered. We found that the animal's choice was significantly influenced by the magnitudes of both actual and fictive rewards in the previous trial. A model-based analysis revealed, however, that the effect of fictive reward was more transient and influenced mostly the choice in the next trial, whereas the effect of actual reward was more sustained, consistent with incremental learning of action values. Our results suggest that the capacity to modify future choices based on fictive outcomes might be shared by many different animal species, but fictive outcomes are less effective than actual outcomes in the incremental value learning system.

Fictive outcomes refer to rewards or punishments that have been observed or inferred but not directly experienced. It is well established that human choice behavior is influenced by actual as well as fictive outcomes^{1,2}. Recent studies have shown that choice behavior in non-human primates is also influenced by fictive outcomes; monkeys tend to choose a target that was associated with a large fictive reward in the previous trial^{3,4}. Thus, cognitive capability to compare actual and fictive outcomes and adjust subsequent choice behavior according to the outcome of the comparison is not unique to human, but also found in other animals.

The capability to adjust future choices according to fictive outcomes would enable an animal to make adaptive choices in the future without directly experiencing all possible outcomes, and hence would be advantageous for its survival in many natural settings. One might then expect that this capability would be widespread in the animal kingdom. On the other hand, adjusting future choices based on the information obtained from both actual and fictive outcomes might require a highly advanced cognitive capacity, thereby limiting its presence in only some animals such as primates. Empirical studies using non-primate animals are needed to resolve this matter. In this regard, a recent study employing a sequential wait-or-skip choice task has shown that, encountering a long-delay choice after skipping a short-delay choice, rats tended to look backwards toward the foregone option and wait for the long delay⁵. Although this study showed potential 'regret'-induced behavioral changes in rats, it did not establish that information on fictive reward can be used by rats in an adaptive manner to increase future gain. To date, the capability to compare actual and fictive rewards and adjust future choices according to the comparison has been demonstrated unequivocally only in primates.

In the present study, in order to address this issue, we examined effects of actual and fictive rewards on choice behavior of the rat, which is one of the most widely used experimental animal models. We found that rat's choice behavior was influenced by both actual and fictive rewards, but the effect of fictive rewards was more transient than that of actual reward.

Results

Effects of actual and fictive rewards on choice. Seven rats were tested in a binary choice task (30 trials per session) in which both targets delivered randomly chosen magnitudes of reward (1, 3 or 5 sucrose pellets). Reward locations were adjacent and they were divided by a transparent wall containing numerous holes (Fig. 1a). Moreover, each magnitude of reward was associated with a distinct number of auditory cues (1, 2 and 3 tones for 1, 3 and 5 sucrose pellets, respectively; Fig. 1b) and fictive reward delivery preceded actual reward delivery (Fig. 1c) in order to provide plenty of sensory information about the magnitude of fictive reward before actual reward became available to the animal.

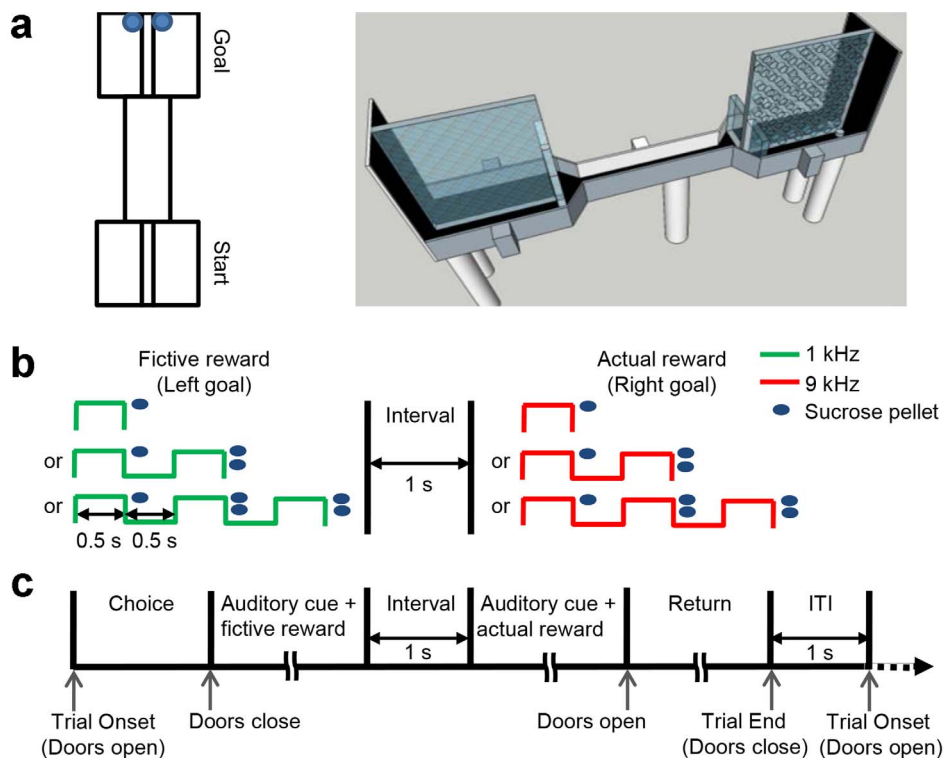


Figure 1 | Behavioral task. (a) A top-down view (left) and a three dimensional drawing (right) of the maze. (b) Schematic diagram showing auditory cues (one, two or three tones; left goal, 1 KHz; right goal, 9 KHz) and associated magnitudes of reward (total one, three or five sucrose pellets, respectively) delivered at the left and right goals. One sucrose pellet was delivered after the first tone, and two pellets were delivered after each additional tone for reward magnitude 3 and 5. (c) Temporal sequence of the task. ITI, inter-trial interval.

All animals showed significant biases towards either the left or right goal (choice bias; binomial test, $p < 0.05$; mean proportion of choosing the preferential goal, $72.9 \pm 5.0\%$), and four animals showed significant biases to repeat the same goal choice as in the previous trial regardless of the magnitudes of actual and fictive rewards (stay bias; binomial test, $p < 0.05$; mean proportion of stay, $57.7 \pm 6.5\%$). Despite these biases, the animal's choice was consistently influenced by the magnitudes of fictive as well as actual rewards in the previous trial. Fig. 2a shows the proportion of stay trials for each of nine combinations of actual and fictive reward magnitudes in the previous trial. In general, the animals tended to repeat the same goal choice as in the previous trial as the magnitude of actual reward in the previous trial increased and, conversely, as the magnitude of fictive reward in the previous trial decreased.

When the animal's choice behavior was analyzed according to the magnitude of actual reward irrespective of the magnitude of fictive reward, the proportion of stay trials increased monotonically as a function of the magnitude of actual reward in the previous trial in all animals except one (rat #4; Fig. 2b, left). Conversely, the proportion of stay trials decreased monotonically as a function of the magnitude of fictive reward in the previous trial in all animals (Fig. 2b, right). Proportions of stay trials significantly deviated from an even distribution (χ^2 -test, $p < 0.05$) across different magnitudes of actual reward in the previous trial in six out of seven animals, and across different magnitudes of fictive reward in the previous trial in four out of seven animals.

We then ran a logistic regression analysis to test whether effects of actual and fictive rewards persist across multiple trials. The animals tended to repeat the previous goal choice as the magnitude of previous actual reward increased (actual reward effect of $t-1$ trial, t -test, $p < 0.05$ in all animals and on average), and, conversely, switch goal choice as the magnitude of previous fictive reward increased (fictive reward effect of $t-1$ trial, $p < 0.05$ in five animals and on average;

Fig. 3). The magnitude of the regression coefficient (i.e., effect size) for $t-1$ trial was not significantly different between actual and fictive rewards (0.428 ± 0.088 vs. 0.289 ± 0.079 ; paired t -test, $p = 0.264$), indicating that effect sizes of actual and fictive rewards on the animal's choice in the next trial were similar. The animals also tended to repeat the goal choice made two trials before more often as the magnitude of actual reward two trials before increased (actual reward effect of $t-2$ trial, $p < 0.05$ in four animals and on average), but the effect of the magnitude of fictive reward two trials before (fictive reward effect of $t-2$ trial) was significant only in two animals and not on average (Fig. 3). Finally, the animals tended to repeat the goal choice made three trials before more often as the magnitude of fictive reward three trials before increased (fictive reward effect of $t-3$ trial, $p < 0.05$ in five animals and on average). Note that fictive reward effect of $t-1$ trial is in opposite direction from fictive reward effects of $t-2$ and $t-3$ trials. Thus, the magnitude of actual, but not fictive, reward had a consistent influence on the animal's goal choice over multiple trials. The regression analysis also revealed significant bias to choose a particular goal (left vs. right) in all animals, significant effect of the location of the start box in five animals, and significant effect of the previous choice in six animals (Table 1).

Learning models of choice behavior. Reinforcement learning (RL) models can explain humans' and animals' choice behavior well in diverse experimental settings⁶⁻⁸. However, recent studies suggest that adding an additional process, such as a win-stay-lose-switch strategy⁹, a perseveration factor¹⁰ or an additional RL term with a short time constant¹¹, to a simple RL model better accounts for humans' and animals' choice behavior in various experimental settings. In the present study, therefore, we used a hybrid model that included a separate term to capture the effect of the reward magnitude in the most recent trial in addition to the standard incremental value updating used in a simple RL algorithm. In

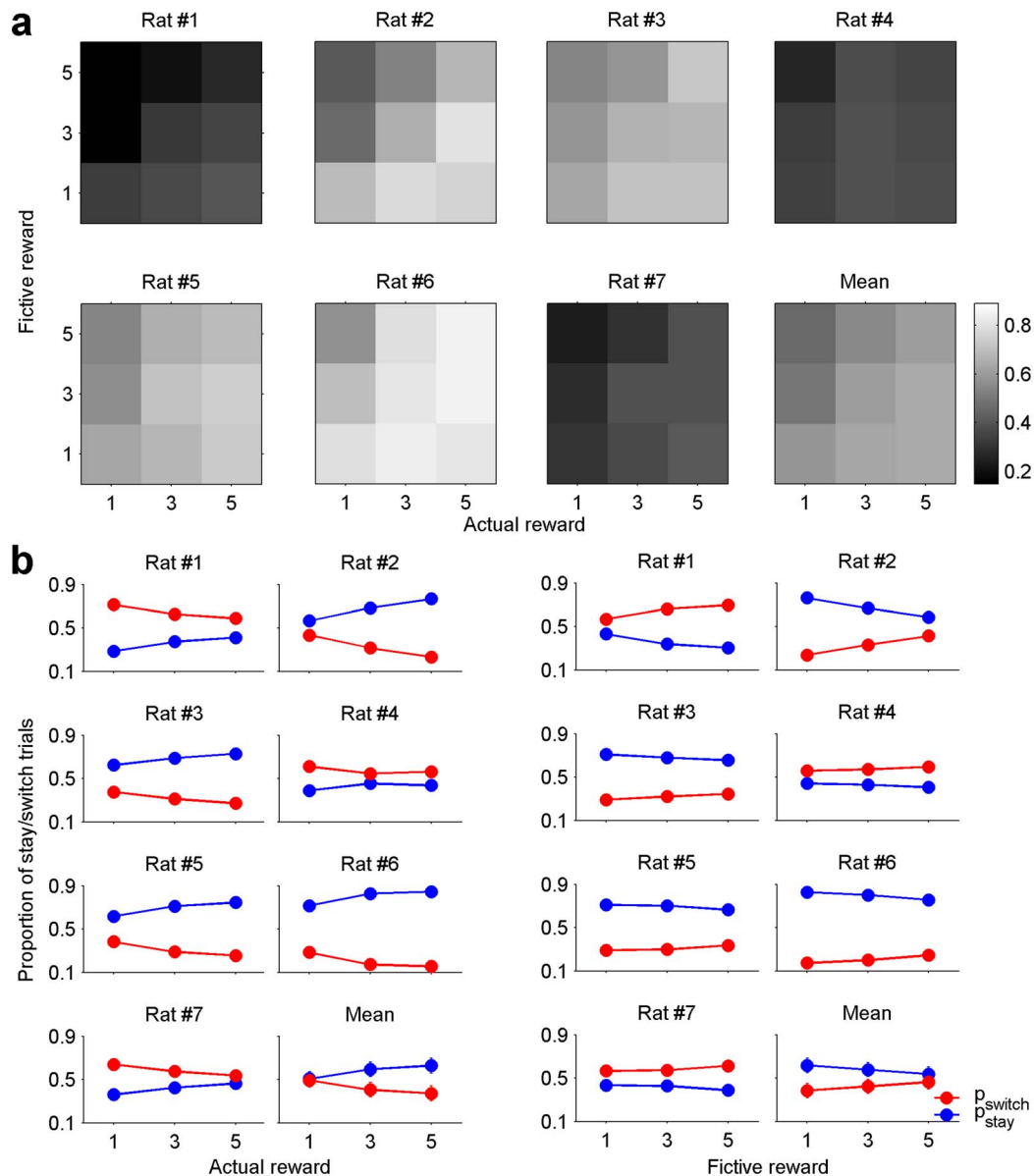


Figure 2 | Dependence of the animal's choice on the magnitudes of actual and fictive rewards. (a) Trials were divided into nine groups according to the combination of actual and fictive reward magnitudes in the previous trial. The proportion of stay trials is indicated in gray scale. Mean, the proportion of stay trials was averaged across animals. (b) Trials were divided into three groups according to the magnitude of actual (left) or fictive (right) reward in the previous trial. Blue circles, proportions of stay trials; red circles, proportions of switch trials. Mean, the proportions of stay and switch trials were averaged across animals (mean \pm SEM).

addition, the value of chosen goal was increased by a constant amount in each trial (see Methods) to reflect the run-length-dependent increase in stay probability seen in the data (Fig. 4). In summary, the full model included the following terms: effect of actual reward magnitude on the animal's choice in the next trial (transient effect of actual reward), effect of fictive reward magnitude on the animal's choice in the next trial (transient effect of fictive reward), effect of actual reward on the value of chosen goal (sustained effect of actual reward), effect of fictive reward on the value of unchosen goal (sustained effect of fictive reward), progressive increase of the value of chosen goal, stay bias, and choice (left vs. right) bias. We compared the full model and all possible reduced models, along with the full model with equal learning rates for actual and fictive rewards and its reduced models, using AIC and BIC¹².

Results of the model comparison are summarized in Table 2, and mean parameters for the full and some reduced models are shown in Table 3. According to AIC, the best model for four animals was either

the full model ($n = 1$), the full model without stay bias ($n = 2$) or the full model with equal learning rates for actual and fictive rewards ($n = 1$). The best model for the remaining three animals was the one without transient ($n = 1$) or sustained ($n = 2$) effect of fictive reward (with or without choice bias). According to BIC, which penalizes the use of additional variables more severely than AIC, the hybrid model without sustained effect of fictive reward and stay bias was the best model for four animals. The best model for other two animals was the hybrid model without transient effect of fictive reward instead of sustained effect of fictive reward, and, for the remaining one animal, the best model was the one containing only transient effects of actual and fictive rewards along with choice and stay biases. Thus, AIC tended to select the model including both transient and sustained effects of actual and fictive rewards, whereas BIC tended to select the model without sustained effect of fictive reward as the best model. In the full model, the learning rate of fictive reward (0.037 ± 0.013) was markedly lower than that of actual reward (0.160 ± 0.009 ; paired t -

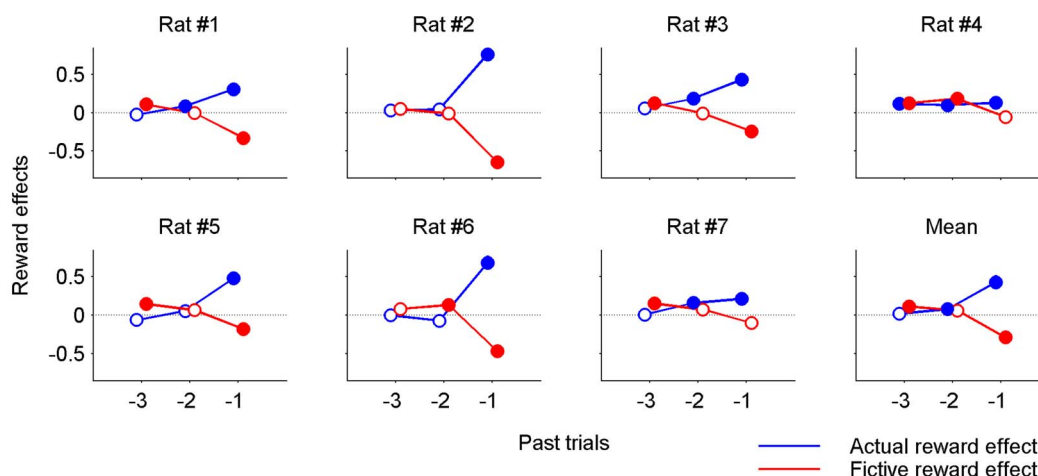


Figure 3 | Results of logistic regression analysis. Effects of actual and fictive rewards in the previous three trials on the rat's current choice were examined by performing a trial-by-trial analysis of rat's choices using a logistic regression model. A positive (or negative) coefficient indicates that the animal's past and current choices tended to be the same (or different). Error bars, standard errors of the coefficient estimates. The animals tended to repeat the same goal choice (stay) as the magnitude of the actual reward in the previous trial ($t-1$) increased (positive coefficients for actual reward for $t-1$ trials), whereas they tended to switch their choice as the magnitude of the fictive reward in the previous trial ($t-1$) increased (negative coefficients for fictive reward for $t-1$ trials). Mean, regression coefficients were averaged across animals (mean \pm SEM; error bars are too small to see). Filled symbols denote statistical significance ($p < 0.05$).

test, $p = 4.6 \times 10^{-6}$). By contrast, transient effects of actual and fictive rewards were similar (0.477 ± 0.108 and 0.322 ± 0.085 , respectively; $p = 0.280$; Fig. 5). These results indicate that actual and fictive rewards influenced the animal's goal choice in the next trial with similar strengths, but fictive reward had little or only a small influence on the value of unchosen goal unlike the effect of actual reward on the value of chosen goal.

To test the stability of choice behavior across sessions (90 for each animal) we divided the entire behavioral data (90 sessions per animal) into early (sessions 1–30), middle (sessions 31–60) and late (sessions 61–90) phases and applied the same analyses to each phase separately. Similar results were obtained across phases (Fig. 6), indicating that effects of actual and fictive rewards on the animal's subsequent choices were consistently maintained across sessions.

Discussion

We examined choice behavior of rats in a binary choice task that revealed the magnitudes of rewards delivered at chosen as well as unchosen goals. The animals' choice was systematically influenced not only by the reward they actually received (actual reward), but also by the reward they could have obtained had they made the alternative choice (fictive reward). A model-based analysis revealed that the effect of fictive reward was more transient than expected from a traditional RL algorithm. These results indicate that rodents are capable of adjusting subsequent choice behavior based on the outcome of the comparison between actual and fictive rewards, but

fictive reward is much less effective than actual reward in incremental value learning.

Effects of fictive reward on choice. The animal's choice was systematically influenced by the magnitudes of both actual and fictive rewards. This conclusion was consistently supported by the analysis examining the proportion of stay trials as a function of actual and/or fictive reward magnitudes (Fig. 2), a logistic regression analysis (Fig. 3), and computational modeling (Table 2). It should be noted that magnitudes of actual and fictive rewards were determined randomly in the present study. By contrast, in a typical two-armed bandit task, choosing the target with a higher estimated value is more likely to yield a higher-value (larger, more probable or less delayed) reward compared to random target selection. In the present study, the animal's choice had no consequence on the magnitudes of subsequent rewards, which might account for the observation that variable levels of choice and stay biases were observed across the animals and that these biases were relatively large in some animals (Fig. 2 and Table 1). Nevertheless, in all of our analyses, significant effect of fictive reward was found in the majority of animals. Moreover, the effect size of fictive reward was as large as that of actual reward, which is consistent with previous findings in monkeys⁴ and humans¹³. These results indicate that the effect of fictive reward on rat's choice behavior was robust.

The outcome of an action (reward or punishment) plays an important role in modifying animal's subsequent behavior¹⁴. However, learning can take place in the absence of reinforcement;

Table 1 | Choice bias, effect of starting location, and effect of the previous choice. Shown are the regression coefficients for choice bias, start box location, and the previous choice of the logistic regression model. All animals showed significant choice bias (preferential choice of the left or right target), five animals showed significant effect of the starting location, and six animals showed significant effect of the previous choice

Rat No.		1	2	3	4	5	6	7
Choice bias	Coeff.	0.440	1.743	1.320	0.235	1.846	1.829	0.459
	p -value	3.9×10^{-20}	7.9×10^{-65}	1.2×10^{-52}	2.1×10^{-7}	2.6×10^{-52}	8.8×10^{-50}	1.1×10^{-21}
Start box	Coeff.	0.089	-0.265	0.163	-0.088	-0.058	-0.066	-0.095
	p -value	0.044	0.014	0.003	0.046	0.297	0.300	0.031
Choice effect	Coeff.	-0.597	-0.514	-0.617	-0.599	-1.175	-0.239	-0.707
	p -value	1.6×10^{-4}	2.4×10^{-6}	0.002	2.0×10^{-4}	1.7×10^{-7}	0.327	1.2×10^{-5}

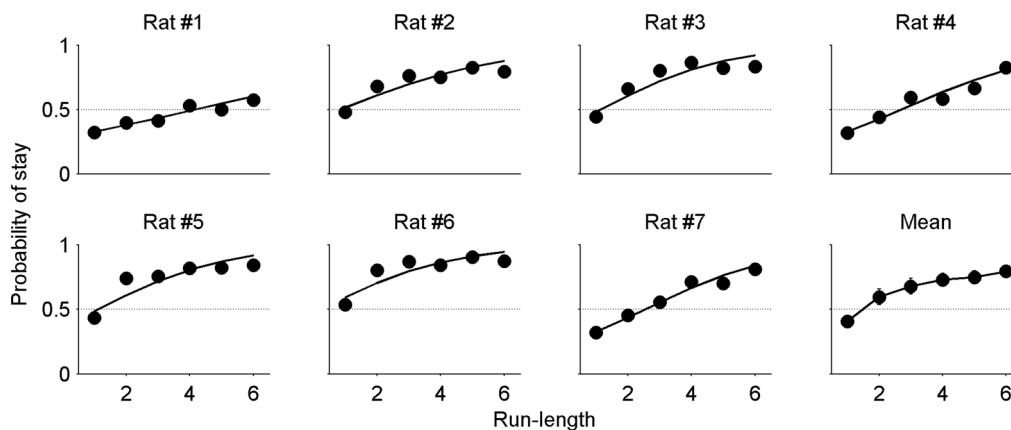


Figure 4 | Dependence of stay choice on run-length. The probability to stay (P_{stay}) was plotted as a function of run-length (the number of consecutive choices of the same goal) for each animal. Each data point was calculated using the choice data across all sessions for a given animal. The lines for individual animals (Rat #1–7) were determined by logistic regression relating animal’s choice and run-length. The animals tended to repeat the same goal choice as run-length increased. Mean, P_{stay} values were averaged across animals (mean \pm SEM; error bars are too small to see for some data points).

rats are capable of acquiring knowledge about the layout of an environment without reinforcement and use this information later in an adaptive manner^{15,16}. Our finding of rat’s capability of using information about fictive reward in deciding which goal to choose in the next trial (i.e., learning without direct reinforcement) is in line with these results. More recent studies^{5,17–19} have found neural activity in the rodent brain that might be related to vicarious trial-and-error²⁰ and foregone choice/outcome. Combined with our findings, these results suggest that rats, as humans and monkeys, have a propensity to compare actual and fictive outcomes and adjust their subsequent choice behavior according to the result of the comparison. It would be advantageous for animals to consider all possible outcomes, experienced or inexperienced, and choose the best option in many natural settings. Our results suggest that the propensity to consider both actual and fictive outcomes and modify choice behavior appropriately might be shared by many different animal species. It remains to be determined how widespread this capability is in the animal kingdom.

Differential effects of actual and fictive rewards on value updating. Both actual and fictive rewards modified the animal’s

next choice and their effect sizes were similar in magnitude. Although unexpected outcomes influence subsequent choices over multiple trials by updating value functions in standard RL algorithms²¹, such value learning was found only for actual reward, and not for fictive reward, in the present study. This indicates that the animals retained the information about both fictive and actual rewards until making a new choice in the next trial, but this information was not used effectively in modifying expected reward magnitude (action value) of the unchosen goal. Previous studies in humans have found fictive reward prediction error signals in the brain that were correlated with upcoming choices of the subjects^{13,22–25}. These findings suggest that the information on fictive reward can be used to update the value of unchosen action in humans. Different results from these human studies and ours might be due to species difference; humans might be able to change the value of unchosen action based on fictive reward more efficiently than rodents. However, they might also arise from the difference in task structures. In the previous human studies, reward probabilities were correlated across trials so that it was advantageous for humans to update expected reward probabilities based on both actual and fictive outcomes. By contrast, in our study,

Table 2 | Results of model comparison. Shown are AIC (top) and BIC (bottom) values for the models that best explained choice behavior of at least one animal. The best model (the smallest AIC or BIC value) for each animal is indicated in bold. Numbers in the parenthesis indicate the number of parameters for each model (including the inverse temperature). V_{act} and V_{fic} , sustained effects (value learning) of actual and fictive rewards, respectively; V , sustained reward effects of the model with equal learning rates for actual and fictive rewards ($V = V_{act} = V_{fic}$); T_{act} and T_{fic} , transient effects of actual and fictive rewards, respectively; V_q , constant increment of chosen value; b_{st} , stay bias; b_L , bias to select the left goal

AIC of optimal models.							
Rat No.	1	2	3	4	5	6	7
$V_{act}, T_{act}, V_{fic}, T_{fic}, V_q, b_{st}, b_L$ [8]	2194	1517	1570	2005	1435	1231	2110
$V_{act}, T_{act}, V_{fic}, T_{fic}, V_q, b_L$ [7]	2184	1513	1572	2001	1431	1228	2112
$V_{act}, T_{act}, V_{fic}, V_q, b_L$ [6]	2242	1556	1582	1999	1437	1238	2111
$V_{act}, T_{act}, T_{fic}, V_q, b_{st}, b_L$ [7]	2196	1515	1570	2015	1433	1234	2137
$V_{act}, T_{act}, T_{fic}, V_q, b_L$ [6]	2184	1513	1573	2017	1431	1232	2153
$V, T_{act}, T_{fic}, V_q, b_{st}, b_L$ [7]	2184	1513	1572	2001	1431	1228	2112
BIC of optimal models.							
Rat No.	1	2	3	4	5	6	7
$V_{act}, T_{act}, V_{fic}, V_q, b_L$ [6]	2275	1589	1615	2032	1470	1271	2144
$V_{act}, T_{act}, T_{fic}, V_q, b_L$ [6]	2217	1546	1606	2050	1464	1265	2186
$T_{act}, T_{fic}, b_{st}, b_L$ [4]	2293	1594	1745	2445	1569	1261	2349



Table 3 | Model parameters. Shown are parameters (mean \pm SEM) of the full and three reduced models (model 1, sustained effect of actual reward only; model 2, sustained and transient effects of actual reward; model 3, sustained effects of actual and fictive rewards). AIC and BIC values are also shown. All models included three bias terms [constant increment of chosen value (V_q), stay bias (b_{st}), and left-choice bias (b_L)]

	Model 1	Model 2	Model 3	Full model
α_{actual}	0.227 \pm 0.086	0.133 \pm 0.006	0.454 \pm 0.187	0.160 \pm 0.009
α_{fictive}	—	—	0.312 \pm 0.173	0.037 \pm 0.013
k_{actual}	—	0.447 \pm 0.095	—	0.461 \pm 0.097
k_{fictive}	—	—	—	0.323 \pm 0.090
β	0.501 \pm 0.051	0.138 \pm 0.055	0.541 \pm 0.067	0.076 \pm 0.037
b_{st}	-0.661 \pm 0.137	-1.561 \pm 0.147	-1.780 \pm 1.067	-1.037 \pm 0.151
b_L	0.563 \pm 0.274	0.485 \pm 0.214	0.729 \pm 0.330	0.493 \pm 0.207
V_q	0.186 \pm 0.119	0.353 \pm 0.038	1.459 \pm 1.129	0.439 \pm 0.042
AIC	1782 \pm 136	1751 \pm 142	1752 \pm 141	1723 \pm 142
BIC	1809 \pm 136	1784 \pm 142	1784 \pm 141	1767 \pm 142
Number of parameters	5	6	6	8

rewards were always delivered and their magnitudes were randomized across trials, so that there was no advantage of keeping track of reward magnitudes. Rats might be able to keep track of values of unchosen actions as efficiently as those of chosen actions if doing so is advantageous for maximizing a long-term sum of rewards, which remains to be tested.

The animals updated the value of chosen goal in the present task, even though this was not advantageous. Numerous studies in humans and animals have shown strong tendency to change their behavior based on experienced reward even when their behavior has no causal relationship with reward delivery (e.g., refs 26–29). These results suggest that humans and animals have a strong propensity to keep track of values based on consequences of their actions. In other words, it is possible that the neural system supporting an RL-like process is always used as a default³⁰. On the other hand, fictive reward, despite its ability to influence next choice behavior, may not have the same privilege to activate value-learning system as strongly as actual reward. It is conceivable, for example, that fictive reward activates primarily cognitive executive control systems, such as dorsolateral prefrontal cortex in primates^{31,32} and medial prefrontal cortex in rats^{33–35}, whereas actual reward activates an additional reward/value representation system that supports a simple RL process, such as the basal ganglia³⁶, so that actual reward automatically activates an RL-like process whereas fictive reward contributes to value learning only when it has a predictive value.

Methods

Subjects. Seven young male Sprague Dawley rats (9–11 weeks old, 300–400 g) were individually housed in their home cages and handled extensively for 5–9 d with free access to food and water. They were then gradually food deprived over 7 d so that their body weights were maintained at 80 ~ 85% of their ad libitum body weights throughout the experiments. The colony room was maintained at 12-h light and 12-h dark cycle (light on: 8 p.m.; light off, 8 a.m.) and behavioral training and testing were done during the dark phase. All experiments were carried out in accordance with the regulations and approval of the Ethics Review Committee for Animal

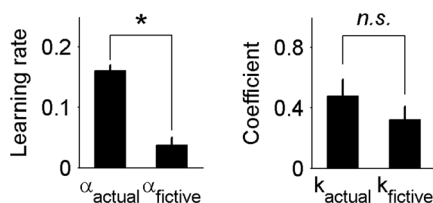


Figure 5 | Sustained and transient effects of actual and fictive rewards. Shown are learning rates (α_{actual} and α_{fictive} ; i.e., coefficients for sustained effects) and coefficients for transient effects (k_{actual} and k_{fictive}) of actual and fictive rewards of the full model averaged across animals (mean \pm SEM). The asterisk denotes statistical significance (paired *t*-test, $p < 0.05$).

Experimentation of Ajou University School of Medicine and Korea Advanced Institute of Science and Technology.

Apparatus. The maze was 100 cm long, 16 cm wide, and elevated 30 cm from the floor (Fig. 1a). It had two start boxes and two goal boxes (30 cm long and 8 cm wide each) along with a central track (40 cm long and 8 cm wide). There were 2 cm high walls along the entire track and 35 cm high walls at both ends of the maze. Each of the starting and goal boxes had a photobeam sensor and a sliding door. The two starting boxes were separated by a transparent acrylic wall (0.5 cm thick and 35 cm high), and the two goal boxes were separated by a transparent acrylic wall (0.5 cm thick and 35 cm high) containing 117 holes (diameter, 0.6 cm).

Behavioral task. The animals were tested in a binary choice task. They were allowed to choose one of two goals freely in each trial (30 trials per session). The same start box was used throughout a given session, and two start boxes were used alternately across successive sessions. The door of the unused start box for a given session was kept closed throughout the session. Each trial began by opening the door of the start box that contained the animal and the doors of both goal boxes (three operating doors for a given session). Once the animal arrived at a goal box, the three operating doors were closed, and food reward was delivered first at the unchosen goal, and then at the chosen goal. The food reward was one, three or five sucrose pellets (20 mg, Dustless Precision Pellet, Bio Serv., NJ, USA), which were delivered manually. The food delivered at the unchosen goal was removed manually before delivering food at the chosen goal. The magnitude of reward (i.e., the number of food pellets) was determined randomly and independently for both chosen and unchosen goals in each trial. Reward locations were adjacent and they were divided by a transparent wall containing numerous holes in order to facilitate sensory information about the fictive reward to be available to the animals. Moreover, each magnitude of reward was associated with a distinct number of auditory tones (1, 2 and 3 tones for 1, 3 and 5 sucrose pellets, respectively; left goal, 1 KHz; right goal, 9 KHz; 0.5 s duration with 0.5 s interval) that preceded each reward delivery (Fig. 1b). Thus, plenty of sensory information about the magnitude of fictive reward was available to the animals. The three operating doors were opened \sim 1 s following reward delivery so that the animals could return to the start box after consuming reward. The three operating doors were closed as soon as the animal returned to the start box, which concluded a trial (Fig. 1c). The next trial began after 1 s of inter-trial interval. The animal's arrival at the start and goal boxes was detected by four sets of photobeam sensors, and opening and closing of doors were controlled automatically by a personal computer using LabView software (National Instruments, TX, USA).

Behavioral training and testing. Following 5–9 d of acclimation to the maze, the animals went through a shaping period during which each of six auditory stimuli (1, 2 or 3 tones at 1 or 9 KHz) was associated with a given magnitude of reward at the start box (5–7 d; location of start box alternated across sessions; reward magnitude randomized across trials; 50–60 daily trials) and then at the goal box (10–14 d; 1 KHz, left goal; 9 KHz, right goal; reward magnitude randomized across trials; the animals were forced to visit a particular goal by opening the door of a randomly chosen goal box; 55–80 daily trials). The animals were then trained to choose between the left and right goals that delivered nine combinations of reward magnitudes (left, 1, 3 or 5; right, 1, 3 or 5 sucrose pellets; fictive reward was delivered before actual reward and each reward was preceded by its corresponding auditory stimulus; 10 consecutive trials for each reward magnitude combination; 90 trials per day; the sequence of reward combinations was randomized). This step was to further teach the animals that the auditory cue at the unchosen goal was associated with a particular magnitude of reward. The animals were trained this way until they made $>70\%$ choices of larger rewards in all unequal combinations (10–14 d of training). They were then tested in the main task during which actual and fictive reward magnitudes were randomly determined in each trial (total 90 sessions per animal; 30 trials per session, 2 sessions per day). The animals were given approximately 10 mg of food (general rat chow) following each session.

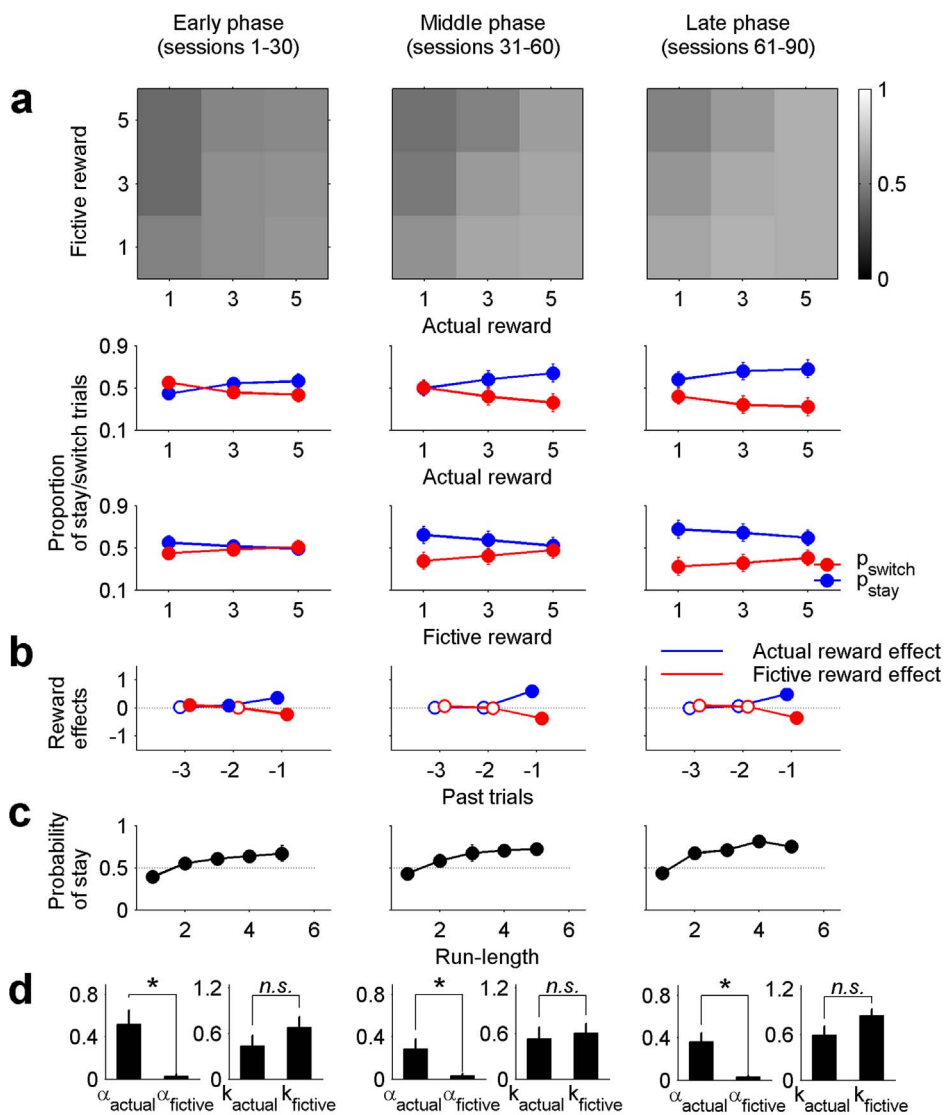


Figure 6 | Effects of actual and fictive reward magnitudes on the animal's choice behavior in different phases. The entire behavioral data (90 sessions) were divided into three phases (early, middle and late) and the same analyses used in Fig. 2–5 were applied to each phase. Shown are values averaged across the animals. (a) Same formats as in Fig. 2. (b) Same formats as in Fig. 3. (c) Same formats as in Fig. 4. (d) Same formats as in Fig. 4.

Analysis. *Stay probability* Stay probability, i.e., the proportion of repeating the same goal choice as in the previous trial, was calculated as a function of actual reward magnitude and/or fictive reward magnitude in order to assess the effects of actual and fictive reward magnitudes on the animal's choice in the next trial.

Logistic regression analysis Effects of the animal's past choices and their outcomes (up to 3 trials) on the animal's choice in the current trial were examined using the following logistic regression model^{37–39}:

$$\log\left(\frac{p_L(i)}{p_R(i)}\right) = \sum_{j=1}^3 \gamma_j^R (R_L(i-j) - R_R(i-j)) - \sum_{j=1}^3 \gamma_j^F (F_L(i-j) - F_R(i-j)) + \gamma^C (C_L(i-1) - C_R(i-1)) + \gamma^S S + \gamma_0, \quad (1)$$

where $p_L(i)$ (or $p_R(i)$) is the probability of selecting the left (or right) goal in the i -th trial. The variable $R_L(i)$ (or $R_R(i)$) is the magnitude of actual reward at the left (or right) goal, $F_L(i)$ (or $F_R(i)$) is the magnitude of fictive reward at the left (or right) goal, $C_L(i-1)$ (or $C_R(i-1)$) is the left (or right) goal choice (0 or 1) in the $(i-1)$ -th trial, and S is the location of the start box (0 or 1). The coefficients γ_j^R , γ_j^F , γ^C and γ^S denote the effects of past actual rewards, past fictive rewards, the previous choice and start box location, respectively, and γ_0 is the bias to choose the left goal (choice bias).

Learning models of choice behavior The probability of selecting the left goal in the $(i+1)$ -th trial was fit to the following model:

$$p_L(i+1) = \frac{e^{A_Q(i)}}{[1 + e^{A_Q(i)}]}, \quad (2)$$

if $C_L(i) = 1$ and $C_R(i) = 0$

$$\Delta Q(i) = \beta [Q_L(i) - Q_R(i)] + [k_{\text{actual}} R_L(i) - k_{\text{fictive}} F_R(i)] + b_{\text{st}} + b_L,$$

else

$$\Delta Q(i) = \beta [Q_L(i) - Q_R(i)] + [k_{\text{fictive}} F_L(i) - k_{\text{actual}} R_R(i)] - b_{\text{st}} + b_L, \quad (3)$$

where b_{st} and b_L are biases to stay (stay bias) and to select the left goal (choice bias), respectively, β is the inverse temperature of the soft-max action selection rule, $Q_L(i)$ and $Q_R(i)$ are action value functions of the left and right goal, respectively, in the i -th trial, and k_{actual} and k_{fictive} are constants. The action value functions were updated as the following:

if $C_L(i) = 1$ and $C_R(i) = 0$

$$\begin{aligned} Q_L(i) &= Q_L(i-1) + \alpha_{\text{actual}} \cdot (R_L(i) - Q_L(i-1)) + V_q/\beta, \\ Q_R(i) &= Q_R(i-1) + \alpha_{\text{fictive}} \cdot (F_R(i) - Q_R(i-1)), \end{aligned} \quad (4)$$

else

$$\begin{aligned} Q_L(i) &= Q_L(i-1) + \alpha_{\text{fictive}} \cdot (F_L(i) - Q_L(i-1)), \\ Q_R(i) &= Q_R(i-1) + \alpha_{\text{actual}} \cdot (R_R(i) - Q_R(i-1)) + V_q/\beta, \end{aligned} \quad (5)$$

where α_{actual} and α_{fictive} are the learning rates for actual and fictive rewards, respectively, and V_q is a constant increment for the value of chosen goal. In reduced models, subsets of terms were excluded from the full model.



Statistical test. Statistical significance of a regression coefficient was determined with a *t*-test (two-tailed). A binomial test was used to determine statistical significance of the proportion of left-choice or stay trials, and a χ^2 -test was used to test whether the proportion of stay trials significantly deviated from an even distribution across different magnitudes of actual or fictive reward in the previous trial. A *p* value <0.05 was used as the criterion for a significant statistical difference. All data are expressed as mean \pm SEM.

- Byrne, R. M. Mental models and counterfactual thoughts about what might have been. *Trends Cogn. Sci.* **6**, 426–431 (2002).
- Epstude, K. & Roese, N. J. The functional theory of counterfactual thinking. *Pers. Soc. Psychol. Rev.* **12**, 168–192 (2008).
- Hayden, B. Y., Pearson, J. M. & Platt, M. L. Fictive reward signals in the anterior cingulate cortex. *Science* **324**, 948–950 (2009).
- Abe, H. & Lee, D. Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron* **70**, 731–741 (2011).
- Steiner, A. P. & Redish, A. D. Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nature Neurosci.* **17**, 995–1002 (2014).
- Dayan, P. & Niv, Y. Reinforcement learning: the good, the bad and the ugly. *Curr. Opin. Neurobiol.* **18**, 185–196 (2008).
- Niv, Y. & Montague, P. R. Theoretical and empirical studies of learning. In *Neuroeconomics: Decision Making and Brain*. Glimcher, P. W., Camerer, C. F., Fehr, E. & Poldrack, R. A., eds, pp.331–351 (Academic Press, New York, 2009).
- Lee, D., Seo, H. & Jung, M. W. Neural basis of reinforcement learning and decision making. *Annu. Rev. Neurosci.* **35**, 287–308 (2012).
- Worthy, D. A. & Todd Maddox, W. A comparison model of reinforcement-learning and win-stay-lose-shift decision-making processes: A tribute to W.K. Estes. *J. Math. Psychol.* **59**, 41–49 (2014).
- Rutledge, R. B. *et al.* Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *J. Neurosci.* **29**, 15104–15114 (2009).
- Beeler, J. A., Daw, N., Frazier, C. R. & Zhuang, X. Tonic dopamine modulates exploitation of reward learning. *Front. Behav. Neurosci.* **4**, 170 (2010).
- Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed. (Springer, New York, 2002).
- Fischer, A. G. & Ullsperger, M. Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron* **79**, 1243–1255 (2013).
- Thorndike, E. L. Animal intelligence: An experimental study of the associative processes in animals. *Psychol. Rev. Monograph Suppl.* **2**, i–109 (1898).
- Tolman, E. C. & Honzik, C. H. Introduction and removal of reward, and maze performance in rats. *Univ. California Pub. Psychol.* **4**, 257–275 (1930).
- Tolman, E. C., Ritchie, B. F. & Kalish, D. Studies in spatial learning. I. Orientation and the short-cut. *J. Exp. Psychol.* **36**, 13–24 (1946).
- Johnson, A. & Redish, A. D. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J. Neurosci.* **27**, 12176–12189 (2007).
- van der Meer, M. A. & Redish, A. D. Covert expectation-of-reward in rat ventral striatum at decision points. *Front. Integr. Neurosci.* **3**, 1–15 (2009).
- Steiner, A. P. & Redish, A. D. The road not taken: neural correlates of decision making in orbitofrontal cortex. *Front. Neurosci.* **6**, 131 (2012).
- Tolman, E. C. The determinants of behavior at a choice point. *Psychol. Rev.* **45**, 1–41 (1938).
- Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, 1998).
- Lohrenz, T., McCabe, K., Camerer, C. F. & Montague, P. R. Neural signature of fictive learning signals in a sequential investment task. *Proc. Natl Acad. Sci. USA* **104**, 9493–9498 (2007).
- Chiu, P. H., Lohrenz, T. M. & Montague, P. R. Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. *Nat. Neurosci.* **11**, 514–520 (2008).
- Boorman, E. D., Behrens, T. E. & Rushworth, M. F. Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS Biol.* **9**, e1001093 (2011).
- D'Ardenne, K., Lohrenz, T., Bartley, K. A. & Montague, P. R. Computational heterogeneity in the human mesencephalic dopamine system. *Cogn Affect Behav Neurosci.* **13**, 747–756 (2013).
- Skinner, B. F. "Superstition" in the pigeon. *J. Exp. Psychol.* **38**, 168–172 (1948).
- Wright, J. C. Consistency and complexity of response sequences as a function of schedules of noncontingent reward. *J. Exp. Psychol.* **63**, 601–609 (1962).
- Davis, H. & Hubbard, J. An analysis of superstitious behavior in the rat. *Behaviour* **43**, 1–12 (1972).
- Matute, H. Human reactions to uncontrollable outcomes: Further evidence for superstitions rather than helplessness. *Quart. J. Exp. Psychol.* **48B**, 142–157 (1995).
- Lee, S. W., Shimojo, S. & O'Doherty, J. P. Neural computations underlying arbitration between model-based and model-free learning. *Neuron* **81**, 687–699 (2014).
- Goldman-Rakic, P. S. Architecture of the prefrontal cortex and the central executive. *Ann. N. Y. Acad. Sci.* **769**, 71–83 (1995).
- Fuster, J. *The Prefrontal Cortex*, 4th ed. (Academic Press, Amsterdam, 2008).
- Kolb, B. Animal models for human PFC-related disorders. *Prog. Brain Res.* **85**, 501–519 (1990).
- Uyilings, H. B., Groenewegen, H. J. & Kolb, B. Do rats have a prefrontal cortex? *Behav. Brain Res.* **146**, 3–17 (2003).
- Vertes, R. P. Interactions among the medial prefrontal cortex, hippocampus and midline thalamus in emotional and cognitive processing in the rat. *Neurosci.* **142**, 1–20 (2006).
- Aggarwal, M., Hyland, B. I. & Wickens, J. R. Neural control of dopamine neurotransmission: implications for reinforcement learning. *Eur. J. Neurosci.* **35**, 1115–1123 (2012).
- Lau, B. & Glimcher, P. W. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* **84**, 555–579 (2005).
- Huh, N., Jo, S., Kim, H., Sul, J. H. & Jung, M. W. Model-based reinforcement learning under concurrent schedules of reinforcement in rodents. *Learn. Mem.* **16**, 315–323 (2009).
- Kim, H., Sul, J. H., Huh, N., Lee, D. & Jung, M. W. Role of striatum in updating values of chosen actions. *J. Neurosci.* **29**, 14701–14712 (2009).

Acknowledgments

This work was supported by the Research Center Program of the Institute for Basic Science (IBS-R002-G1; M.W.J.) and NIH Grant R01 DA029330 (D.L.).

Author contributions

K.-U.K., D.L. and M.W.J. designed the study, K.-U.K. performed behavioral experiments, K.-U.K., N.H., Y.J., D.L. and M.W.J. analyzed the data, M.W.J. wrote the manuscript with inputs from other authors, and all authors reviewed the manuscript.

Additional information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Kim, K.-U., Huh, N., Jang, Y., Lee, D. & Jung, M.W. Effects of fictive reward on rat's choice behavior. *Sci. Rep.* **5**, 8040; DOI:10.1038/srep08040 (2015).



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>