論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Choice beyond probability : behavioral anomalies induced by utility of action
著者(和文)	恩蔵絢子
Author(English)	Ayako Onzo
出典(和文)	学位:博士（学術）, 学位授与機関:東京工業大学, 報告番号:甲第6944号, 授与年月日:2007年3月26日, 学位の種別:課程博士, 審査員:
Citation(English)	Degree:Doctor of Philosophy, Conferring organization: Tokyo Institute of Technology, Report number:甲第6944号, Conferred date:2007/3/26, Degree Type:Course doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Choice beyond Probability:

Behavioral Anomalies Induced by Utility of Action

Department of Computational Intelligence and Systems Science

Ayako Onzo

Abstract

Animals, including humans, have to make a decision with limited environmental information. Animal learning theories assume that they behave so as to maximize reward they receive [1, 2]. This basic feature of learning is generally assumed to be relevant also in human learning. However, in the modern world, the nature of reward has become increasingly complex, highlighting the significance of uncertainty. A recent study found that uncertainty itself is rewarding for animals if it is presented in an appropriate context [3]. In behavioral economics, which deals with the actual human's behavior in the presence of uncertainty, evidences have accumulated that people do not necessarily behave in a way predicted simple objective measures of reward [4, 5]. To explain the deviations from simple reward maximization, a variety of theories on human's decision making have been proposed. In expected utility theory and prospect theory, the personal preference of uncertainty and other factors are reflected in the value function of reward [6-9]. These theories provide powerful insights into choice behavior, where they handle the preference of humans with statistical methods. Here I investigate the humans' internal dynamics of decision making using a simple gambling game in which any optimal behavior cannot be defined ("flat reward" condition) (Study 1). The results suggest that even in the flat reward condition, the subjects do not employ a random betting strategy. The rich internal dynamics loosely coupled with the external reward structure reveals the underlying principles involved in choice. Further sets of experiments show the difference of behavior in games where the objective reward structures are identical in terms of all the statistical measures, induced by the subjects' perception of agency (Study 2 & 3). The previous theories in behavioral economics assume that the

statistically identical options have the same values. The present study suggests that the uncertainty involved in a certain probability can be heterogeneous, being affected by the subject's agency structure. From these results, I will discuss open-endedness in human's learning.

Acknowledgements

First of all, I would like to deeply appreciate my supervisor, Dr. Kenichiro Mogi, for his advice and support in various situations. He gave me a great deal of invaluable experience. I could also have second opinions from his best friends, Prof. Takashi Ikegami at Tokyo University and Prof. Yukio Gunji at Kobe University. In discussion with these people, I could learn the most important things in science and life.

I also thank all the members of Mogi Lab., Dr. Qi Zhang, Dr. Fumihiko Taya, Mr. Hisayuki Nagashima, Mr. Kei Omata, Mr. Toru Yanagawa, Mr. Takayasu Sekine, Ms. Tamami Sudo, Ms. Fumiko Tanabe, Mr. Eiichi Hoshino, Mr. Tomomitsu Herai, Ms. Fumi Okubo, Mr. Tetsuo Ishikawa, and Mr. Shinichi Nozawa, for their help.

Finally, I would like to express my gratitude for the helpful comments of Prof. Kiyohiko Nakamura, Prof. Eizo Miyashita, Prof. Masayuki Yamamura, and Prof. Shigenobu Kobayashi, who kindly refereed this thesis.

Contents

Introduction

This thesis consists of three main studies on human's decision making under uncertainty using simple gambling games and related discussions. First, evidence is shown that people's internal dynamics play an important role in the condition where a single optimum behavior cannot be defined (Study 1). Secondly, a study is conducted on people's behavioral anomaly in betting games where the rewards for the available choices are statistically identical. The differential betting tendencies in these conditions are likely to be induced by the sense of "agency", the perception of being involved in an action [10] (Study 2). Finally, an aspect of cognitive mechanisms involved in the robust handling of uncertainty is revealed through an experiment on the role of "commitment [11]" (Study3).

The basic idea behind these experiments has been that humans are not just a machine making stimulus-response associations. In real life situations, we are often obliged to decide without knowing the correct answer in the presence of uncertainty. The human brain is equipped with mechanisms with which working functionality is supported even in the absence of sufficient information. For example, illusory perception can make up for the lack of complete information as is reported from several studies of vision (e.g. Kanizsa triangle [12]). Within the same general principle, some aspects of human decision-making process are based on the prediction of the future outcome through self-induce voluntary processes which do not necessarily depend on externally provided information. Furthermore, such cognitive processes are likely to come to terms with "onceness", the concept which describes real life situations in which we experience significant events once and for all, making it difficult to apply statistical learning principles even when one is implemented. For example, we do not always feel the same when seeing an identical object, and are obliged to make our own interpretation of the external world every time we encounter novel and sometimes even familiar phenomena. Therefore, onceness in life's events puts a constraint on and induces the spontaneity in cognitive processes of decision making.

Although there are anecdotal evidences that onceness is important in human cognitive development, it is difficult to study scientifically, since standard statistical treatments often fail to capture its essence, with regard to "repeatability" in particular. A good strategy to give a scientific grounding to the problem of onceness is to link it with learning process. Every choice that a human being makes reflects, and facilitates in turn, the process of learning. It is important in this respect to investigate the change in choice under the same condition. In this thesis, I will address the issue of onceness and open-endedness reflected in a series of choice by investigating the nature of human’s decision making paying attention to its aspects that develop independent of the statistically treatable measures of reward.

Chapter 1: Decision making under uncertainty (Related works)

1.1 Reward functions in animal learning theory

Animals, including humans, encounter novel stimuli in the course of life, incurring cognitive uncertainty. How animals coordinate their actions in such an uncertain environment is one of the crucial questions to be asked of cognition.

Reward is one of the most important influences shaping behavior. In the early 20th century, Pavlov proposed that the function of reward is producing a change in behavior, also called learning [13]. A dog salivates to a bell only after the sound has been paired with a food, but not to a different, non-paired sound, suggesting that its behavioral response (salivation) has changed after food conditioning. Such an associative learning between a neutral stimulus ("Conditioned Stimulus", CS) and a reward ("Unconditioned Stimulus", US) is called Pavlovian learning. Around the same time, Thorndike's Law of Effect postulated that a reward increases the frequency and intensity of a specific action that has resulted in a reward before [14]. The association between an action and a reward is named Instrumental conditioning, distinguished from Pavlovian learning in that it deals with the modification of voluntary behavior through the use of consequences.

More recently, Schultz distinguished three function of reward [1]. First, rewards induce learning, as they make an agent come back for more (positive reinforcement); second, they induce approach and consuming behavior for acquiring the reward object; and third, they induce positive emotions. Rewards serve as goals of behavior if the reward and contingency between action and reward are represented in the brain during the action. By contrast, punishers induce avoidance learning, withdrawal behavior and negative emotions.

Based on these psychological observations, a computational approach to learning, Reinforcement learning [15], has been developed. In the reinforcement learning, an agent interacting with the environment learns by trial-and-error in order to maximize reward in the long run. The agent basically reinforces actions that have led previously to desirable outcomes. However, an essential problem arises here, which is of the balance between exploitation and exploration. To keep exploiting only certain sources of reward is tantamount to dismissing opportunities to explore alternative sources of reward, and might work unfavorably for maximizing the cumulative rewards. Although avoiding uncertainty is not necessarily adaptive, exploring uncertainty sometimes poses a danger of death or serious injury. Then, how do animals overcome uncertainty? What determine the balance between exploration and exploration?

In the developmental process, the psychological "secure base" provided by caretakers is considered to be a necessary basis for infant's voluntary exploration of novel stimuli [16, 17]. Perception of "secure base" as a basis for exploration is likely to be relevant also in mature humans. I will investigate how mature humans explore the uncertain environment in Study 1 (Chapter 2).

1.2 Reward system in the brain

Food, water and sexual stimuli are called primary rewards as they reinforce behavior without being learned. The reward mechanisms involved in such stimuli are mostly innate, for they are essential for survival and reproduction. Other stimuli, such as cultural goods or money, are called secondary rewards as they reinforce behavior only after learning. They acquire reward value through the association with primary rewards.

In 1954, Olds and Milner discovered that rats lever-press at high rates to obtain brief stimulation pulses to certain brain regions [18] (Fig. 1. 1). Subsequent studies revealed that electrical stimulation of the dopaminergic pathway which has its origin in the ventral tegmental area (VTA) A10 is critical for the behavior. Rats often choose self-stimulation over other rewards such as food or water and press the lever repeatedly until they are exhausted. Dopamine receptor blockers such as the antipsychotic drug haloperidol reduce the rewarding effect of food and intracranial self-stimulation, while hedonic drugs such as cocaine or amphetamine increase the level of dopamine release at the projection terminals of the VTA and facilitate reinforcing behavior (addiction) [19-21]. Thus, dopamine is regarded as a mediator of reward in the brain.

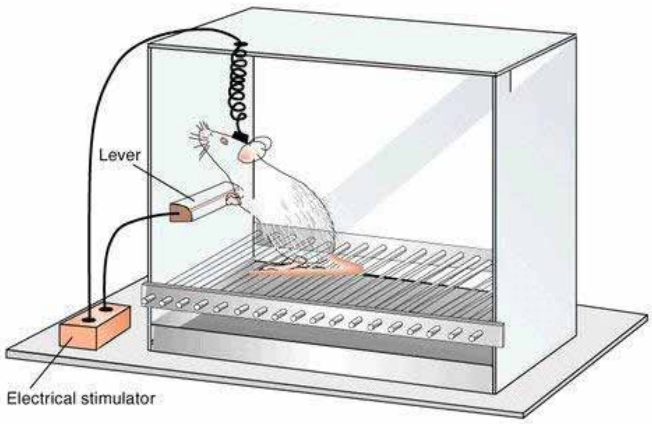


Fig. 1. 1 Self-stimulation in rats.
[Extracted from http://www.cerebromente.org.br/n18/history/stimulation_i.htm]

In humans, the mesolimbic-mesocortical dopamine system is known as the reward system in the brain. Similarly to rats, the system originates in the VTA (A10) and projects to the mesial components of the limbic system, such as the ventral striatum including the nucleus accumbens (NAcc), the amygdala and the mesial frontal cortex, and the prefrontal cortex (Fig. 1. 2). In particular, the ventral striatum, the amygdala, and the orbitofrontal cortex in these projection sites are considered to be important for reward processing [22, 23]

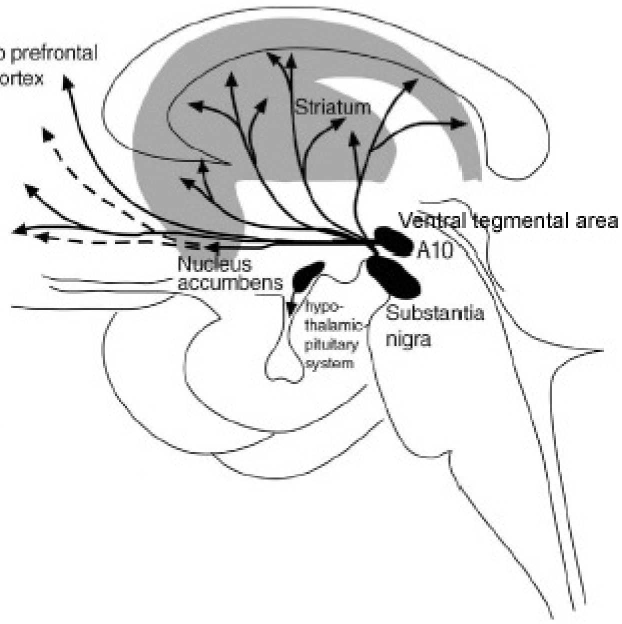


Fig. 1. 2 Human's reward system, the mesolimbic-mesocortical dopamine pathways (modified from [23]).

Schultz and his colleagues revealed the properties of reward processing in the brain from the experiments in which activity of single dopamine neurons was recorded in monkeys while they received rewards [24]. When a monkey is presented unexpectedly with various appetitive stimuli (e.g. fruit juice), dopamine neurons respond with short phasic burst of activity (Fig. 1. 3. a). Aversive stimuli like air puffs to the hand or drops of saline to the mouth do not cause these transient activations. Thus, dopamine neurons are activated only by unpredictable stimuli that elicit reward. Furthermore, after repeated pairings of visual and auditory cues (CS) followed with a fixed interval by reward (US), dopamine neurons change the time of their phasic activation from just after the time of reward delivery to the time of cue onset (Fig. 1. 3. b). The changes in dopamine neurons activity strongly resemble the transfer of the animal's appetitive behavioral reaction (e.g. licking) from the US to the CS. Since the CS predicts reward after learning, the CS becomes appetitive. While the US is no longer unpredictable, the timing of CS delivery is uncertain. In trials where the reward is not delivered at the appropriate time after the CS, the firing rate of dopamine neurons decreases below the base line at the exact time that the reward should have occurred (Fig. 1. 3. c). These results suggest that the firings of dopamine neurons appear to code for an error between the actual reward and the predictions of the timing and magnitude of reward (i.e. the relative "goodness" of environmental events to learned predictions about those). Therefore, learning is considered to be driven by the prediction-error of reward.

Fiorillo and his co-workers found an additional property of dopamine neurons [3]. They suggested in the following experiment that uncertainty itself has rewarding or reinforcing properties if it is presented in an appropriate context. Monkeys were conditioned in a Pavlovian procedure with distinct visual stimuli indicating the probability ($p = 0$, 0.25, 0.5, 0.75, and 1.0) of liquid reward being delivered after a 2-s delay. This new parameter was introduced as a measure of uncertainty. Uncertainty is maximal at $p = 0.5$, but absent at the two extremes ($p = 0$ and 1). The phasic activities of dopamine neurons varied monotonically with reward probability (Fig. 1. 4. a). While the magnitude of the phasic activations at the time of reward delivery increased as probability decreased, conditioned stimuli elicited the phasic activations with their magnitude increasing with increasing reward probability. These results are consistent with the prediction error hypothesis. However, a previously unreported activation of dopamine neurons is also discovered in this study.

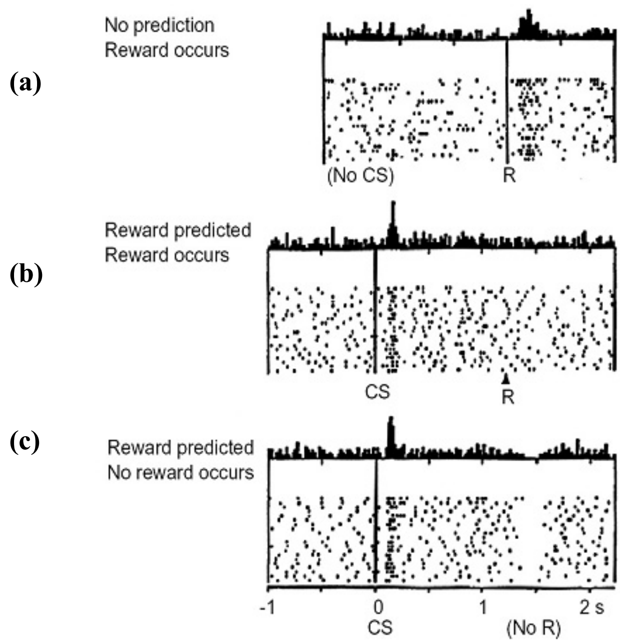
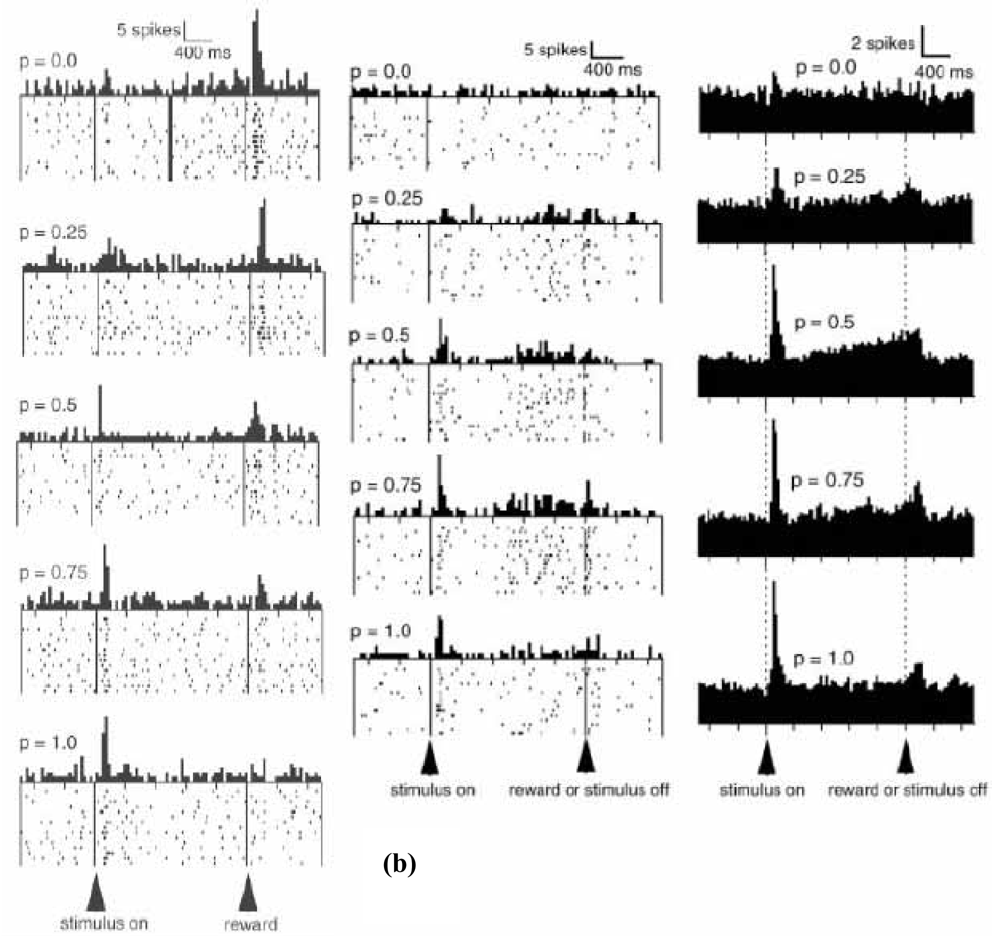


Fig. 1. 3 Reward-prediction error coding in a single dopamine neuron (extracted from [24]). **(a)** Before learning, the dopamine neuron fires at the occurrence of reward. **(b)** After learning, the dopamine neuron is activated by the reward-predicting stimulus (the conditioned stimulus) but fails to be activated by the predicted reward. **(c)** After learning, the conditioned stimulus predicts a reward, but the reward fails to occur because of a mistake in the behavioral response in the monkey. The activity of the dopamine neuron is depressed exactly at the time when the reward would have occurred.

There was a sustained increase in activity that grew from the onset of the conditioned stimulus to the expected time of reward (Fig. 1. 4. b). The sustained activation was maximal at p=0.5, less pronounced at p=0.25 and 0.75 and absent at p=0 and 1, co-varied with uncertainty. Furthermore, the peak of the sustained activation occurred at the time of potential reward, which corresponded to the moment of greatest uncertainty. The coding of uncertainty in dopamine neurons indicates that uncertainty could be the secondary reward for animals in this procedure. This argument gives an explanation for gambling addiction.

As we saw in exploration vs. exploitation dilemma, it is necessary to consider not only actual rewards but also unpredictability or uncertainty in order to explain animals' behaviors. In the next section, I will review the studies about how uncertainty affects our behavior in the real life situations.



(a)

Fig. 1. 4 Dopamine neurons' responses at various reward probabilities ranging from 0.0 (top) to 1.0 (bottom) in Pavlovian procedure (extracted from [3]). Bin width=20 ms. **(a)** Phasic activation. Rasters and histograms of activity in a single cell illustrate responses to the conditioned stimuli and reward. Only rewarded trials are shown. **(b)** Sustained activation before potential reward at all three intermediate probabilities. Both rewarded and unrewarded trials are shown. (Left) Rasters and histograms of activity in a single cell. (Right) Population histograms.

1.3 Behavioral economics

To make uncertainty tractable, Pascal formed Probability theory around 1650. He considered that it would be rational to choose an option with the highest value in all options after weighing the potential outcomes by their associated probabilities, that is, to make a decision according to expected values (Expected value theory) [25, 26]. But nearly one hundred years later, Bernoulli reported an odd paradox, named St. Petersburg paradox [25, 26]. Which of the following would you prefer, $40 or a lottery ticket that pays according to the outcomes of one or more fair coin tosses: heads you get $2 and the game ends, tails you get another toss and the game repeats, but now if the second toss lands heads up you get $4, and so on. If the *n*th toss is the first to land heads up, you get 2^n dollars. The game continues until the coin lands heads up. The expected value of this lottery ticket is infinite. A rational person should prefer the ticket with the infinite expected value. But actually, people tend to choose the sure option of $40. Bernoulli reasoned this was because humans are rationally prudent, and thus averse to taking the risks associated with the lottery [6].

In behavioral economics in which psychology and economics are combined, actual human's decision making in the presence of uncertainty has been investigated. People faced with decision under uncertainty do not necessarily behave in a way predicted by simple objective measures of reward, exhibiting various anomalies. A number of attempts, including Expected utility theory, Prospect theory, and other models in behavioral economics have been developed to account for the observed discrepancies [6-9].

1.3.1 Expected utility theory

People's tendency of risk aversion is explained by Expected utility theory [6, 7] where it is assumed that people do not make decisions based on the expected values (EV) of choices, but on the expected utilities (EU). “Utility” refers to subjective desirability of reward [27]. EU is calculated as

$$\mathrm{EU} = \sum_{\mathrm{i}} (p_i \cdot u(x_i));\ i = 1,\ n;\ n = \text{number of reward}$$

where p_i and $u(x_i)$ represent the probability and utility of a gain x_i, respectively, corresponding to EV. The utility of a gain does not increase linearly but frequently follows a concave function (Fig. 1. 5). For example, the utility of a cup of juice is not constant at all the time. The utility is larger when we are thirsty than when we are not. It is called “Diminishing marginal utility” that utility from one additional unit is inversely related to the number of units already owned.

Here, consider the following lotteries:

A: a sure gain of 5 units

B: 50 % chance of winning 1 and 9 units

C: 50 % chance of winning 4 and 6 units

where EV is 5 units in each lottery. EU(B) is considerably lower than EU(A) when using a concave utility function (Fig. 1. 5), reflecting people’s tendency of risk-aversion. Furthermore, even though the probability of winning is also the same between Lottery B and C, EU(C) is higher than EU(B) (Fig. 1. 5), indicating the smaller risk in Lottery C because of a smaller range of reward magnitudes.

The concave utility function is often modeled as $u(x) = -e^{-bx}$, where b is a constant. Then, EU results in

$$\mathrm{EU} = \sum_i (p_i \cdot (-e^{-bx_i})),$$

which can be developed by the Laplace transform into

$$\mathrm{EU} = -e^{-b(\mathrm{EV} - b/2 \cdot \mathrm{Var})},$$

where Var is variance, and the probability distribution p_i is Gaussian. Thus, EU can be expressed as $f(\mathrm{EV}, \mathrm{Var})$. Namely, the deviation from Expected value theory such as risk-aversion is explained by the differences in statistical measures of higher order such as variance in Expected utility theory.

However, Expected utility theory is also violated. The famous violations are Ellsberg's paradox [28] and Framing effect [29, 30] (Section 1. 3. 2). Other theories have attempted to modify Expected utility theory and provided important insights into choice behavior (Section 1. 3. 3), but not yet become a global theory of choice that can truly replace Expected utility theory.

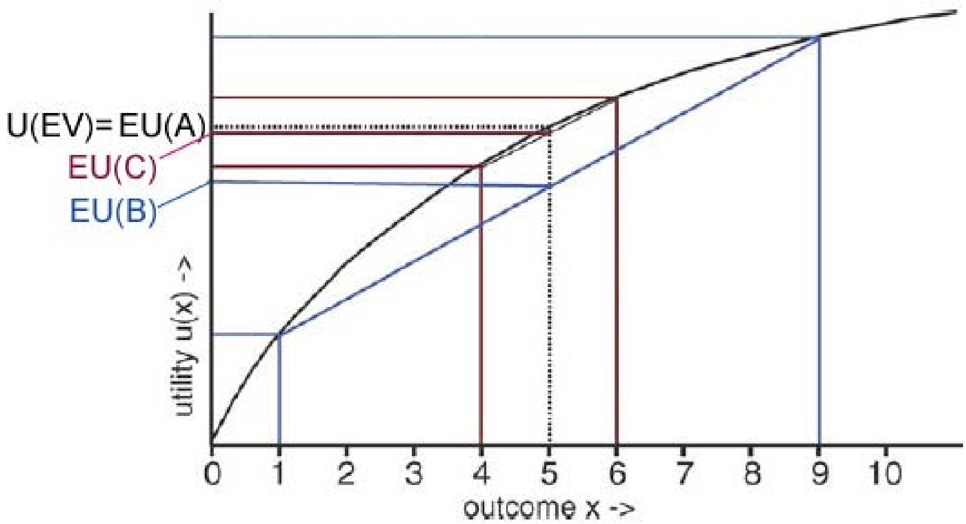


Fig. 1. 5 A hypothetical utility function (modified from [2]). The utility of outcome does not increase linearly but frequently follows a concave function. The utility from one additional unit (marginal utility) varies inversely with the number of units already owned (Diminishing marginal utility). The risk aversion can be reflected in this function: Even though EV (expected value) is 5 in each lottery, EU (expected utility) differs among the lotteries (EU(A)>EU(C)>EU(B)).

1.3.2 Violations of Expected utility theory

(1) Ambiguity aversion

In Ellsberg's paradox (Fig. 1. 6), people are presented with an urn that contains 90 balls. Of these, 30 are blue and 60 are red or yellow; any proportion is possible. They are offered a choice between red and blue (Choice 1). If they draw a ball with the color they chose, they can get $100. Here, while the probability of a blue draw is defined ($p=1/3$), the probability of a red draw is ambiguous: It is a choice between an event with a known probability and an event with unknown probability. In this case, people typically choose blue. Next, they are asked to choose between blue/yellow and red/yellow (Choice 2). If either color in the pair they chose is drawn, they can get $100. Under this condition, the likelihood of winning is clear in the red/yellow pair ($p=2/3$), but ambiguous in the blue/yellow pair ($1/3 \le p \le 1$). People in choice 2 typically choose the red/yellow pair. According to Expected utility theory, when x represents the number of red balls, people should expect $x<30$ in Choice 1, and $30<x$ (because $30+(60-x)<60$) in Choice 2. This is a contradiction. This paradox suggests that when a probability of an event is unknown or ambiguous, Expected utility theory fails, and that people tend to dislike ambiguity.

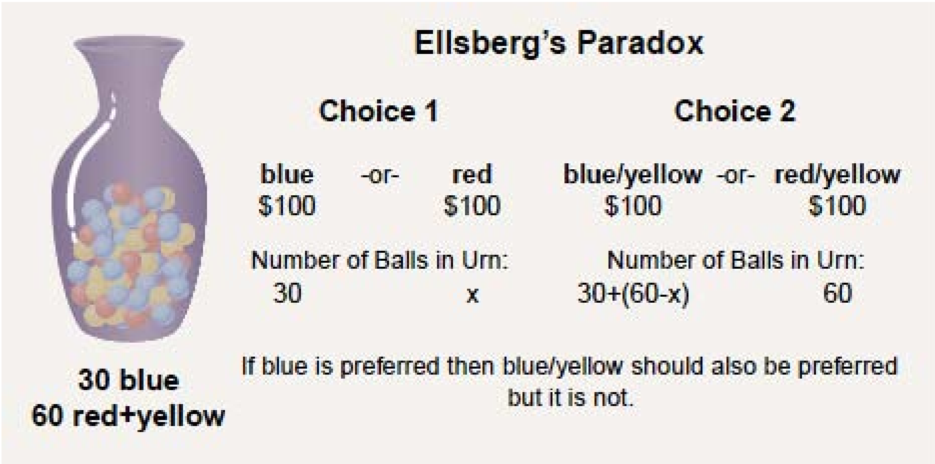
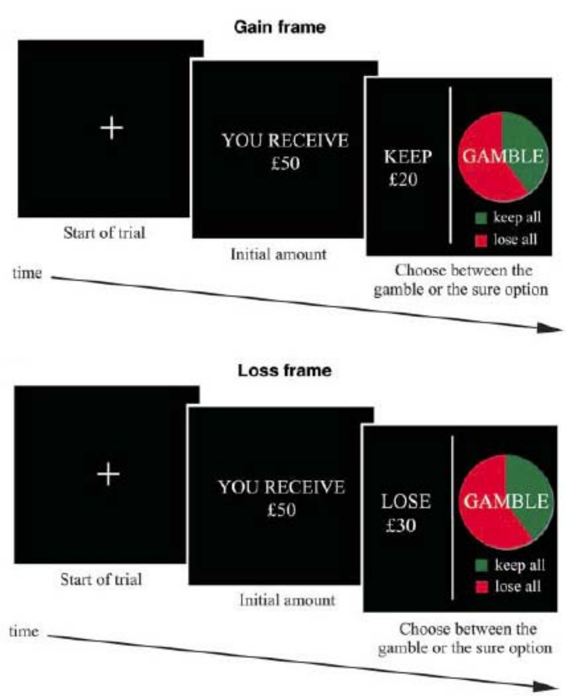


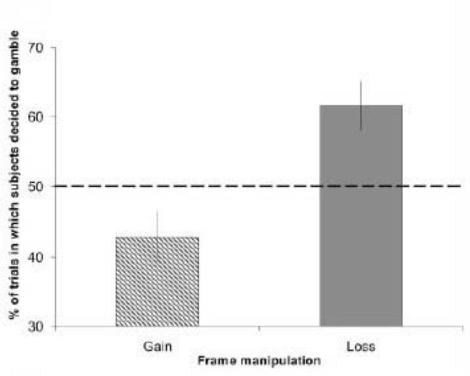
Fig. 1. 6 Ellsberg's paradox (extracted from [26]). If blue is preferred in Choice 1, then blue/yellow should logically be preferred in Choice 2. But people tend not to choose ambiguous options.

(2) Framing effect

Expected utility theory is based on the implicit assumption that people's preferences are invariant under different representations ("frames") of equivalent choice problems. However, many studies have demonstrated systematic reversals of preference when the same problem is presented in different ways, which is termed the "framing effect" [29, 30].

Consider the following. You initially receive £50. You then have to choose between a "sure" option and a "gamble" option presented in the context of two different frames (Fig. 1. 7. a). The "sure" option is formulated as either the amount of money retained from the initial starting amount (e.g., keep £20 of the £50; "Gain" frame) or as the amount of money lost from the initial amount (e.g., lose £30 of the £50; "Loss" frame). The "gamble" option is identical in both frames. The expected outcomes of the gamble and sure options are equivalent. Therefore the objective reward conditions are the same in both frames (i.e. the expected value is £20 whichever option they choose in either frame). However, people's decisions are significantly affected by the framing manipulation (Fig. 1. 7. b). They are risk-averse in the Gain frame, tending to choose the sure option over the gamble option, and risk-seeking in the Loss frame, preferring the gamble option.





(b)

(a)

Fig. 1. 7 Framing effect (extracted from [30]). **(a)** Experimental settings. Mathematically equivalent options (Sure and Gamble) were presented in two different frames (Gain/Loss frame). **(b)** Percentage of trials in which subjects chose the gamble option in the Gain frame and the Loss frames. Subjects showed a significant increase of the rate in the Loss frame with respect to the Gain frame.

1.3.3 Prospect theory

Kahneman and Tversky proposed Prospect theory [8, 31] as an alternative to Expected utility theory. This theory consists of two main functions: (1) the value function, which corresponds to the utility function in Expected utility theory, and (2) the weighting function.

(1) The value function

There are three important features in the value function: (a) reference dependence, the idea that value is judged in terms of gains and losses from a multidimensional reference point, (b) loss aversion, the idea that when the same amounts of gain and loss are given, the loss is estimated steeper than the gain, and (c) diminishing sensitivity, the notion that the marginal value of both gains and losses diminishes with their size (Fig. 1. 8). This value function can account for the framing effect (i.e. the frames can be given by the reference dependence, and the diminishing sensitivity represents our different risk preference between the gain and loss frame, because the function becomes a concave in the gain frame and a convex in the loss frame.)

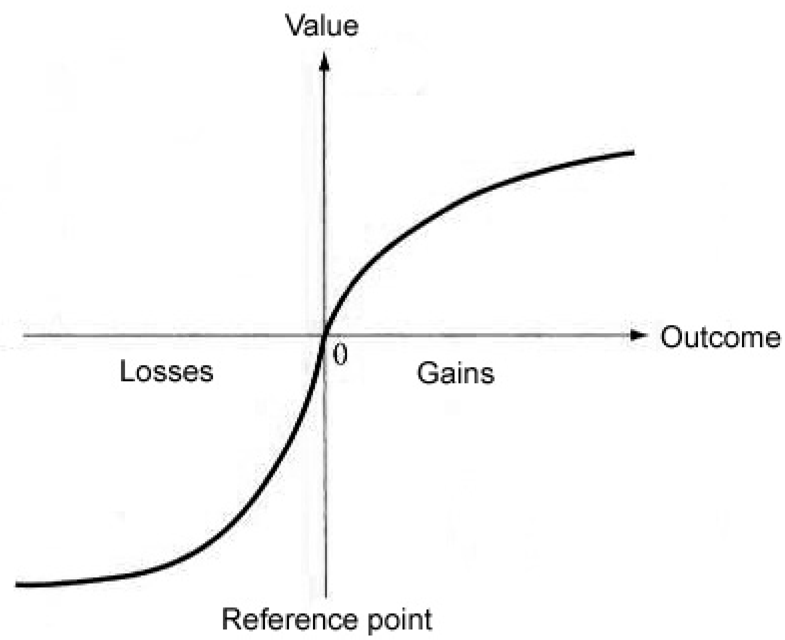


Fig. 1. 8 Value function. Prospect theory focuses on changes from a reference point in outcome. The difference in risk preference between the Gain and the Loss frames and the loss aversion are expressed in this function.

(2) The weighting function

In Prospect theory, the value of each outcome is multiplied by a decision weight. In other words, probabilities of events are replaced by decision weights. People tend to overestimate unlikely events (e.g. an airline accident) but underestimate likely events (e.g. a car accident). Decision weights are generally lower than the corresponding probabilities, except in the range of low probabilities (Fig. 1. 9).

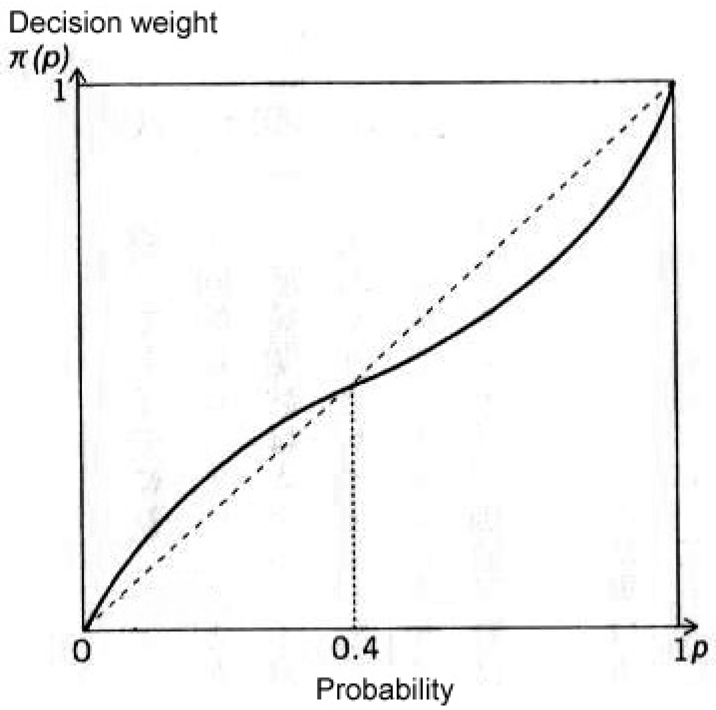


Fig. 1. 9 Weighting function. People tend to overestimate unlikely events but underestimate likely events.

We do not behave in an objectively rational way. However, biasing or framing can be rather adaptive because we have to make a decision with limited or complex environmental information. On the neural basis of biasing or framing, I will overview the studies of neuroeconomics in the next section.

The theories presented in this section describe very "static" feature of human's preference. Two remained questions are (1) on the dynamics of decision making because once assigned subjective values may not be fixed over time, and (2) whether the completely equivalent options in terms of statistical measures of reward and of presented contexts have the same subjective values or not. In the real world, is there a substantial difference between ambiguity and risk? It is possible that various kinds of uncertainty are involved even in a certain probability. I will address these issues in my studies, (1) in Study 1 and (2) in Study 2 and 3.

1.4 Neuroeconomics

To reveal the underlying neural mechanisms of choice, neuroeconomics in which behavioral economics and neuroscience are combined has been established recently.

Lesion studies made it clear that logical inferences would not always lead to good decisions. Patients with damage to the orbitofrontal cortex (OFC) have an impaired ability to make appropriate decisions under some conditions (e.g. planning their work and choosing friends), and show lack of emotion, even though their intellectual abilities are normal (e.g. logical thinking, attention and memory performance) [32, 33]. These observations have led Damasio to propose the "somatic marker hypothesis" that the emotional system mediates decision biases [32]. Consistently with this hypothesis, the activations of the emotional system including the OFC and the amygdala are necessary in choices with ambiguity as in Ellsberg's paradox [34], and are also associated with the framing effect [30].

These areas receive dopaminergic inputs from the VTA (Fig. 1. 2). Coupled with the function of reward, they are likely to contribute to make a subjective value of reward (e.g. utility). Social communication, which is one of the most important features in human's activity, inevitably involves uncertainty. The role of emotion and reward systems in social communication has been illustrated in studies using simple interpersonal games such as the Ultimatum game, the Prisoner's Dilemma game or the Trust game [35-37]. In the Ultimatum Game, two players split a sum of money. One player proposes a division and the other can accept or reject this offer. If it is accepted, the money is split as proposed, but if the responder rejects the offer, then neither player receives anything. The rational solution to this game is for the proposer to offer the smallest sum of money possible to the responder and for the responder to accept this offer because any monetary amount is preferable to none. However, considerable behavioral researches indicate that responders reject unfair offers. Participants with stronger activation of the anterior insula which has been implicated in the evaluation and representation of negative emotional states, to unfair offers reject those at higher rates [35]. Furthermore, punishments for players who behave unfairly in these interpersonal games activate the reward-related regions (the striatum), correlated with an expressed desire for revenge [36, 37], suggesting that the reward system not only processes reinforcing stimuli that serve basic survival needs, but also is more generally involved in the evaluation of stimuli with social interaction.

Chapter 2: Decision making without optimal behavior (Study 1)

Previous studies have reported deviations from rational behaviors. Here I examine human's betting behavior in a very simple gambling game where any optimum or rational behavior cannot be defined. The probability and magnitude of reward are designed to give a "zero" expected net reward ("flat reward condition" (Fig. 2. 1)). No matter how the subject behaves, there is on average no change in one's resources, and therefore every possible sequence of action has the same value. Investigating people's betting behavior is one of the simple and effective tools for uncovering the nature of the cognition of uncertainty. In most studies, the game is designed in such a way that a change in betting behavior leads to a change in the expected gain, reflecting real life situations. On the other hand, it is possible that people's betting behavior exhibit internal dynamics of its own, in dissociation with externally defined reward structure. The flat reward condition may reveal the nature of internal neural dynamics involved in people's judgment under uncertainty, which is loosely coupled with externally defined reward. As a result, even in the flat reward condition, the subjects are found to behave not in a random manner but in ways showing characteristic tendencies, reflecting the rich and complex dynamics of brain's reward system.

2.1 Methods

I conducted a betting game that consisted of 20 trials. The participants were 12 healthy young adults (ages 22-29, 6M & 3F). The subjects were undergraduate and graduate students in several universities in Tokyo area, and possessed basic knowledge of probability. The subjects were aware that the outcomes of betting would be determined by a simulated random process in a computer, and that there was no way to influence the outcome by any system of betting. The subjects were instructed about the general conditions involved in the game beforehand and gave informed consent. The subjects started with an initial resource of 5 units, and tried to increase the amount as best as they could. The buttons and letters were displayed on a laptop computer (Sony VAIO PCG-505G/B).

In each trial, the subject had a choice of either betting or escaping, by pushing the "Bet" or "Escape" button that became active 5 seconds after pushing the "Next" button (Fig. 2. 2). The trial number, probability of winning (=0.25, always the same), and the current resource were displayed. For the very first turn, the “Bet” and “Escape” buttons became active 5 seconds after launching the game software. There was no explicit time pressure, and the subjects were allowed to use as much time as they liked when making the decisions. The delay from the moment when the "Bet" and "Escape" buttons became available to the moment when the subjects made the choice was recorded. When the subject pushed the "Bet" button, 1 unit was taken from the subject's resource. The probability of winning was fixed at 0.25. If the subject won, 4 units were added to the resource. The expected net gain when betting was therefore zero ($EV = 0.25 \times 4 + 0.75 \times 0 - 1 = 0$). When the subject pushed the "Escape" button, the resource remained the same. Thus, there was no change in the expected value of gain no matter which choice ("Bet" or "Escape") the subject made (flat reward condition, Fig. 2. 1). From the probabilistic point of view, the subject had no rational motivation to bet or escape with a particular strategy, and any deviations from a random betting pattern can only be explained in terms of cognitive bias and/or illusion, with no actual contribution to the net gain.

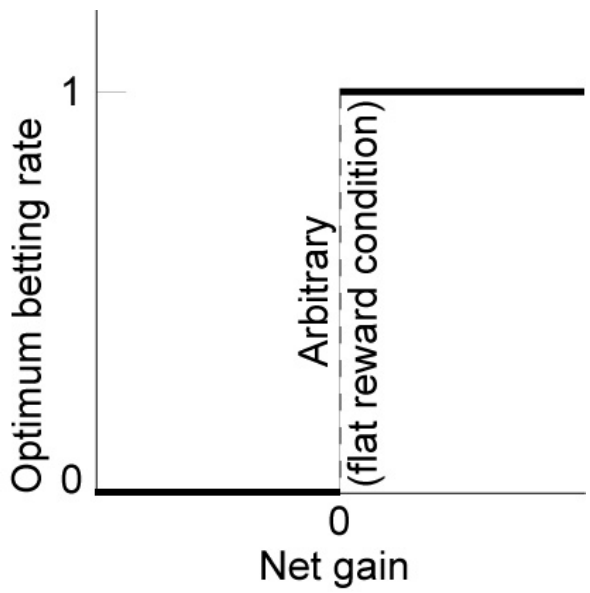


Fig. 2. 1 Flat reward condition. If the expected value of a gambling game is positive, the optimum behavior can be defined as betting, and if negative, not betting. Only when the expected net gain is zero, the optimum betting rate becomes arbitrary (flat reward condition).

The feedback (“You Win!” “You Lose!” or “You Escape!”) was displayed immediately after the subject pushed the "Bet" or "Escape" button. The screen remained the same until the subject pushed the "Next" button. The game was over when the resource became 0 or the trial number reached 20. Each subject repeated 30 games. When the game was over, the resource was set back to the initial value of 5, and the gains and losses were not carried over to the next session.

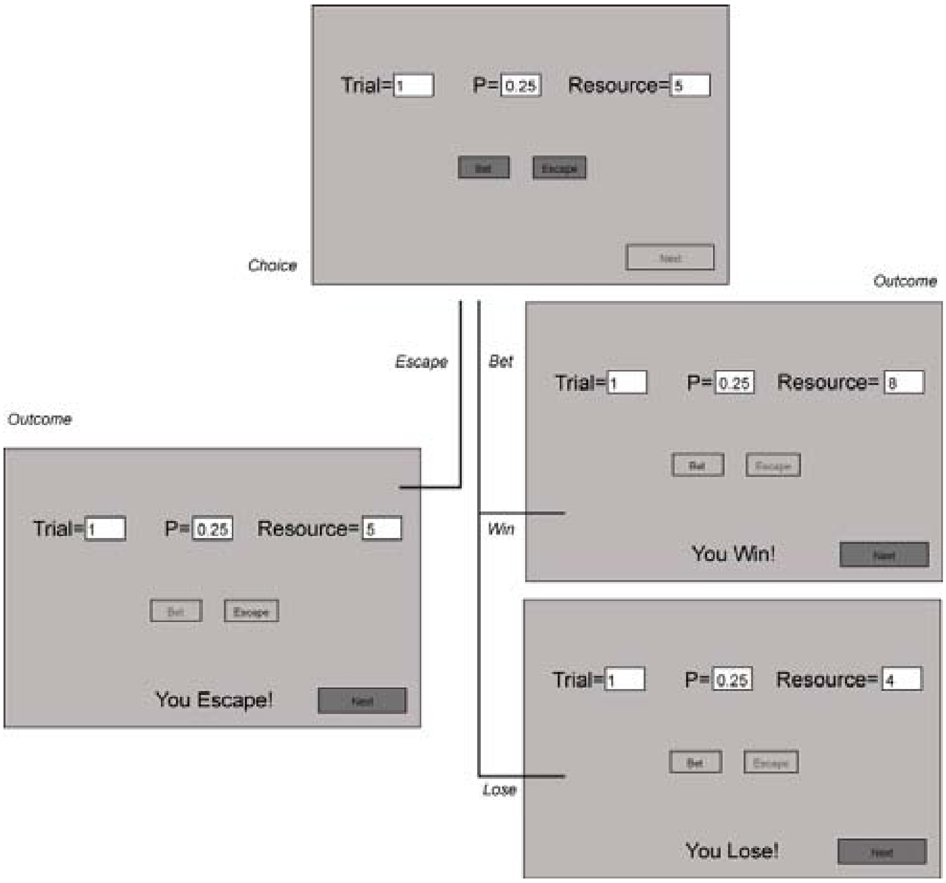


Fig. 2. 2 Experimental protocol.

2.2 Results

The behavior of the subjects can be characterized by the probability of the subjects to bet (betting rate), N_{bet}/N_{trial}, where N_{bet} and N_{trial} are the number of betting and trials, respectively. The betting rates were obtained under various criteria. From the probabilistic point of view, there is no rationale for the subjects to bet using any specific strategy. Any tendency away from the random betting behavior would suggest the existence of (unconscious or conscious) strategy that the subjects employed, in which the internal dynamics of the brain's reward and decision-associated areas is reflected.

The mean betting rate ± s.d. of the subjects was 0.81±0.18 and significantly above the chance level (=0.5) ($p<0.05$, Chi-square test), suggesting that the subjects do not behave in a random manner even in the flat reward condition. The higher betting rate over the low winning probability (=0.25) is consistent with the prospect theory in that people overestimate unlikely events [8] (Section 1. 3).

However, they do not always keep betting throughout a game. Although the external reward structure is the same every trial, the subjects choose to bet at some trials and to escape at others. What makes the subjects choose the different options under the fixed external conditions? The following factors are likely involved: (1) the perception of “secure base [17] (Section 1. 1)” such as the remaining trials or resource, (2) the emotional memory of the previous outcomes (Lose, Win or Escape), and (3) the perception of winning probability.

Fig. 2. 3 shows the relation between the amount of resource and betting rate. The rate is flat for amount of resource up to ~16 (about a threefold increase from the initial resource), and then again stays at a higher fixed level for larger amounts of resource. As there are considerable individual differences, the average as well as individual data are plotted in this and following figures. The subject numbers are used consistently throughout. Since the game is over when the resource becomes zero, it is psychologically natural to bet more frequently when more resource is accumulated through winning ("secure base" effect). Note that there is a slight increase of betting rate when the resource is smaller than the initial resource, which cannot be explained by the secure base effect. The subjects may have taken risks in order to compensate for the loss from the initial resource.

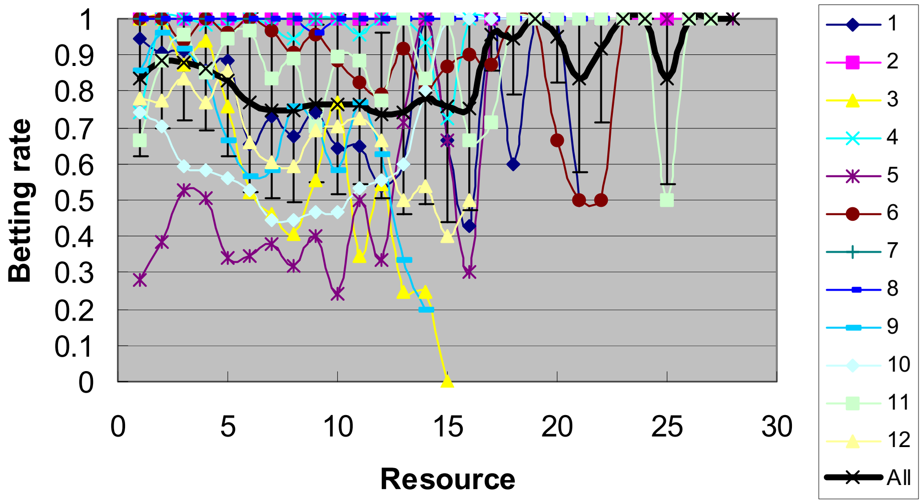


Fig. 2. 3 Resource and betting rate. The average betting rate (black line) is flat for amount of resource up to ~16 and then again stays at a higher fixed level for larger amounts of resource ("secure base" effect) (n=12).

The trial number could also influence the betting rate. A subject could, for example, take a greater risk towards the end of the 20 trials set within a game. Note that the subjects were aware of the trial number, as it is displayed clearly on the game screen (see Methods). Fig. 2. 4 shows the relation between the trial number and betting rate. There is no apparent tendency for the subjects to bet more frequently towards the end of the session. The absence of the effect of trial number is consistent throughout the subjects, although the betting rate differs among subjects.

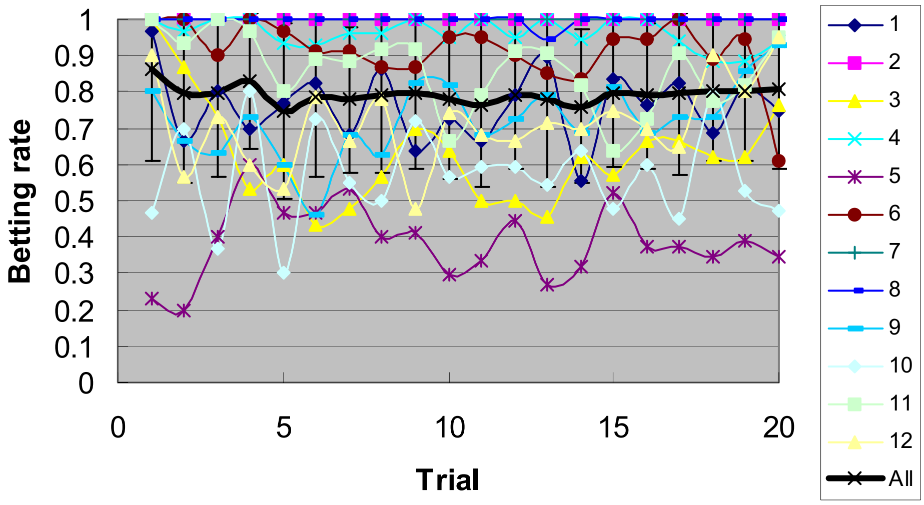


Fig. 2. 4 Trials number and betting behavior. There is no apparent effect of trial (n=12).

Since the probability of winning in a trial is independent of past outcomes, there can be no rational explanation for the dependence of betting behavior, if any, on the previous outcomes. Fig. 2. 5 shows the dependence of betting rate on the immediately previous outcome (Lose, Win, or Escape). There is a statistically significant tendency to bet more frequently after losing, than winning ($p<0.05$, paired-t test). Note the considerable individual differences in the betting rate.

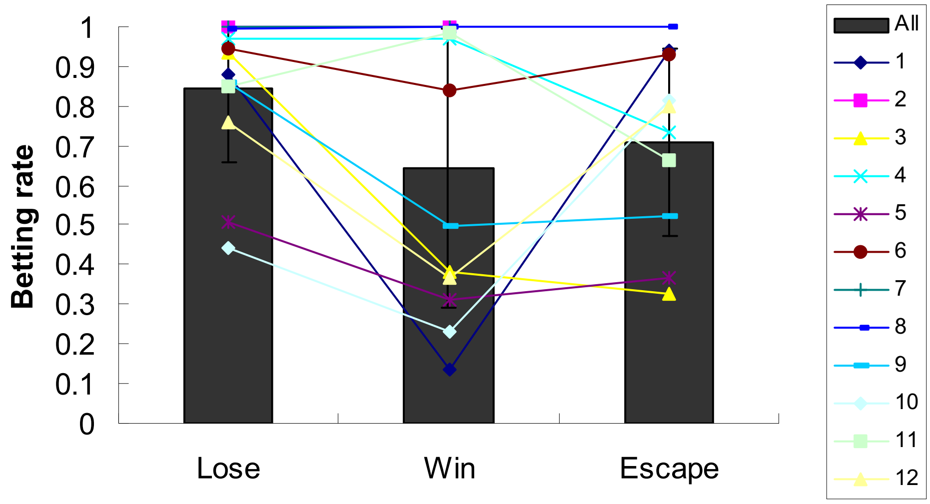


Fig. 2. 5 Previous outcome and betting behavior. The average betting rate after “Lose” is significantly higher than after “Win” (P<0.05, paired t-test) (n=12).

When averaged over the subjects, there was no significant difference in the reaction time no matter which choice ("Bet" or "Escape") the subject made. In addition, no significant dependence of the reaction time on the previous outcome ("Escape", "Lose", or "Win") was found (Fig. 2. 6). This statistical insignificance, however, is due to the large standard deviation in the reaction time. The individual plots for average reaction time for making the "Bet" or "Escape" choice indicate that the subjects do actually take longer when choosing to bet than to escape (Fig. 2. 7). The longer delay before deciding to "bet" likely reflects the nature of cortical processing involved. Note that two subjects (number 2 and 7) opted to keep betting and never escaped (often resulting in the termination of the game before 20 trials). Therefore, their data do not appear in Fig. 2. 6.

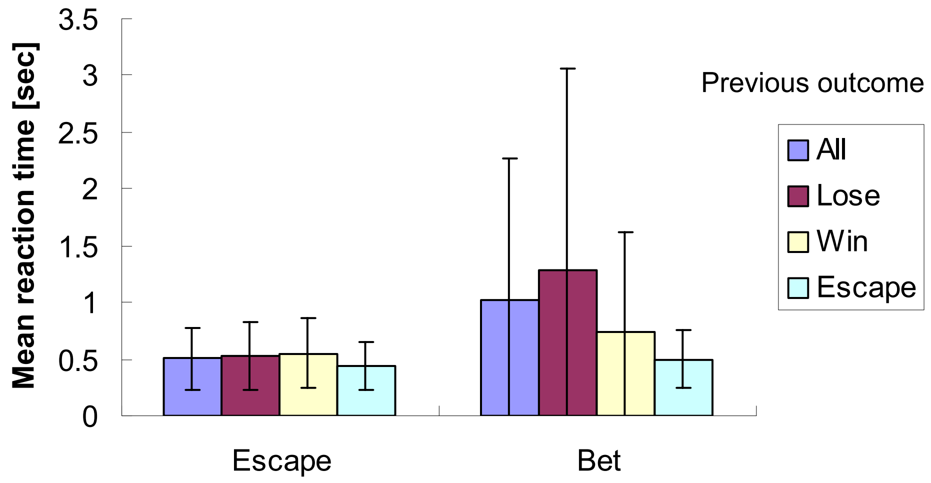


Fig. 2. 6 Betting behavior and reaction time. On average, there is no significant difference in reaction time between "Bet" and "Escape" and no significant dependence on the previous result ("Escape", "Lose", or "Win") (n=12).

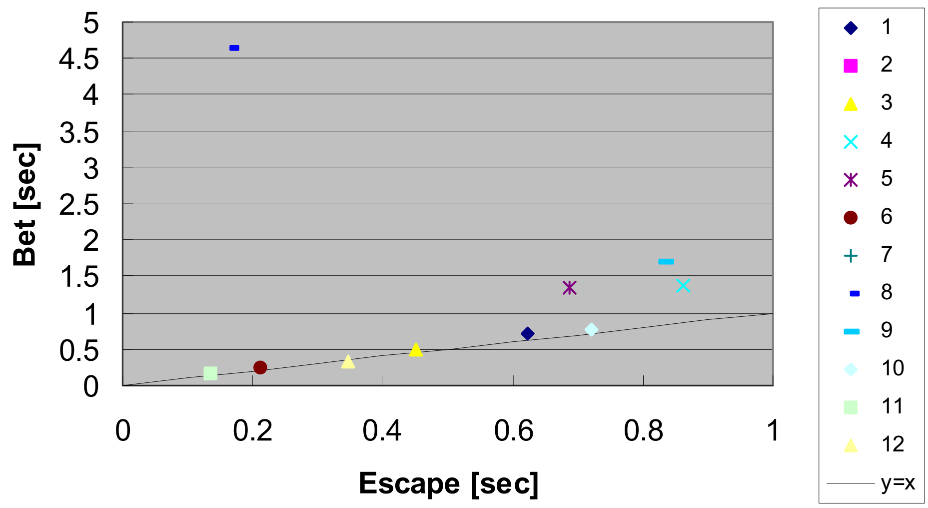


Fig. 2. 7 Reaction time of individual subjects for "Bet" and "Escape" choices. Although there is no difference in reaction time on average (Fig. 2. 6), the individual plot reveals that the subjects take longer when choosing to bet than to escape.

Data analysis so far suggests that there are considerable individual differences in betting behavior. There are important aspects of betting dynamics that are not apparent when averaged. The dynamics of betting behavior might be heterogeneous even in an individual. It seemed possible that the betting behavior as is indicated in Fig. 2. 5, which can be approximated as a Markov process, is actually composed of heterogeneous modes of betting dynamics, with the subjects behaving in distinct manners away from the random behavior.

Since there are 3^k possible outcomes for k consecutive trials, it is difficult to analyze all possible trajectories of betting behavior. After some preliminary analysis, I focused on a particular mode of betting behavior, where the subject keeps betting for several trials without escaping ("betting streak").

Fig. 2. 8 shows the "cornering effect" in a betting streak. Here, the betting rate is plotted against the number of consecutive choices of betting already made. The betting rate keeps increasing, away from the random behavior where the betting rate should stay at a constant level. When a subject keeps betting, it appears that he or she is "cornered" into a state where the betting rate continues to increase, resource permitting.

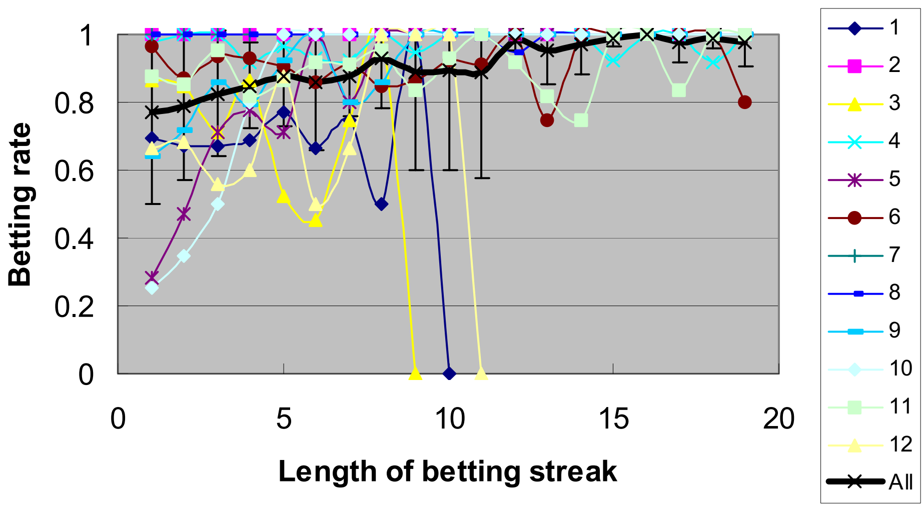


Fig. 2. 8 Betting streak. The average betting rate (black line) increases with the length of betting streak (“cornering effect”) (n=12).

It is possible that a betting streak is driven by consecutive winnings or losses. In consecutive winnings, the subjects might "feel good" and keep betting. In consecutive losses, they might be motivated to make up for the loss. I analyzed the betting rate in "winning streaks" (where the subject keeps betting and winning) and "losing streaks" (where the subject keeps betting and losing). Note that winning streaks and losing streaks are subsets of betting streaks. Comparison between the data suggests that the outcome does not have significant effect on the subject's behavior in a betting streak (Fig. 2. 9). The subject seems to be determined to keep betting, regardless of the result.

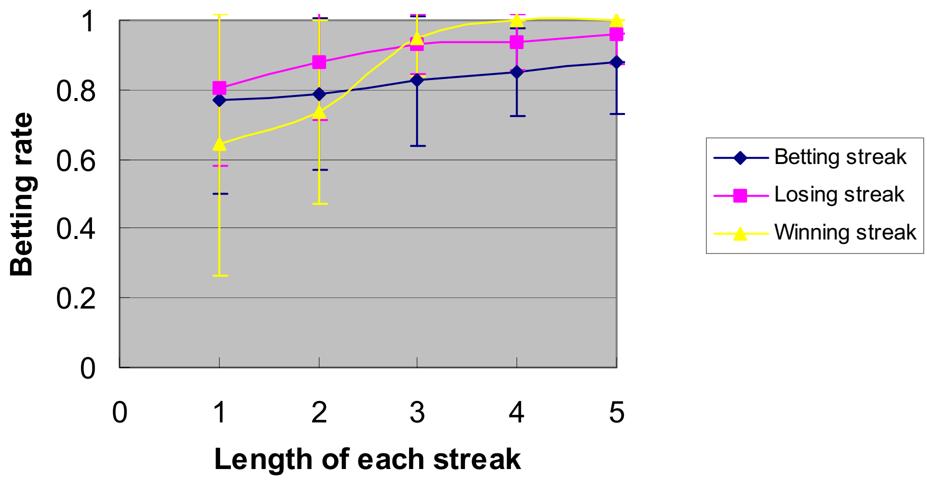


Fig. 2. 9 Betting rate in betting, losing and winning streaks. There is no significant difference between the losing streak and the winning streak, suggesting the outcome does not have significant effect on the subject’s behavior in a betting streak.

Although no statistically significant difference in betting rate was found in the comparison between the three kinds of streaks, there indeed was a small tendency that a betting streak is induced by consecutive winnings. Fig. 2. 10 provides evidence that supports the hypothesis that a betting streak is induced by winning repeatedly. When I calculate the cumulative winning rate (i.e., the average winning rate within a betting streak), it becomes significantly higher as the length of the betting streak is increased. This result suggests that at least some betting streaks are induced by an event where the subject had the fortune to win above the chance level of 0.25. Note that when averaged over all trials, the betting rate is higher after losing, rather than winning (Fig. 2. 5). Thus, the behavior of the subjects in a betting streak is a phenomenon separate from the average tendency.

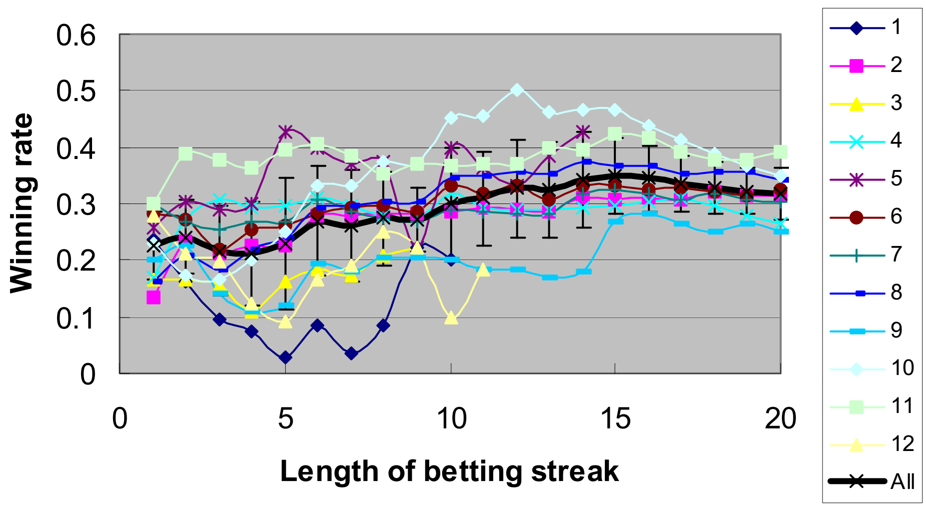


Fig. 2. 10 Cumulative winning rate in a betting streak. The average winning rate within a betting streak (black line) becomes significantly higher as the length of the betting streak is increased.

2.3 Discussions

The robust handling of uncertainty is important for survival. It is interesting to ask how we recognize and judge uncertainty, and choose appropriate actions. Subjective bias and illusions are integral part of the brain's cognitive process of uncertainty handling, and merits investigation apart from the context of external reward optimization

A large variety of studies in economics and neuroeconomics have taken subjective factors that influence a subject's decision making (Chapter 1). It is important to study the cognition of uncertainty in the full richness of its dynamics, since our behavior is embedded in the constantly changing situations in daily life. In my experiment reported here, the subjects exhibited characteristic betting dynamics in a simply designed game. Since the expected gain is constant regardless of the subject's behavior, differential behavior cannot be explained on the basis of reward optimization. The differential behavior of the subject can only arise from the brain's internal dynamics, reflecting neural mechanism for evaluating reward and making decisions in an uncertain environment. Informal interviews with the subjects after the experiment suggested that they were in general unaware of the fact that they were behaving differentially depending on the previous outcome. Therefore, the betting dynamics is likely to reflect unconscious tendencies, rather than an application of conscious strategies.

Several factors seem to affect the betting behavior. The perception of the resource as secure base, memory of recent results, and perception of the probability of winning are some elements affecting the betting behavior observed. The complex neural computation involving the representation of these factors, finally culminating in a winner-take-all type decision-making, is likely to be an integral part of the robust handling of uncertainty. The betting dynamics under the flat reward condition is neutral to reward, and the evolution of the neural dynamics might be understood in a context neutral to the immediate reward [38]. At least some aspects of betting dynamics (e.g. higher betting rate after losing compared to winning) cannot be explained by the immediate optimization of gain. Needless to say, these cognitive processes need to contribute to, or at lease be compatible with, the efficient utilization of external reward in the environment. The enrichment of neural dynamics in a context neutral to expected gain might contribute to the final fitness. The exact logic behind such a development of neutral dynamics, however, needs to be clarified in future investigations.

Finally, there was a considerable individual difference of betting behavior. As already mentioned, two subjects chose to keep betting and never escape. Apart from these extreme cases, there is a wide variety of betting patterns among the subjects. Anecdotal evidence suggest that such heterogeneity of strategy is typically observed in gaming under the presence of uncertainty, and might reflect a general tendency of the neural system involved in the robust handling of uncertain situations. In particular, in situations involving interaction with another agents, such as the ultimatum game [35] (Section 1. 4), and prisoner's dilemma [39], heterogeneity of strategies might induce rich interpersonal dynamics, increasing the complexity of social interaction and contributing to the overall fitness of the group of people involved.

Chapter 3: Behavioral anomaly induced by the utility of action (Study 2)

Here I report a behavioral anomaly induced by the intrinsic utility of action, which is influenced by the way betting opportunities are presented in space and time.

Rewards such as food and water are important factors in the evolution and learning of animal behavior. In the modern world, the nature of uncertainty faced by the average person has become increasingly complex. Choices and actions are driven by the brain's reward system, the understanding of which poses an interesting empirical and theoretical problem. Abstract rewards involving eye contact [40], interpersonal interaction [35-37, 41], and money [42] play important roles in people's behavior. In economics, "utility" is the measure of happiness or satisfaction gained from a reward [27], where choices associated with a larger value of utility are given preference in the decision making process.

Evidences have accumulated that humans do not necessarily behave in a way predicted by objective reward, especially in the presence of uncertainty. Bernoulli, in his classical study of the St. Petersburg paradox [6], showed how the effect of variances can even override that of the average value, resulting in the choice of a lower expected gain. When given a choice in lottery, subjects can prefer one to the other even when the expected gain is the same [5]. People faced with uncertainty tend to avoid ambiguity [28], overestimate losses [8, 31], and are influenced by the context within which the choices are presented [29, 30], resulting in various anomalies in behavior. Uncovering novel anomalies are important in pinning down essential factors to be incorporated into the utility function used to explain what people choose.

There are a number of models in behavioral economics [4, 5] and other fields aimed at explaining human behavior in the presence of uncertainty. Expected utility theory [6, 7] attempts to explain seemingly anomalous human behavior involving risk. The prospect theory [8] describes how people evaluate gains and losses. In general, even when the average values for rewards are the same, differences in statistical measures of higher order such as variance can cause humans to give priorities to a particular choice. Recently, the robust handling of uncertainty in the reward system of the human brain has been the subject of the burgeoning field of neuroeconomics [25, 26].

When the nature of decision-making process is considered within the context of brain function, factors other than classical utilities emerge as important. The perception of being able to choose one's own action and influence the outcome ("agency" [10, 44-46]) is an important element in decision making. Time is a precious resource, so that prolonged wait time is an effective punishment in a task [47]. Less wait time results in more actions per session, increasing the sense of agency. When the number of actions taken is the same, the quality of actions becomes important in agency perception. It is thus interesting to investigate how the sense of agency and the nature of wait time affect human decision in the presence of uncertainty.

Here I examined the behavior of human subjects using three conditions of simple gambling games, where the probability of winning and the magnitude of the reward are statistically identical ($p=0.25$) (Fig. 3. 1). In each condition, the subjects made a choice of either to bet or to escape. When the subject decided to bet, 1 unit was taken from the resource. Winning the bet was rewarded with 4 units. Thus the expected net gain (the expected value of reward minus the cost for betting) was 0 whether the subject chose to bet or not. This "flat reward" condition (Fig. 2. 1, Chapter 2) was introduced to make the nominal gain to be independent of the specific choices the subjects made, revealing other factors contributing to utility.

In the "bet button" (B) condition, the subjects made a bet by clicking on the “Bet” button displayed on the computer screen. In the "one box" (O) condition, a box was presented in place of the “Bet” button. The subject chose to bet by clicking on the box, followed by the same random process as in condition B. In the "four boxes" (F) condition, four identical boxes were displayed in a horizontal alignment on the screen, one of which was guaranteed to be a "win". The winning box was changed in a random manner. The subject made a bet by clicking on one of the boxes or escaped. The "Escape" button was common for all conditions.

Note that in all three conditions, the winning probabilities are mathematically identical, in terms of the average expected value as well as any higher order statistical measures such as variance. Thus, theories based on the statistical properties of the nominal reward [4-9, 25, 26] would predict that the subject's behavior would be the same for all conditions. Any differences in the subject's betting behavior need to be accounted for by a utility function involving factors other than the nominal reward.

3.1 Methods

The participants were undergraduate and graduate students from universities in the greater Tokyo area (ages 22-29, 10 males & 2 females). Pre-trial interviews ensured that they possessed a basic knowledge of probability theory, being aware that there was no real chance of winning above the statistical average no matter what "betting system" they used. The subjects were instructed about the general conditions involved in the game beforehand and gave informed consent. The subjects were provided with an initial resource of 5 units and tried to increase it as best they could. In each game, the subjects were asked to make a choice whether or not to bet having been informed of the probability of winning (p=0.25), until the resource became zero. One game consisted of a maximum of 20 trials. The game was over when the resource became 0 or the trial number reached 20. When the game was over, the resource was set back to the initial value of 5, and the gains and losses were not carried over to the next session. The subjects played 30 games each under the "bet button" (B), “one box” (O) and “four boxes” (F) conditions (Fig. 3. 1). The numbers of trials and initial resource were designed to motivate the subjects to bet with some care, in order not to run out of resource. The average numbers of trials actually played were 15.7 ± 1.05, 16.2 ± 1.75, and 14.8 ± 1.17 for conditions B, O, and F, respectively. The subjects on average were able to play the maximum 20 trials in 0.611 ± 0.088, 0.622 ± 0.173, and 0.55 ± 0.098 of total games for conditions B, O, and F, respectively. The betting and feedback screens were displayed on a laptop computer (Sony VAIO PCG-505G/B). The buttons and boxes became active 5 seconds after clicking the "Next" button. The trial number, probability of winning (=0.25 for all three conditions), and the current resource were displayed at the time of choice. For the very first turn, the buttons became active 5 seconds after launching the game software. The feedback (“Win” “Lose” or “Escape”) was displayed immediately after the subject made the choice. In the four boxes condition, there was no feedback of the "correct" box in the case of "Lose". After the result was presented, the screen did not change until the subject pushed the "Next" button.

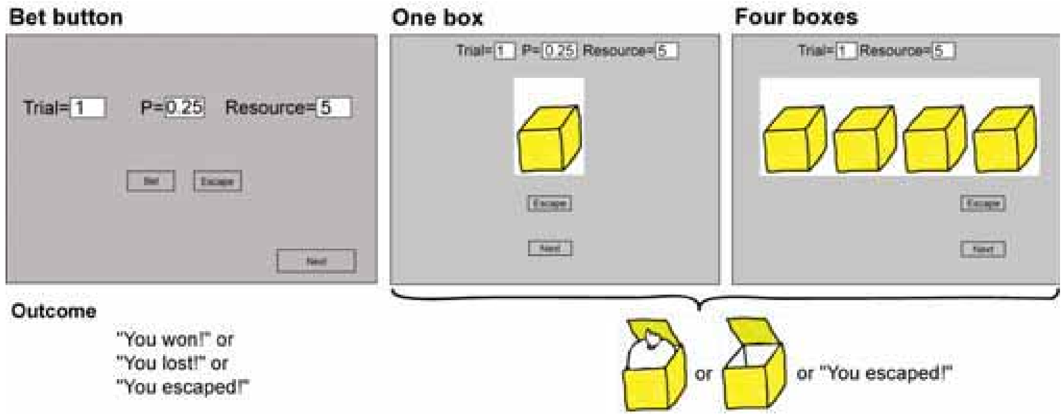


Fig. 3. 1 Experimental set up. In the "bet button" (B) condition, trial numbers, probability of winning (always 0.25), and the remaining resource were presented above the bet and escape buttons. In the "One box" (O) condition, a yellow box was presented instead of the bet button. In the "four boxes" (F) condition, four yellow boxes were presented. Probability of winning was not explicitly shown in the F condition. In all conditions, the subjects clicked the "next" button to go to the next trial.

3.2 Results

The average betting rates (mean ± s.d.) for the three betting conditions were 0.812 ± 0.183, 0.723 ± 0.240, 0.996 ± 0.005 for the bet button, one box, and four boxes conditions, respectively (Fig. 3. 2). One-Way Repeated Measures ANOVA found a significant difference among these betting rates ($F(2,22)=7.89$, $P=0.003<0.005$). Post hoc Tukey HSD test found a significant difference between the betting rates for conditions B and F ($P=0.040<0.05$) and conditions O and F ($P=0.002<0.005$). There was no significant difference between conditions B and O ($P= 0.423$), consistent with the interpretation that the specific visual appearance of the interface does not have a crucial influence on the betting behavior possibly driven by an abstract reward structure [48].

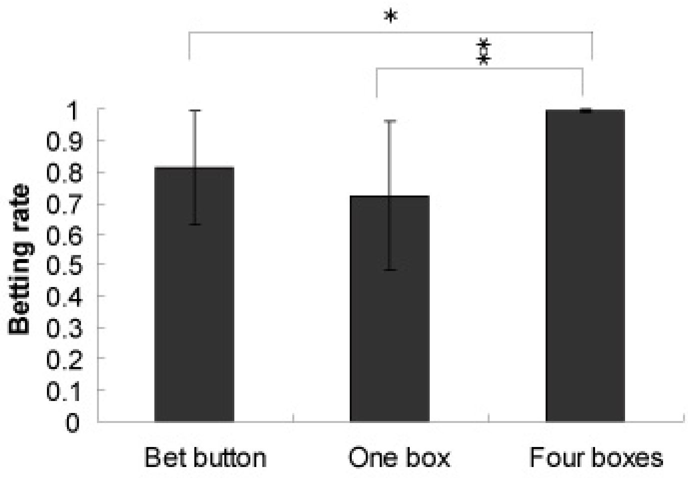


Fig. 3. 2 Average betting rates (± s.d.). The single and double asterisks represent P<0.05 and P<0.005 in the Tukey HSD test, respectively.

The subject's behavior could be influenced by a number of elements other than the difference in betting conditions as described above. For example, there could be an effect induced by neighboring trials, leading to a failure of the assumption that each trial can be treated independently. To countercheck this possibility, I conducted yet another series of trials in a "randomized" (R) series where one of the three conditions (B, O, or F) were randomly presented in each trial. The subjects and general conditions were the same as in previous experiments. The resulting betting rates (mean ± s.d.) were 0.644 ± 0.350, 0.709 ± 0.289, and 0.901 ± 0.129 for conditions B, and O, and F, respectively. One-Way Repeated Measures ANOVA found a significant difference among these betting rates (F (2,22)=6.21, P=0.007<0.01). Post hoc Tukey HSD test found a significant difference between the betting rates for conditions B and F (P=0.007<0.01) and conditions O and F (P=0.049<0.05). There was no significant difference between conditions B and O (P=0.668).

The significantly higher betting rate observed for condition F compared to conditions B and O cannot be explained in terms of the difference in the probability structure of reward, which is identical for all three conditions, not only in the average value but also in statistical measures of arbitrarily higher order. The manner in which the choices are presented in space and time in the four boxes condition, however, is different from that for the bet button and one box conditions. A hypothetical "almighty" subject who foresees the outcomes would be motivated to bet every time, since one of the four boxes is guaranteed to be the winning one. An illusory sense of agency (i.e., in this case, the erroneous feeling of being able to choose the winning box out of the four) might cause the anomalous betting behavior observed, in a manner different from the so-called "gambler's fallacy" (the illusion that past events will affect the outcomes for future events), or modes of behavior predicted by models of choice based on differences in statistical measures of nominal reward.

I assessed the subjects' perception of agency in the three conditions by a questionnaire taken after the games. The subjects were given the question “did you have a feeling that you could tell the winning box at the time of choice?" and were required to answer in a five-point scale (1="not at all”, 2=“hardly”, 3=“neither yes nor no”, 4=“slightly”, 5 =“very much so”, the original questions and scale descriptions in Japanese). The average perceived agency (mean ± s.d.) were 2.33 ± 1.23, 2.67 ± 1.30, and 4.00 ± 0.74, in conditions B, O, and F, respectively (Fig. 3. 3). One-Way Repeated Measures ANOVA found a significant difference among these values ($F(2,22)=7.90$, $P=0.003<0.005$). Post hoc Tukey HSD test found a significant difference between the perceived agency for conditions B and F ($P=0.003<0.005$) and conditions O and F ($P=0.017<0.05$). There was no significant difference in perceived agency between conditions B and O ($P= 0.736$).

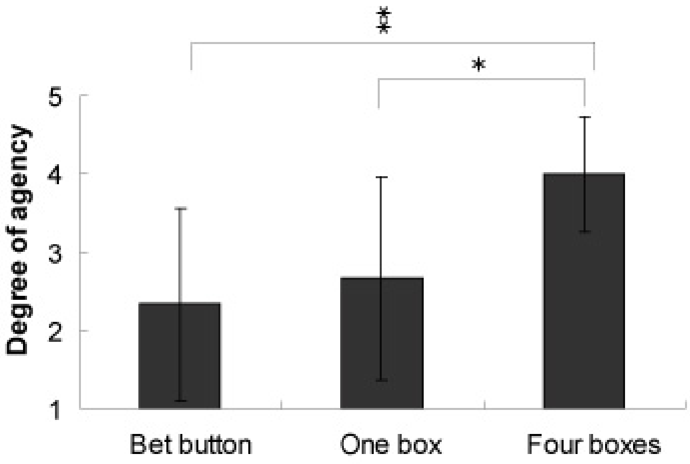


Fig. 3. 3 Average perceived agency (± s.d.). The single and double asterisks represent P<0.05 and P<0.005 in the Tukey HSD test, respectively.

I have reported an anomaly in people's betting behavior, which cannot be explained by models on human decision-making which depends on some statistical measures of external reward. These results are consistent with the interpretation that the anomaly in betting behavior was caused by an illusory sense of agency, where betting more frequently leads to more utilization of the illusory agency and less wait time. The nature of the utility of action is influenced by the way the betting opportunities are presented in space and time, even when the probability structure in terms of ensembles are exactly the same. Note that this anomaly can neither explained by ambiguity aversion nor by framing effect. The reward structures in my experiments are defined by a certain probability and in a “positive” frame. Furthermore, this anomaly does not result from social interactions.

3.3 Discussions

Dopamine also plays an important role in motor control. Decrease in dopamine level in the striatum causes the motor deficit in Parkinson’s disease. The double role of reward and motor processing in dopamine may support the utility of action.

Studies have implicated brain areas such as the right inferior parietal cortex and the right insula [44, 45] for the processing of the sense of agency. Activity of the striatum is higher when a reward is gained with one's own effort, compared to the case of a passive reception of reward [49]. The dorsal striatum (caudate nucleus) has been implicated for the instrumental learning in associating one's own action and reward [50]. Neural circuits including these areas are likely to underlie the intrinsic utility of action.

Choices and preferences should contribute to a better chance of survival, ultimately leading to an improved fitness [51]. Models of human choices based primary on monetary and other external rewards need to be expanded to a more general formulation of the utility function, reflecting the general conditions of life. Interpersonal relationships are known to be effective inducers of anomalies in behavior. Recent models of social preferences define utility functions incorporating a motive to sanction violations of fairness and cooperation norms, leading to a better prediction than models based on self-interested preferences [37]. The intrinsic utility of action, influenced by the way betting opportunities are presented in space and time, can be an important predictor of behavior in the absence of other people. It is likely to play an important role in various real life situations, where prediction of people's behavior poses interesting theoretical and practical (and often financial) challenges, as in the "click economy" on the Internet.

Chapter 4: Robust handling of uncertainty by the effect of commitment (Study 3)

In Study 2, I found the behavioral difference induced by the intrinsic utility of action in the betting games where the mathematical reward structures are equivalent. It is, however, not obvious that people prefer the game where they bet more frequently. To examine people's preference directly, I conducted another game in which they were asked to play betting games under the condition they prefer. They were forced to choose between the one box and four boxes conditions in each trial. In addition to the confirmation of their preference for the four boxes condition, the selection of the condition was found to affect the subsequent betting behavior. People rarely chose to escape, whichever game they chose. Thus, their preference of the condition cannot always be disclosed in their betting rates.

In economics, it has been known that "commitment", in which the subject renounces one or more choices, has an effect on the subsequent behavior of the subject himself or others [11]. We report here that the alternative constraining, which enables the subjects to make a commitment, can change the betting behavior in a drastic manner. The present result suggests that uncertainty characterized by the same winning probability can be heterogeneous, being affected by such elements as the perceived agency and commitment.

4.1 Methods

The participants were 7 graduate school students from universities in Tokyo (ages 22-30, 5 males & 2 females) and possessed a basic knowledge of probability theory. They were instructed about the general conditions involved in the game beforehand and gave informed consent. Before the main experiment, I examined the behavioral difference between the one box and four boxes conditions again by excluding the possibility of the subjects' betting preference being influenced by the exhaustion of resource in the course of a game. The effect of the difference in size between a box and the "Escape" button ("Control" games) was also examined. The experimental settings are the same as in Study 2 except the amount of the initial resource = 20 (= the number of trials in a game) and the size of the "Escape" button (Fig.4. 1). The mean betting rates were 0.732±0.227 and 0.921±0.141 for the one box and four boxes conditions, respectively. Consistently with Study 2, the rate in the four boxes condition was significantly higher ($p=0.020<0.05$, two-tailed paired t-test) (the black bars in Fig. 4. 3).

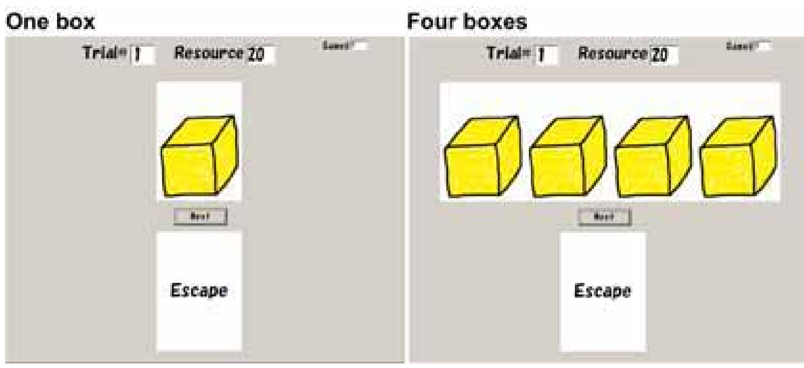


Fig. 4. 1 Control games. The subjects played independent series of games in the one box condition (left) and the four boxes condition (right) before the main experiment (30 games for each). Given the initial resource of 20 units, they were guaranteed not to run out of the resource in the course of a game. In addition, the box and "Escape" button were designed to be the same size.

Next, I tested whether higher betting rates indicate a correspondingly higher perceived values for the utility of the game. At the beginning of every trial, the subjects were asked to choose either the one box or the four boxes condition by pushing the “1box” or “ 4 boxes” button (Fig. 4. 2). This step was followed by a choice between the “Bet” and “Escape” buttons. The boxes and the “Escape” button became active 5 seconds after the chosen condition was presented (“After-selection” game).

The feedback (“You Win!” “You Lose!” or “You Escape!”) was displayed immediately after the subjects made the second choice (i.e. “Bet” or “Escape”). The screen remained the same until the subject pushed the “Next” button. One game consisted of 20 trials. Each subject repeated 30 games. When the game was over, the resource was set back to the initial value of 20, and the gains and losses were not carried over to the next session. The subjects’ preference and betting rate for each condition in the after-selection game were analyzed.

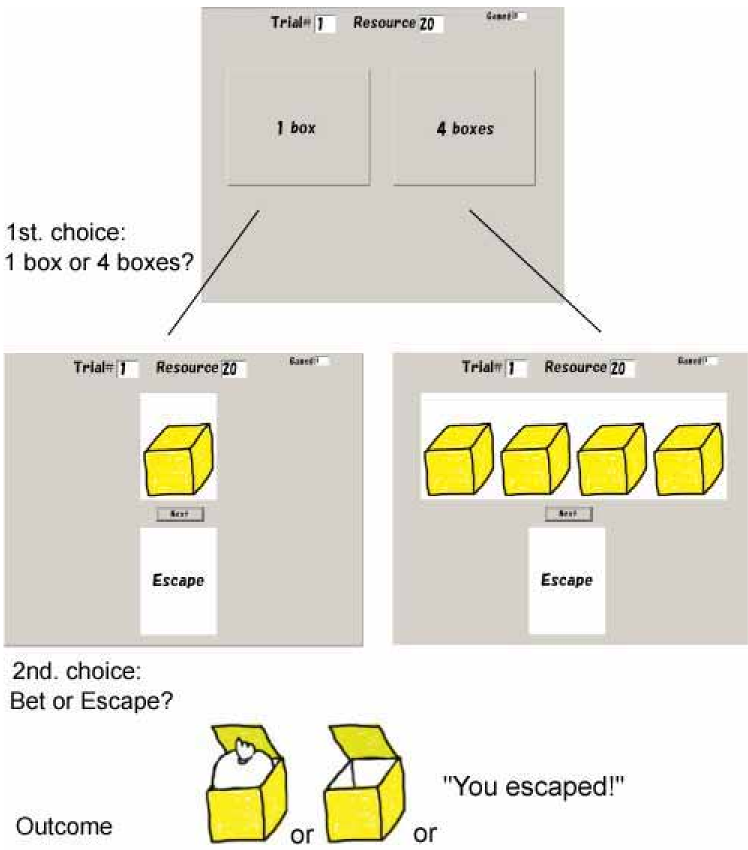


Fig. 4. 2 Experimental protocol of the main experiment (“After-selection” game).

4.2 Results

The subjects significantly preferred the four boxes condition to the one box condition (73:27 of 600 trials, $p=9.76 \cdot 10^{-30}<0.01$, Chi-square test), consistent with the interpretation that they were willing to play in the four boxes condition where they could have the stronger sense of agency.

The mean betting rates in the after-selection game were significantly different from those in the control game ($F(1,6)=10.65$, $P=0.017<0.05$, Two-way repeated measure ANOVA). The rates (± s.d.) are 0.926 ± 0.152 and 0.963 ± 0.072 for the one box and four boxes conditions, respectively (Fig. 4. 3). Two-tailed paired t-test revealed no significant difference between the conditions in the after-selection game ($P=0.316$).

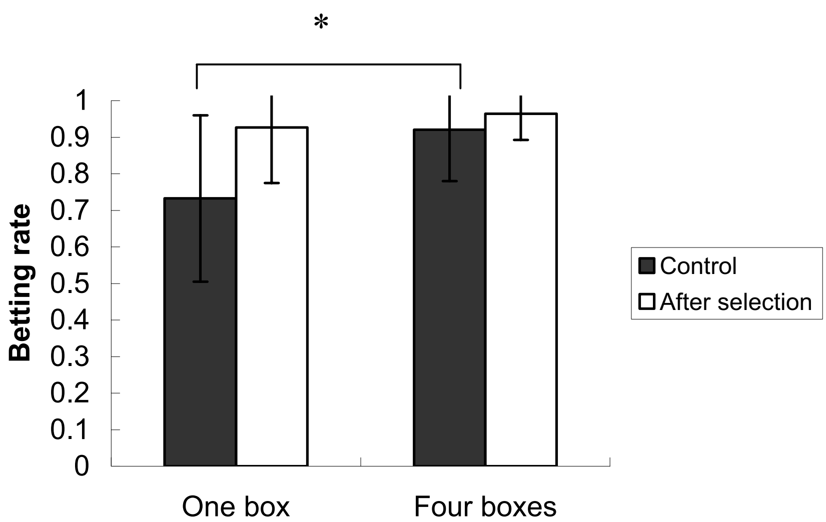


Fig. 4. 3 Average betting rates (± s.d.) in the control game (Black bars) and the after-selection game (White bar). Although subjects bet more frequently in the four boxes condition in the control game (*$P<0.05$, two-tailed paired *t*-test), the betting rates become almost the same (nearly 1) in the After-selection game.

4.3 Discussions

Under the condition studied here, the difference in the subjective value of utility between the one box and four boxes conditions remains significant. However, it is not necessarily reflected in the respective betting rates. The rates were almost 1 in either condition, unlike in the case of playing those conditions independently. This is clearly a measurable effect of commitment.

I have not constructed a randomized game in which the one box and four boxes condition were randomly (and thus passively) presented in each trial in Study 3. It is better to compare the betting rates in the after-selection game with those in the randomized game from the perspective of active versus passive alternative choice. Yet, in Study 2, there was no statistical significance between the randomized and independent series ($p=2.33$, Two-way repeated measures ANOVA). In conclusion, it is likely that the commitment (or an active choice) is the main factor that affects the subjects' expectancy of winning.

The sunk cost effect or the Concorde fallacy refers to a maladaptive economics behavior that is manifested in a greater tendency to continue an endeavor once an investment in money, effort, or time has been made [52]. A prior investment or commitment should not influence one's consideration of current options; only the incremental costs and benefits of the current options should influence one's decision. According to Arkes and Ayton, such irrational behaviors cannot be seen in human infants nor lower animals, but are remarkably observed in human adults. Our ability to generate abstract rules (e.g. past investments predict future benefits) and to self-justify (e.g. the previous waste) may sometimes invoke detrimental effects [52]. However, such abilities are, of course, not disadvantageous when the effect is considered as a total. Rather, those must have been developed in compensation for our computational limit to identify the situations in the environment a priori.

In this experiment, although the feedbacks of betting were given, the subjects could not modify their behavior. This result shows an interesting aspect of humans' cognition of uncertainty, which can have general implications. Elements involved in the subjects' agency structure can affect the assessment of the subjective value in a robust way. Previous studies have shown that voluntary actions affect the contingency perception between one's own actions and the sensory consequences [53] and conversely, that the mismatch between actions and effects modify the perceived agency [44, 45, 54, 55]. However, the present results suggest that the agency which is not influenced by any external feedback reassigns values of stimuli, leading to the different behavior. In previous learning theories, values continue to be updated until the prediction error becomes zero [24].

Based on the results reported here, I propose an alternative learning scheme in which not only the inconsistency between external and internal information but also the internal feedbacks of intentionality and agency contribute to the results of learning (Fig. 4. 4). Although this is at present only a rough sketch, this kind of internal feedback loop is likely to support the open-endedness in human learning process, where even when the external environment becomes fully predictable, the possibility of change remains in the internal loop.

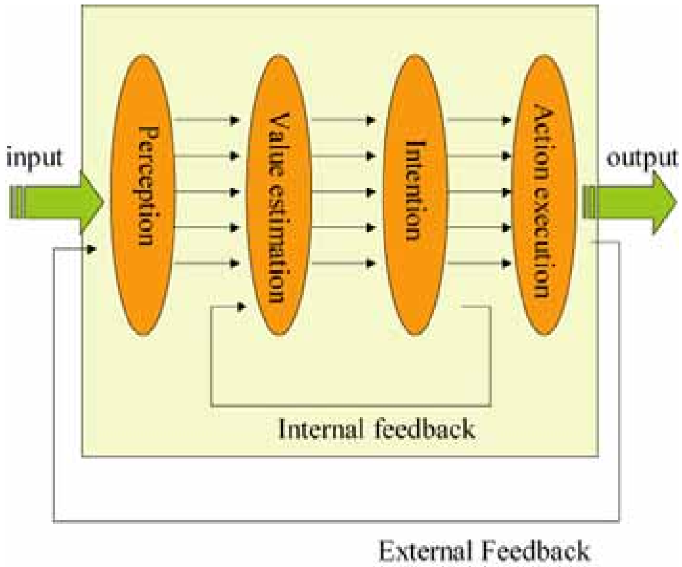


Fig. 4. 4 Learning model. Not only external feedbacks but also internal feedbacks such as agency update the value of reward.

Summary

Some studies in behavioral economics and neuroeconomics have indicated that apparently irrational behaviors of humans are caused by the necessity to develop abilities such as biasing, abstraction, or forming concepts which play essential roles in the adaptation to uncertainties encountered in the interaction with the environment.

In this study, the cognitive processes of perceiving and adapting to certain external reward structures have been investigated. Human behaviors reflecting past outcomes (Study 1), the choice among statistically equivalent alternatives (Study 2), and the effect of commitment (Study 3) have been examined. In addition to previously reported factors, the present study has revealed the significance of perception related to agency in assessing the value of a choice.

Learning has been considered to be complete when, based on the internal model, the outcomes involved in the execution of action directed to a specific goal has become fully predictable. While keeping this basic thesis, it is possible to propose three possible ways to explain open-endedness in human learning: (1) there is in practice no case where the external information becomes perfectly predictable in the real life situations, (2) the goal set by the agent changes with time. Here it is conjectured by adapting both of these points that the residual unpredictability associated with any successful learning necessitates a proactive mechanism which compensates the deficit in information, through which various kinds of goals are created. In the results reported here, robust behavioral differences have been observed even when the external conditions are the same, consistent with the existence of such a mechanism. Continuously incoming external inputs may change our goals in

themselves. In addition, the changes in goals would also depend on the internal dynamics of the cognitive process involved in the processing of agency.

The cognitive processes of uncertainty based on the reward and emotion systems, in association with the higher cognitive systems which support metacognitive factors such as agency, interfere with one another and change the perceived value in a dynamical manner. Probability associated with an uncertain event becomes, therefore, equivocal. The learning process in humans becomes open-ended in its essence, and values keep changing even after they are once assigned. Open-endedness is thus associated with onceness encountered in the course of living. Further investigations into this matter would necessitate the construction of a mathematical model of open-endedness and/or onceness.

References

[1] W. Schultz, “Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology.” *Curr. Opin. Neurobiol.* **14**, 139-147 (2004).

[2] W. Schultz, “Behavioral theories and the neurophysiology of reward.” *Annu. Rev. Psychol.* **57**, 87-115 (2006).

[3] C. D. Fiorillo, P. N. Tobler, and W. Schultz, “Discrete coding of reward probability and uncertainty by dopamine neurons.” *Science* **299**, 1898-1902 (2003).

[4] D. Kahneman, P. Slovic, and A. Tversky, *Judgment under Uncertainty: Heuristics and Biases* (Cambridge Univ. Press, New York, 1982).

[5] D. Kahneman and A. Tversky, *Choice, Values and Frames* (Cambridge Univ. Press, New York, 2000).

[6] D. Bernoulli, “Exposition of a new theory on the measurement of risk,” *Econometrica* **22**, 23-36 (1954).

[7] J. V. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton. Univ. Press, Princeton, NJ, 1944).

[8] D. Kahneman and A. Tversky, “Prospect theory: an analysis of decision under risk,” *Econometrica* **47**, 263-291 (1979).

[9] L. J. Savage, *The foundations of Statistics* (Dover, New York, 1954).

[10] N. Sebanz and W. Prinz (Eds.), *Disorders of Volition* (MIT Press, Cambridge, MA, 2006).

[11] R. H. Frank, *Passions within Reason: The Strategic Role of Emotion* (Norton, New York, 1988).

[12] G. Kanizsa, “Subjective contours”, *Sci. Am.* **235**, 48-52 (1976).

[13] P. I. Pavlov, *Conditioned Reflexes* (Oxford Univ. Press, London, 1927).

[14] E. L. Thorndike, *Animal Intelligence: Experimental Studies* (MacMillan, New York, 1911)

[15] R. S. Sutton and A. G. Barto, *Reinforcement learning: An Introduction* (MIT Press, Cambridge, MA, 1998).

[16] M. D. S. Ainsworth, “Object relations, dependency, and attachment: a theoretical review of the infant-mother relationship.” *Child Dev.* **40**, 969-1025 (1969)

[17] J. Bowlby, *Attachment and loss: Attachment* (Basic Books, New York, 1969).

[18] J. Olds and P. Milner, “Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain,” *J. Comp. Physiol. Psychol.* **47**, 419-427 (1954).

[19] R. A. Wise, “Drug-activation of brain reward pathways”, *Drug. Alcohol. Depend.* **51**, 13–22 (1998).

[20] K. McFarland and A. Ettenberg, “Haloperidol differentially affects reinforcement and motivational processes in rats running an alley for intravenous heroin.” *Psychopharmacology* **122**, 346-350 (1995).

[21] R. A. Wise, “Dopamine, learning and motivation,” *Nat. Rev. Neurosci.* **5**, 483-494 (2004).

[22] R. Elliott, J. L. Newman, O. A. Longe, and J. F. Deakin, “Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study,” *J. Neurosci.* **23**, 303-307 (2003).

[23] H. Walter, B. Abler, A. Ciaramidaro, and S. Erk, “Motivating forces of human actions. Neuroimaging reward and social interaction,” *Brain Res. Bull.* **67**, 368-381 (2005).

[24] W. Schultz, P. Dayan, P. R. Montague, “A neural substrate of prediction and reward,”

Science **275**, 1593-1599 (1997).

[25] P. W. Glimcher*, Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics* (MIT Press, Cambridge, MA, 2003)

[26] P. W. Glimcher and A. Rustichini, "Neuroeconomics: The consilience of Brain and Decision," *Science* **306**, 447-452 (2004)

[27] N. G. Mankiw, *Principles of Economics* (South-Western College Pub., Mason, OH, 2003)

[28] D. Ellsberg, "Risk, ambiguity and the Savage axioms," *Quart. J. Econ.* **75**, 643-699 (1961).

[29] H. C. Breiter, I. Aharon, D. Kahneman, A. Dale and P. Shizgal, "Functional imaging of neural responses to expectancy and experience of monetary gains and losses," *Neuron* **30**, 619-639 (2001).

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

[31] A. Tversky and D. Kahneman, "Loss aversion in riskless choice: A reference dependent model," *Quart. J. Econ*. **106**, 1039-1061 (1991).

[32] A. R. Damasio, *Descartes' Error: Emotion, Reason, And the Human Brain* (Putnam, New York, 1994)

[33] A. Bechara, H. Damasio, D. Tranel and A. R. Damasio, "Deciding advantageously before knowing the advantageous strategy," *Science* **275**, 1293-1295 (1997).

[34] M. Hsu, M. Bhatt, R. Adolphs, D. Tranel and C. F. Camerer, "Neural system responding to degree of uncertainty in human decision-making," *Science* **310**, 1680-1683 (2005).

[35] A. G. Sanfey, J. K. Rilling, J. A. Aronson, L. E. Nystrom and J. D. Cohen, "The neural basis of economic decision-making in the Ultimatum Game," *Science* **300**, 1755-1758 (2003).

[36] T. Singer, B. Seymour, J. P. O'Doherty, K. E. Stephan, R. J. Dolan and C. D. Frith,

"Empathic neural responses are modulated by the perceived fairness of others," *Nature* **439**, 466-469 (2006).

[37] D. J. de Quervain, U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck and E. Fehr, "The neural basis of altruistic punishment," *Science* **305**, 1254-1258 (2004).

[38] M. Kimura, "Evolutionary rate at the molecular level," *Nature* **217**, 624-626 (1968).

[39] M. Milinsky and C. Wedekind, "Working memory constraints human cooperation in the prisoner's dilemma," *Proc. Natl. Acad. Sci.* **95**, 13755-13758 (1998).

[40] K. K. W. Kampe, C. D. Frith, R. J. Dolan and U. Frith, "Reward value of attractiveness and gaze," *Nature* **413**, 589 (2001)

[41] J. Rilling, D. Gutman, T. Zeh, G. Pagnoni, G. Berns, C. Kilts, "A neural basis for social cooperation," *Neuron* **35**, 395-405 (2002).

[42] B. Knutson, A. Westdorp, E. Kaiser, and D. Hommer, "FMRI visualization of brain activity during a monetary incentive delay task," *Neuroimage* **12**, 20-27 (2000).

[44] C. Farrer, N. Franck, N. Georgieff, C. D. Frith, J. Decety and M. Jeannerod, "Modulating the experience of agency: a positron emission tomography study," *Neuroimage* **18**, 324-333 (2003).

[45] C. Farrer and C. D. Frith, "Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency," *Neuroimage* **15**, 596-603 (2002).

[46] F. de Vignemont and P. Fourneret, "The sense of agency: a philosophical and empirical review of the "Who" system," *Conscious Cogn*. **13**, 1-19 (2004).

[47] M. R. Roesch and C. R. Olson, "Neuronal activity related to reward value and motivation in primate frontal cortex," *Science* **304**, 307-310 (2004).

[48] J. O'Doherty, M. L. Kringelbach, E. T. Rolls, J. Hornak, and C. Andrews, "Abstract reward and punishment representations in the human orbitofrontal cortex," *Nat. Neurosci.* **4**,

95-102 (2001).

[49] C. F. Zink, G. Pagnoni, M. E. Martib-Skurski, J. C. Chappelow and G. S. Berns, "Human striatal responses to monetary reward depend on saliency," *Neuron* **42**, 509-517 (2004).

[50] J. O'Doherty, P. Dayan, J. Schultz, R. Deichmann, K. Friston, and R. J. Dolan, "Dissociable roles of ventral and dorsal striatum in instrumental conditioning," *Science* **304**, 452-454 (2004).

[51] J. B. S. Haldane, "A mathematical theory of natural and artificial selection. Part I." *Trans. Camb. Philos. Soc.* **23**, 19-41 (1924).

[52] H. R. Arkes and P. Ayton, "The sunk cost and Concorde effects: are humans less rational than lower animals?," *Psychol. Bull.* **125**, 591-600 (1999).

[53] P. Haggard, S. Clark and J. Kalogeras, "Voluntary action and conscious awareness," *Nat. Neurosci.* **5**, 382-385 (2002).

[54] S. J. Blakemore, D. M. Wolpert, C. D. Frith, "Abnormalities in the awareness of action," *Trends Cogn. Sci*. **6**, 237-242 (2002).

[55] P. J. Jackson and J. Decety, "Motor cognition: a new paradigm to study self-other interactions," *Curr. Opin. Neurobiol.* 14, 259-263 (2004).

Publications

Journal papers

[1] A. Onzo and K. Mogi, “Dynamics of betting behavior under flat reward condition,” *International Journal of Neural Systems*, Vol. 15(1-2): 93-99 (2005).

[2] A. Onzo and K. Mogi, “Anomaly in people's risk taking behavior induced by the intrinsic utility of action,” (2006) Submitted.

Conference paper

[3] A. Onzo and K. Mogi, “Cognitive Process of Emotion under Uncertainty,” *ICONIP 2004*, 300-305 (2004).

*Study 1 and 2 in this thesis correspond to [1][3] and [2], respectively.