Processing of the Incentive for Social Approval in the Ventral Striatum during Charitable Donation

Keise Izuma[1,2], **Daisuke N. Saito**[1], and **Norihiro Sadato**[1,2,3,4]

Abstract

■ Human behaviors are motivated not only by materialistic rewards but also by abstract social rewards, such as the approval of others. When choosing an action in social situations, to evaluate each action, the brain must convert different types of reward (such as money or social approval) into a common scale. Here using fMRI, we investigated the neural correlates of such valuation computations while individuals freely decided whether to donate to real charities or to take the money for themselves in the presence or absence of observers. Behavioral evidence showed that the mere presence of observers increased donation rates, and neuroimaging results revealed that activation in the ventral striatum before the same choice ("donate" or "not donate") was significantly modulated by the presence of observers. Particularly high striatal activations were observed when a high social reward was expected (donation in public) and when there was the potential for monetary gain without social cost (no donation in the absence of observers). These findings highlight the importance of this area in representing both social and monetary rewards as a "decision utility" and add to the understanding of how the brain makes a choice using a "common neural currency" in social situations. ■

INTRODUCTION

Value-based decision-making is fundamental for all animals. To select an appropriate course of action, an animal has to compare several possible alternatives based on their reward values. This basic process seems to be essential not only for wild animals to survive in physical environments but also for humans to navigate complex social environments successfully. Social exchange theory (Blau, 1964; Homans, 1961; Thibaut & Kelley, 1959) states that the behaviors of individuals in social situations are no different from their economic behaviors, as each person tries to maximize the ratio of rewards to costs. The important point is that not only materialistic rewards, such as food or money, but also nonmaterial social rewards, such as social approval or a good reputation, play a key role in social decision-making. In the psychology and economics literature, this theory has provided a possible answer to the question of why individuals help others in the absence of an apparent material benefit. Social scientists have long appreciated the influence of social approval or a "prestige benefit" on prosocial behaviors, such as donation, as one of the incentives affecting helping behaviors (Benabou & Tirole, 2006; Harbaugh, 1998a, 1998b; Holländer, 1990). Behavioral studies have also shown that introducing social approval incentives increased the contribution rate among strangers in the public goods game (Rege & Telle, 2004) and that even subtle cues suggestive of being observed by others (e.g.,, pictures of eyes) were sufficient to enhance prosocial behaviors in both laboratory (Kurzban, DeScioli, & O'Brien, 2007; Haley & Fessler, 2005) and real-life (Bateson, Nettle, & Roberts, 2006) situations.

Despite its considerable impact on human social behaviors, how the reward value of social approval is processed in the brain during decision-making remains unexplored. Thus, the present study sought to identify the brain areas involved in representing the expected values of both money and social approval during decision making. To this end, subjects engaged in a donation judgment where they were required to evaluate the expected reward values of donating (social reward) and of not donating (money) to select their action.

When individuals are deciding whether to donate when there is an expectation of social approval, the brain must convert various types of reward, including both the materialistic reward of money and the nonmaterial reward of social approval, into a common scale (currency). In this way, subjects can evaluate which action (i.e., donating or not donating) brings them more benefit (Montague & Berns, 2002). We predicted that the striatum is involved in such computations. The striatum is known to play a pivotal role in reward processing (Delgado, 2007; Schultz, Tremblay, & Hollerman, 2000) and is activated by primary

---

[1] National Institute for Physiological Sciences (NIPS), Okazaki, Aichi, Japan, [2] The Graduate University for Advanced Studies, Hayama, Kanagawa, Japan, [3] Japan Science and Technology Agency (JST)/Research Institute of Science and Technology for Society (RISTEX), Kawaguchi, Saitama, Japan, [4] University of Fukui, Eiheiji, Fukui, Japan

rewards, such as food and drink (Berns, McClure, Pagnoni, & Montague, 2001) or sexual stimuli (Redoute et al., 2000), as well as by secondary rewards, such as money (Delgado, Nystrom, Fissell, Noll, & Fiez, 2000; Elliott, Friston, & Dolan, 2000; Knutson, Westdorp, Kaiser, & Hommer, 2000). In addition, monkey single-cell recording studies have shown that neurons in the striatum encode the reward value of each action before the action is actually executed (Lau & Glimcher, 2008; Samejima, Ueda, Doya, & Kimura, 2005). Similarly, a human neuroimaging study indicated that the striatum plays a role in the value representation of a given choice during decision-making (Tom, Fox, Trepel, & Poldrack, 2007). This study demonstrated that the same striatal area showed increasing activity with an increase in the potential monetary gain and decreasing activity with an increase in the potential monetary loss, suggesting that both positive and negative values are represented in a common valuation system (Tom et al., 2007). Studies from the newly emerging field of neuroeconomics have also reported that human striatal activity is modulated by social preferences (Fehr & Camerer, 2007). Moreover, we previously found that perceiving one's good reputation among others (social approval) activated the same striatal areas that were activated by the acquisition of monetary rewards (Izuma, Saito, & Sadato, 2008). Therefore, we predict that the striatum plays an important role in representing both materialistic and abstract social rewards during prosocial decision-making.

We scanned 23 subjects using fMRI while they engaged in a simple decision task, in which they decided whether to donate ¥500 (~€3 or US$5) to charity or to take the money for themselves. Although the effects of both intrinsic and extrinsic (or "prestige") benefits on prosocial behaviors have been acknowledged (Harbaugh, 1998a), previous neuroimaging studies investigating brain activations during charitable donations (Harbaugh, Mayr, & Burghart, 2007; Moll et al., 2006) focused only on the "intrinsic benefit" of donation by guaranteeing anonymity in the experiments. Moll et al. (2006) reported striatal activation in response to pure monetary gain and to noncostly donations. Similarly, Harbaugh et al. (2007) observed higher striatal activation both when subjects expected to obtain more money and when they anticipated that a charity would get more money during mandatory transfers ("pure altruism"). In addition, the ventral striatum showed higher activations when the transfers were voluntary compared with mandatory transfers ("warm glow"). Here, we exclusively focused on the extrinsic "prestige benefit" of social approval (Harbaugh, 1998a) and systematically manipulated the incentive for this extrinsic social reward by asking the subjects to decide whether to donate in the presence or absence of an audience (Figure 1). This manipulation is known to greatly influence prosocial tendencies (Kurzban et al., 2007; Bateson et al., 2006; Haley & Fessler, 2005). For one half of the sessions, the performance of subjects on the donation task was constantly observed by two persons (both of whom were actors), and the subjects could see their faces through a video camera during the task (the presence condition; Figure 1, top). In the other half of the sessions, the subjects viewed the tops of two chairs, indicating that no one was watching their performance (the absence condition; Figure 1, bottom). Although all of the subjects believed that these observers were truly present, in reality the images that the subjects saw during fMRI scanning were prerecorded videotapes. During the presence and absence conditions, the same 78 real charitable

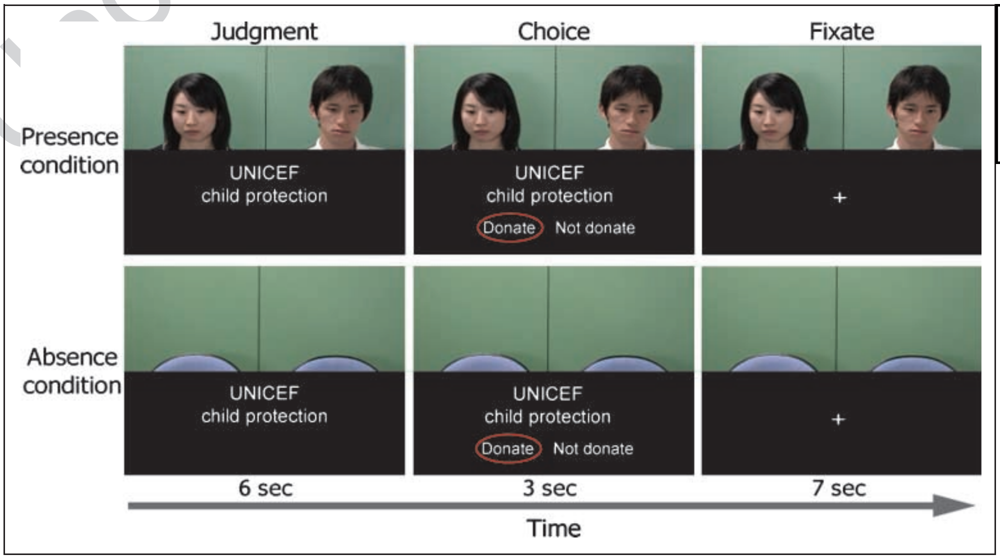

Figure 1. Sequence of events during trials in the presence (top) and absence (bottom) conditions. In the presence condition, two observers were shown on the top half of the screen. Subjects were led to believe that the observers were present in the fMRI control room and were watching their performance. In the absence condition, the tops of two chairs were shown to indicate that nobody was watching. In each trial, the name of a charity and its brief mission statement were presented for 6 sec, during which subjects were asked to decide whether to donate ¥500 to the organization (subjects could keep any money that was not donated). Then, two choices ("donate" and "not donate") were presented randomly on the right or left side of the screen for 3 sec, during which subjects made their selection. Their choice was highlighted by a red circle and thus was observable by others. The intertrial interval was 7 sec. We focused on the brain activity during the decision-making period (the 6-sec judgment period).

Table 1. Examples of Charitable Organizations Used in the Present Study

	Organizations	*Web Addresses*
1	United Nations Children's Fund (UNICEF) Japan	http://www.unicef.org/
2	Global Sports Alliance	http://gsa-world.org/english/
3	Japan Foundation for AIDS Prevention	http://www.jfap.or.jp/english/index.htm
4	BirdLife Asia	http://www.birdlife-asia.org/eng/about/index.html
5	World Food Programme	http://www.wfp.org/english/
6	Retired Weapons	http://www.retired.jp/
7	Sea Turtle Association of Japan	http://www.umigame.org/E/ETop.html
8	Mozilla Japan	http://www.mozilla.org/foundation/
9	Japanese Red Cross Society	http://www.jrc.or.jp/english/index.html
10	Japan Spinal Cord Foundation	ttp://www.jscf.org/english/index.html

In total, 78 charitable organizations were used in the study, and most of the organizations were chosen from the Yahoo Japan volunteer Web page at http://volunteer.yahoo.co.jp/donation/index.html.

organizations of varying perceived importance (Table 1) were presented one at a time.

As in the previous study using the donation task with multiple organizations (Moll et al., 2006), we differentiated between trials according to subjects' subsequent choices. We employed a 2 (Presence vs. Absence of observers) × 2 (subjects' choice of ¥500 to Charity vs. Self) factorial design to analyze the imaging data. On the basis of the findings of previous behavioral studies (Kurzban et al., 2007; Bateson et al., 2006; Haley & Fessler, 2005), subjects' estimations of both social approval (via the act of donating) and social disapproval (via not donating) were considered to be much higher in the presence condition than in the absence condition, whereas all other costs and benefits for each action were considered to be the same between these two conditions (Figure 2). Accordingly, although subjects subsequently selected the same "donate" choice, the expected value of the action would be larger in the presence condition than in the absence condition, because of the additional reward of social approval brought about by making a donation in public. By contrast, when subjects subsequently chose not to donate, the expected reward value of the action might be reduced in the presence condition relative to the absence condition due to the negative reward value of potential social disapproval as the result

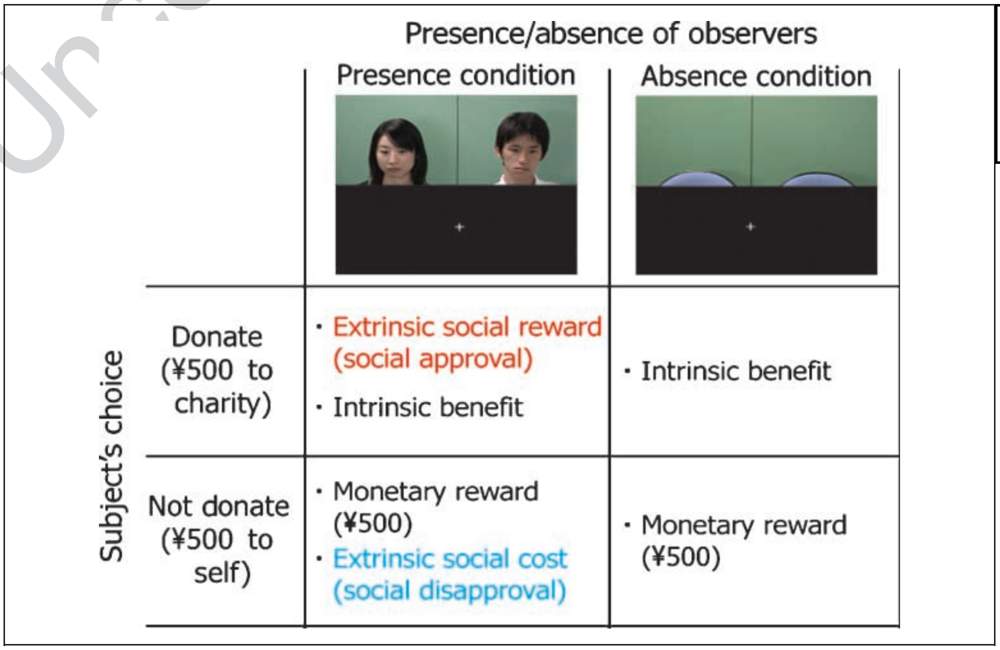

Figure 2. Experimental design and relevant rewards for each choice ("donate" or "not donate") in each condition (presence or absence). During decision-making, subjects' estimation of the intrinsic benefit of the act of donating should be the same regardless of the presence/absence of observers. However, the extrinsic reward of social approval was a factor only when their decision was observed by others. Therefore, the reward value of social approval was processed only in the presence condition. Similarly, although subjects estimated the positive reward value of ¥500 to an equal extent during the presence and absence conditions, the negative reward value of social disapproval was processed only in the presence condition.

of selfish behavior in front of others. Therefore, we specifically predicted that if the same striatal area encodes the reward values of both money and extrinsic social approval/disapproval duringprosocial decision-making, it should show a 2 (Presence vs. Absence) × 2 (¥500 to Charity vs. Self) interaction. Particularly high activations were predicted when subjects expected high extrinsic social rewards (donating ¥500 to charity in the presence condition) and when they expected a monetary reward with no social costs (opting to keep the ¥500 for themselves in the absence condition; Figure 2).

METHODS

Participants

In total, 29 healthy nonsmoking subjects participated in the study. The reported analyses were based on 23 subjects (11 males; mean age = 22.7 ± 4.7 years; one left-handed). Five subjects were excluded from the analysis because they failed to provide sufficient data to model: two of the five subjects chose a "donate" option fewer than three times per session (i.e., <12%; one session consisted of 26 trials) and/or <12 times in total (i.e., <15%) during the three presence or absence sessions; the remaining three subjects were excluded because they chose a "not donate" option fewer than three times per session and/or <12 times in total during the three presence or absence sessions. One further subject was excluded because during the interview after the experiment, she reported that she did not believe the observer manipulation and strongly doubted that her performance was being observed even in the absence condition. None of the subjects had a history of neurological or psychiatric illness. All subjects gave written informed consent for participation, and the study was approved by the Ethical Committee of the National Institute for Physiological Sciences, Japan.

Procedure

When the subjects arrived at the scanner control room, they were informed that they would take part in various judgment tasks regarding charitable contributions within the fMRI scanner. Then, all of the subjects were asked to read a brief written description of 78 charitable organizations for ~50 min. All of the organizations were real and were actually asking for donations through the Internet. Although the activities of the organizations were all more or less socially desirable, their perceived importance varied considerably, from those intending to save people's lives to those whose goal was to distribute free software on the web (for examples of the organizations included, see Table 1). To prevent the subjects from making their donation decisions before the fMRI scanning, we explained the details of the fMRI task after they had finished reading the descriptions. Also, to manipulate the potential social approval that might result from a donation, the individuals needed to know how important other people considered each organization to be. To provide this information briefly, we ordered the descriptions of the 78 organizations according to their importance, as rated by another sample of 28 subjects (11 male, mean age = 24.1 ± 4.4 years). It should be noted that unlike Moll et al. (2006), we did not include any organizations whose activities were related to politically sensitive issues such as abortion, death penalty, and euthanasia to ensure that there were strong positive correlations between the perceived importance ratings made by the subjects and by the 28 other participants (see the Behavioral results section). Therefore, in this phase of the experiment, the subjects learned about the activities of each organization and the extent to which each was considered socially important by other people.

Furthermore, the subjects were advised that two students from the National Institute for Physiological Sciences (who were actually one male actor and one female actor) were participating in the study and that in three of the six sessions these two people would monitor their performance through a video camera in the room next to the fMRI scanner. Outside the scanning room, all of the subjects were shown the experimental setup, which consisted of two chairs, a video camera, a screen, and a screen splitter. The subjects were led to believe that their responses during scanning would be shown on the screen and that they would see the faces of the two observers who were sitting in the chairs and watching their performance during the task (the presence condition). The subjects were also told that these observers had learned about the organizations and their activities, and when these observers were on the screen, they were engaged in an "impression formation task." The details of this task were not explained to the subjects, but it was emphasized that the two observers would be paying attention to their performance. Also, the subjects were told that when only the upper portions of the two chairs were visible, the observers would be engaged in different tasks that were unrelated to the subject's task (the absence condition). For most of the subjects, the two observers were completely unfamiliar individuals whom they had not previously met. Two of the 23 subjects had met the male observer once before in another behavioral experiment run by our laboratory, but they were not personal acquaintances; the main findings were largely unchanged when the data from these two subjects were excluded from the analysis.

After completing a 3-min practice session in the scanner control room (the organizations used in the practice session were not included in the main paradigm), all of the subjects met and were introduced to at least one of the two observers before entering the fMRI room. If one of the observers could not meet a subject, that individual was told that the observer was on their way and would arrive shortly. Moreover, to make the interaction with the two observers during scanning more meaningful, the subjects were told that they would engage in some other tasks together with the two observers after the fMRI scanning. All of the subjects believed that the two observers were in the room next to the fMRI scanner room during the task (this was confirmed during an interview at the end of the experiment); however, in reality, a prerecorded video was used throughout the sessions. To control for possible confounding factors elicited by changes in the observers' facial expressions or head and eye movements, their expressions were kept neutral, and their eye gaze was fixed on the bottom half of the screen, as if they were looking at the subjects' responses. Head and eye movements were kept to a minimum throughout the video (all of the subjects were told in advance that the two observers would be asked to stay still and to focus on their task, in order not to distract the subjects).

After scanning, the subjects were asked to complete a postexperimental questionnaire, which included rating the importance and familiarity of each of the 78 organizations presented during fMRI using a 7-point scale (for the importance rating, 1 = *not important at all* to 7 = *very important*; for the familiarity rating, 1 = *not familiar at all* to 7 = *very familiar*) and answering the yes/no question "Did the presence of observers have any influence on your choice during the donation task?" Also, the subjects were interviewed to examine whether they had any doubt or suspicions regarding the experimental paradigm. Finally, all of the subjects were fully debriefed as to the purpose of the study and the use of the video and were paid a fixed amount of ¥8,000 (plus ¥500 if they had chosen to "not donate" in a randomly selected trial).

Experimental Task

During fMRI, the subjects were asked to decide whether they wanted to donate ¥500 (~€3 or US$5) to each charity presented on the bottom half of the screen. When subjects chose to "not donate," they could keep the ¥500. The subjects made a total of 156 such donation decisions (78 each in the presence and the absence conditions). Also, the subjects were told that after the experiment, one trial would be chosen randomly using a bingo machine, and their decision during that trial would count in reality: If the subjects had chosen to donate during that trial, they would actually donate ¥500 to the charity through the Internet; if they had not donated, they would receive ¥500 in addition to the fixed amount paid for their participation. Before the fMRI scanning, all of the subjects were shown the Internet page through which they could make their donation and a Japanese ¥500 coin.

During scanning, in each trial, the subjects were presented with the name of a charitable organization and its brief mission statement for 6 sec, during which they were asked to decide whether to donate (Figure 1). After the 6 sec, the two choices of "donate" and "not donate" were presented on the screen for 3 sec, during which the subjects were asked to select what they would like to do by using their right index or middle finger. The subjects were encouraged to choose freely and were told that it was fine to choose to "not donate" and to take the ¥500 in all trials if they would like to do so. Also, it was emphasized that the subjects should ignore the two observers on the screen and should concentrate on their own task. The choices of "donate" or "not donate" were presented randomly on either the right or the left side of the screen, so that the button they pressed (index or middle fingers) and the choice they made ("donate" or "not donate") were uncorrelated. Soon after they made their choice, the chosen option was highlighted by a red circle (and was thus clearly visible to the observers). Then, a cross hair was presented for 7 sec before the onset of the next trial. To ensure that there were equal numbers of donated and not-donated trials in each session, we divided the 78 organizations into three groups of 26 organizations, each of which had a similar average importance rating according to another sample of subjects ($n = 28$). Within each session, the order of presentation was randomized for each subject, and the order of the three sessions within the presence and the absence conditions was counterbalanced across subjects. One half of the subjects performed the three presence sessions followed by the three absence sessions, and the other half performed the experiment in the reverse order.

All of the stimuli used in the task were prepared and displayed using Presentation software (Neurobehavioral Systems) on a microcomputer (Dimension 8200; Dell Computer Co.). The videos showing the two observers and the two chairs were played by a digital video cassette player (GV-D1000; Sony). Using a liquid-crystal display projector (DLA-M200L; Victor), the visual stimuli were projected onto a half-transparent viewing screen via a screen splitter (MV-40F; FOR-A) so that the video was shown on the top half of the screen, and the stimuli for the tasks were presented on the bottom half of the screen. The screen was located behind the head coil, and the subjects viewed the stimuli through a mirror. All of the stimuli for the tasks were written in Japanese and were presented as white letters against a black background.

Image Acquisition and Analysis

Images were acquired using a 3-T MR imager (Allegra, Siemens). For functional imaging, interleaved T2*-weighted gradient-echo EPI sequences were used to produce 38 continuous 3-mm-thick transaxial slices covering nearly the entire cerebrum (repetition time = 2,000 msec; echo time = 25 msec; flip angle = 85°; field of view = 192 mm; 64×64 matrix; voxel dimensions = $3.0 \times 3.0 \times 3.0$ mm). A high-resolution anatomical T1-weighted image was also acquired by magnetization-prepared rapid gradient-echo imaging (repetition time = 2.5 s; echo time = 4.38 msec; flip angle = 8°; 256×256 matrix; 192 slices; voxel dimensions = $0.75 \times 0.75 \times 1$ mm) for each subject.

After discarding the first six volumes to allow for stabilization of the magnetization, the remaining 208 volumes per session in the experiment (a total of 1,248 volumes per subject for six sessions) were used for analysis. The data were analyzed using SPM5 software (Wellcome Department of Imaging Neuroscience) implemented in Matlab 7.1 (Mathworks). After correcting for differences in slice timing within each image volume, head motion was corrected using the realignment program of SPM5. Following realignment, the volumes were normalized to Montreal Neurological Institute space using a transformation matrix obtained from the normalization process of the first EPI image of each individual subject to the EPI template. The normalized fMRI data were spatially smoothed with a Gaussian kernel of 6 mm (FWHM) in the x-, y-, and z-axes.

Statistical analysis was conducted at two levels. First, the individual task-related activation was evaluated. Second, the summary data for each individual were incorporated into a second-level analysis using a random-effect model (Friston, Holmes, & Worsley, 1999) to make inferences at a population level.

In the individual analysis, the signal was scaled proportionally by setting the whole-brain mean value to 100 arbitrary units. The signal time course for each subject was modeled with a general linear model. Regressors of interest (condition effects) were generated using a boxcar function convolved with a hemodynamic response function. We included three regressors of interest in each of the presence and absence conditions: First, the judgment period (duration = 6 sec) for trials in which the subject's subsequent choice was to "donate"; second, the judgment period (6 sec) for trials in which the subject's subsequent choice was to "not donate"; and third, the choice period (duration = 3 sec), regardless of the subject's decision. Regressors that were of no interest, such as the session effect, and high-pass filtering (128 sec) were also included. In addition, failed trials (i.e., trials in which subjects did not press any button within the 3-sec choice period) were modeled separately as a regressor of no interest.

The weighted sum of the parameters estimated in the individual analysis consisted of "contrast" images that were used for the group analyses. In this second-level analysis, we used a full factorial design with four contrast images from each subject: first, "donate (¥500 to Charity)" in the presence condition; second, "donate (¥500 to Charity)" in the absence condition; third, "not donate (¥500 to Self)" in the presence condition; and fourth, "not donate (¥500 to Self)" in the absence condition. The contrast images obtained by the individual analyses represented the normalized increment of the fMRI signal for each subject. The SPM$\{t\}$ for the contrast images was created as described above. Significant signal changes for each contrast were assessed by means of t statistics on a voxel-by-voxel basis.

On the basis of previous reports (Izuma et al., 2008; Lau & Glimcher, 2008; Fehr & Camerer, 2007; Harbaugh et al., 2007; Tom et al., 2007; Moll et al., 2006; Samejima et al., 2005), we focused our ROIs in the striatum, and we defined ROIs both anatomically and functionally. Anatomical ROIs in the striatum (the caudate and the putamen bilaterally) were generated using the WFU PickAtlas toolbox for SPM (Maldjian, Laurienti, Kraft, & Burdette, 2003) with a dilation factor of 1. To further define the ROIs functionally, we intersected the anatomical ROIs with voxels showing significant increases ($p < .001$, uncorrected for multiple comparisons) in activation both during the "donate" trials in the presence condition and during the "not donate" trials in the absence condition. Our predicted patterns of interaction should be characterized by particularly high increases in activation during these two conditions rather than by decreases in the other two conditions from the baseline ("donate" in the absence condition and "not donate" in the presence condition). We reasoned that before making a donation in front of others, the subject's estimation of the extrinsic social reward is especially high, and thus in this condition, the striatum should show a significant increase in activation relative to the baseline condition. Also, before deciding not to donate in the absence of others, subjects estimated the reward value of the money (¥500) without regard for the presence/absence of observers. However, it might be that, in the presence of others, taking ¥500 for oneself (i.e., not donating) might carry the cost of social disapproval. A previous study reported that activations in the same striatal area were positively correlated with monetary gain and negatively correlated with monetary loss during decision-making (Tom et al., 2007); the striatal activity might be similarly affected by the negative reward value of social disapproval. Thus, if the same striatal regions process not only social rewards but also monetary rewards, the striatal activity during the decision to "not donate," especially in the absence condition (i.e., monetary reward with no social cost), should show a reliable increase relative to the baseline.

Within the ROIs, we explored the positive interaction between (Presence − Absence) × (Donate − Not Donate). The results are reported at $p_{\mathrm{FWE}} < .05$, with an extent threshold of >10 contiguous voxels. If activation was found in one hemisphere, the threshold was lowered to $p < .005$ (uncorrected) to establish whether there was activation in the same area of the other hemisphere. Activations outside the ROIs were also reported at $p < .001$ (uncorrected for multiple comparisons) with an extent threshold of >10 contiguous voxels.

RESULTS

Behavioral Results

There were few trials in which subjects did not press any button within the 3-sec choice period, and there was no significant difference in the mean number of failed trials between the presence (0.35) and the absence (0.83) conditions (paired t test, $p = .22$, *ns*).

Our behavioral data confirmed that the presence of observers had a powerful effect on prosocial behaviors. Although subjects were presented with the same 78 organizations, they donated more often in the presence condition (33.3 times, on average) than in the absence condition (31.3 times, on average), paired t test, $t(22) = 2.54$, $p = .009$. It was further predicted that the impact of the observer manipulation on the donation decision would be greater when the charities were perceived to be moderately important, whereas if the organizations were regarded as highly important, the subjects would donateregardless of the presence of observers, and if the charities were perceived as having almost no social importance, the subjects would not donate even in the presence of observers. In the current study, on average, 21.9 of 78 charitable organizations were rated as highly important (an importance rating of 6 or 7 on the 7-point scale), 43.4 organizations were rated as moderately important (an importance rating of 3–5), and the remaining 12.7 organizations were rated as relatively unimportant (an importance rating of 1 or 2). The effects of observers on the average number of donations were plotted according to the three categories of perceived importance (Figure 3). A 2 (presence or absence of observers) × 3 (perceived importance of charities: high, middle, or low) repeated measures ANOVA revealed the main effects of both observers, $F(1,22) = 6.45$, $p = .019$, and the perceived importance of the charity, $F(2,44) = 23.95$, $p < .001$. More importantly, as predicted, it revealed a significant interaction, $F(2,44) = 4.08$, $p = .024$. Paired t tests further showed that although there were no differences in the number of donations between the presence versus absence conditions for the high and the low importance categories ($p > .56$, *ns*, for both), the subjects donated more often in the presence condition than in the absence condition for those charities that had a moderate level of perceived importance, $t(22) = 2.34$, $p = .014$.

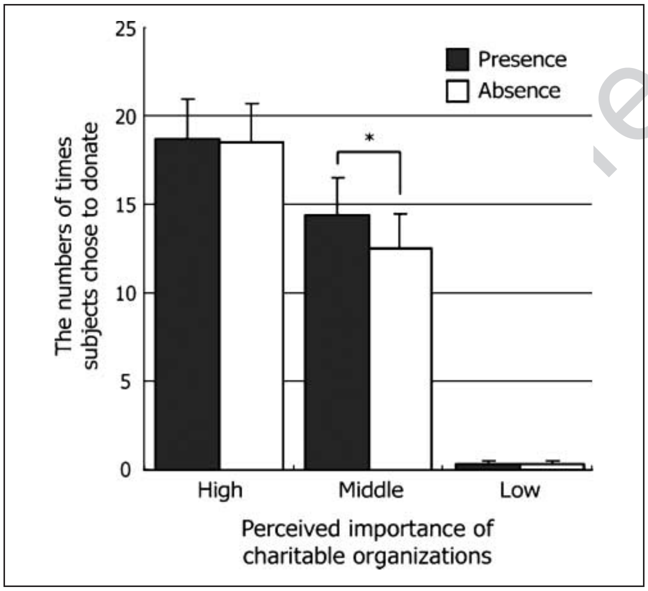

Figure 3. Behavioral results. The average numbers of times subjects chose to donate in the presence and the absence conditions were plotted separately for the three levels of perceived importance of the charitable organizations. Each charity was categorized into one of three importance levels according to each subject's perceived importance rating of that charity. The organizations with an importance rating of 6 or 7 on the 7-point scale were classified in the high importance category (on average, 21.9 of 78 organizations). The middle importance category included 43.4 organizations with an importance rating of 3–5. The low importance category included 12.7 organizations with an importance rating of 1 or 2. Error bars denote the *SEM*; $^*p < .05$ (paired t test).

The importance rating that each subject gave to the 78 charitable organizations significantly correlated with the mean importance ratings assigned by another sample of 28 subjects, indicating that all subjects had normal, representative attitudes toward these charities (mean of the 23 within-subject Pearson correlation coefficients, $r = 0.63$), one-sample t test, $t(22) = 22.6$, $p < .001$; for all analyses hereafter, the within-subject correlations were converted to z scores using Fisher's z transformation, but the mean values reported here have been converted back to correlation coefficients. Also, although subjects were more likely to donate in the presence condition than in the absence condition, their decisions were largely consistent across the two conditions ($r = 0.76$), one-sample t test, $t(22) = 11.9$, $p < .001$.

As shown in Figure 3, subjects' donation decisions during fMRI scanning also were highly correlated with the importance ratings that they assigned to the 78 organizations (mean of the within-subject Pearson correlation coefficients, $r = 0.67$), one-sample t test, $t(22) = 16.9$, $p < .001$; for this computation, the choice to "donate" was coded as 1 and the choice to "not donate" was coded as 0. This indicates that the subjects were more likely to donate when they considered the activity of an organization to be comparatively important. This tendency did not differ between the presence and the absence conditions (paired t test, $p = .82$, *ns*).

The familiarity ratings were positively correlated with the donation decision ($r = 0.41$), one-sample t test, $t(22) = 12.2$, $p < .001$, indicating that subjects tended to donate to organizations with which they were more familiar. However, this was probably due to the highly positive correlation between the importance and the familiarity ratings ($r = 0.53$), one-sample t test, $t(22) = 9.80$, $p < .001$. The correlations between the donation decision and the familiarity ratings did not differ significantly between the presence ($r = 0.42$) and the absence ($r = 0.39$) conditions (paired t test, $p = .08$, *ns*).

We conducted a 2 (Presence or Absence of observers) × 2 (subjects' choice of ¥500 to Charity vs. Self) repeated measures ANOVA on the RT data and found neither significant main effects nor significant interaction (all p values $> .64$, *ns*). Overall, subjects' average RTs were 945.2 msec ($SD = 228.1$ msec).

The postexperimental questionnaire included the yes/no question: "Did the presence of observers have any influence on your choice during the donation task?" In total, 21 of 22 subjects (data were not collected from one subject) answered "no" and denied any influence of social incentive on their decisions (chi-square test, $p < .001$). Although other interpretations are also possible, these data suggest that the cost-benefit analysis performed during the donation decision is not necessarily a conscious process.

Imaging Results

Our fMRI data showed that, within the ROIs of the striatum, there was a significant interaction effect during the judgment period in the left ventral striatum ($p_{\mathrm{FWE}} < .05$). When we lowered the threshold to $p < .005$ (uncorrected), the activation was also found in the right ventral striatum (Figure 4A). For each subject, beta values were extracted from the peak voxels in the two regions of the ventral striatum for all four conditions, and the patterns of activation also confirmed our predictions (Figure 4B). The data showed that, although subjects eventually made the same choice (i.e., to "donate" or "not donate"), their striatal activities during decision-making were significantly modulated by the presence or absence of observers (i.e., the possibility of gaining an extrinsic social reward). The direct comparison of activations in "donate" or "not donate" trials between the presence and the absence conditions showed that even when subjects made a similar choice to donate, the left striatal activations were significantly higher when the donations were made with observers than when no one was observing the donations, paired *t* test, $t(22) = 2.87$, $p = .004$, and the same comparison in the right striatum showed a similar nonsignificant trend, paired *t* test, $t(22) = 1.41$, $p = .09$, *ns*. By contrast, although the striatal activations for "not donate" trials tended to be higher in the absence condition than in the presence condition, the differences were not significant for both the right and the left striatum ($p > .18$, *ns*, for both).

Outside the a priori ROIs, the areas showing significant interaction ($p < .001$, uncorrected) included ventral striatum, parahippocampal gyrus, and midbrain (Table 2).

DISCUSSION

To our knowledge, this is the first evidence that, during prosocial decision-making in front of others, the extrinsic reward of social approval is processed in the same striatal region that encodes monetary rewards. Our behavioral data showed that the mere presence of observers enhanced a subject's tendency to donate, which is consistent with the results of previous behavioral studies (Kurzban et al., 2007; Bateson et al., 2006; Haley & Fessler, 2005). This behavioral evidence indicates that there was an

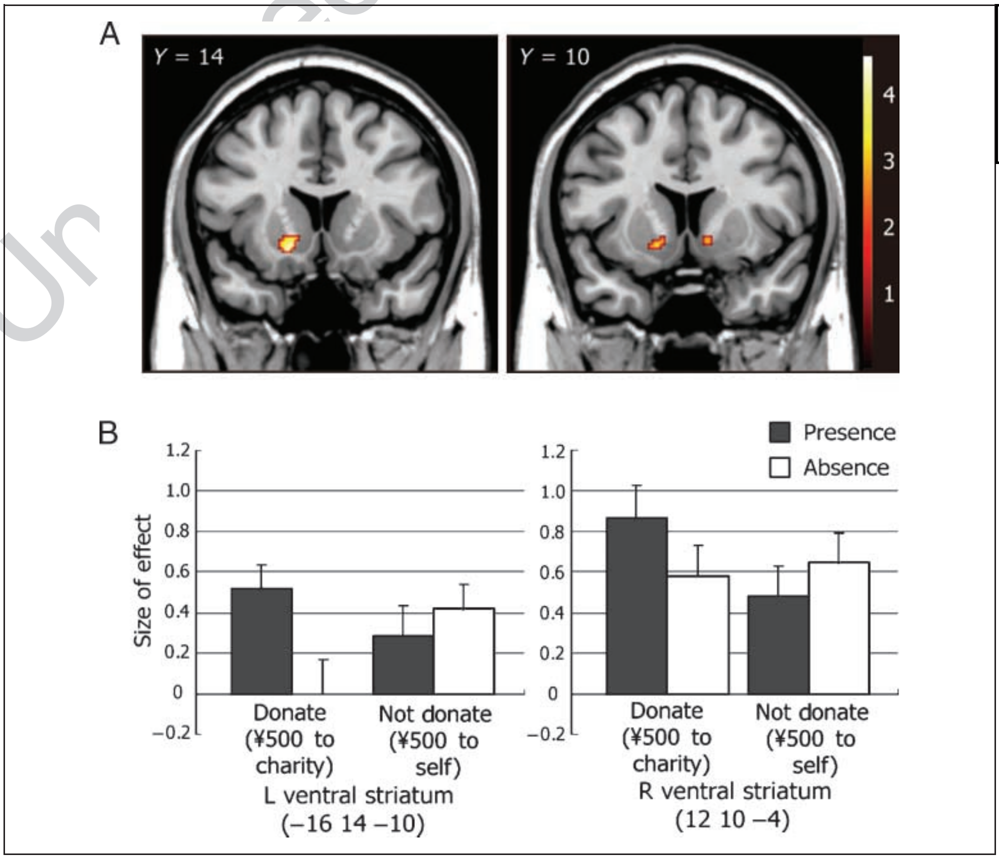

Figure 4. fMRI results. (A) Coronal slices showing significant ventral striatum activations ($p < .005$, uncorrected) in the 2 (Presence vs. Absence) × 2 (¥500 to Charity vs. Self) interaction contrast. This contrast was explored within the a priori ROIs. The activation in the left ventral striatum was found at a threshold of $p_{\mathrm{FWE}} < .05$, and the right ventral striatum showed significant activation when the threshold was lowered to $p < .005$ (uncorrected). The scale shows the *t* values. (B) Bar graphs indicate the effect sizes at the peaks in the ventral striatum in each condition. Error bars denote the *SEM*.

Table 2. Areas Activated by the Interaction Contrast of (Donate − Not Donate) × (Presence − Absence) outside the ROIs

			Montreal Neurological Institute Coordinate				
Location	*Side*	*Brodmann's area*	*x*	*y*	*z*	*Z*	*Size (Voxel)*
Ventral striatum[a]	L	−10	6	−12	3.70	15	
Parahippocampal gyrus	R	34	10	−6	−20	3.56	22
Midbrain	R		10	−20	−16	3.58	32

The statistical threshold was set at $p < .001$ (uncorrected for multiple comparisons) with a cluster extent threshold of >10 contiguous voxels.

[a] Although in the striatum, this cluster was not within our predefined ROIs and is different from the cluster reported in Figure 4.

additional incentive (social approval) to donate in the presence condition, thus validating the psychological framework illustrated in Figure 2. Our neuroimaging data showed that although subjects subsequently made the same choice ("donate" or "not donate"), this manipulation of social situations significantly affected the striatal activity during decision-making. Furthermore, in these striatal regions, particularly high activations were observed both when subjects estimated a high extrinsic social reward ("donate" trials in the presence condition) and when they estimated a monetary reward without any social cost ("not donate" trials in the absence condition). These data indicate that the same ventral striatal areas encoded both extrinsic social and monetary rewards during prosocial decision-making.

The current study provides an important insight into the role of the striatum in representing social rewards and the neural basis of social decision-making. First, these data extend our previous finding (Izuma et al., 2008) that monetary rewards and the more abstract reward of social approval (or a good reputation) are both processed in the human striatum as an "experienced utility," which is the satisfaction derived from the outcome that is actually obtained. The present study demonstrated that the ventral striatum also processes social approval (or a good reputation) as a "decision utility," which refers to the weight that individuals assign to a decision's outcome (Kahneman, Wakker, & Sarin, 1997). The role of the striatum in representing a decision utility is consistent with a previous study, which used a gambling paradigm with monetary rewards (Tom et al., 2007). It should be stressed that, because the present study focused on how the brain processes social approval not as an "experienced utility" but as a "decision utility," by using the prerecorded video, we ensured that the subjects did not receive any social feedback from the observers. Throughout the video, the observers' facial expressions were kept neutral so that subjects did not receive any feedback based on these cues (such as approval or disgust). Second, adding to previous studies reporting that the intrinsic benefit of donation activated the ventral striatum (Harbaugh et al., 2007; Moll et al., 2006), the present study found that the extrinsic benefit of donation was also processed in the ventral striatum during decision-making.

Taken together, these findings add to the understanding of the neural basis of human social decision-making. These important advances suggest that the striatum, particularly the ventral striatum, plays an essential role in representing a "common currency" for rewards, including monetary rewards and both intrinsic and extrinsic social rewards. This allows for the precise valuation of each possible action in a given situation, and the best course of action can be chosen depending on the assessed values of the potential actions (Glimcher, Dorris, & Bayer, 2005; Montague & Berns, 2002).

The present findings, therefore, improve our understanding of the neural basis of human social decision-making. Our data clearly showed that although it was not obvious, the extrinsic social reward of the approval of other people strongly influenced donation behaviors and modulated the activity of the brain's reward system, particularly the ventral striatum. Thus, our results suggest that human behaviors in social environments are not unusual, as, like other behaviors, they attempt to maximize the expected utility. Also, despite the apparent cost (such as time, energy, or money), prosocial behavior can, in fact, be considered rational in terms of its neural representation, which is an idea that is consistent with physiological utility theory (Glimcher et al., 2005). Therefore, the brain activation patterns in the present study illustrate that individuals are, at least in part, exchanging money for social rewards by donating to charity, especially in front of other people, as suggested by social exchange theory (Blau, 1964; Homans, 1961; Thibaut & Kelley, 1959).

The present study demonstrated that the reward values of both money and social approval were processed in the ventral striatum during charitable donation. However, the findings did not provide sufficient evidence to establish that the ventral striatum represented a "common currency" for rewards during decision-making because we did not explicitly investigate where in the brain the values of the two actions associated with different rewards were compared. The common neural currency is important because it allows an individual to compare several actions, each of which is associated with various reward stimuli, using a common scale, and to choose the most appropriate action in a given situation based on this comparison (Montague & Berns, 2002). Therefore, it is essential not only to show that the striatum codes the reward value of different stimuli, including social and monetary rewards, but also to establish that the neural comparison of actions takes place within the striatum to establish that it represents a "common currency." However, in the present study, we found that both social and monetary rewards were represented as a "decision utility" in the striatum. Thus, it is conceivable that the striatum might be involved not only in processing different rewards but also in comparing them. In other words, the strength of the activations in a given single trial might represent the difference between the values of the two actions ("donate" vs. "not donate") in that trial. In support of this theory, it was reported that the activity of neurons in the striatum in response to a certain reward cue (before the actual action occurs) was modulated by the value of the alternative reward in a given trial block, indicating that striatal neurons process the relative values of two rewards (Cromwell, Hassani, & Schultz, 2005). Samejima et al. (2005) also reported the existence of striatal neurons that encode the difference between two action values in the free-choice paradigm. However, the area where the comparison between two actions takes place in the brain should be tested systematically in future research.

The limitations of the present study should be noted. First, although we did not find any significant difference in RT, our data were not indicative of the actual decision time because subjects were instructed to press a button during the choice period (6 sec after a charitable organization was presented; Figure 2). Therefore, our RT data are more likely to be a simple reflection of motor readiness or attention level at the onset of the choice period. Allowing subjects to respond as soon as they made their decision would make it possible to gain a deeper insight into how individuals make prosocial decisions in front of others. Furthermore, it might also make it possible to obtain some evidence of where in the brain the difference between two actions is coded (i.e., where the values of two actions are compared) because the RTs should be correlated with the difference in value between two actions (e.g., if one action is clearly beneficial compared with another action, the RT should be faster than when two actions have similar values). Second, although it is plausible that people make a donation not only to gain extrinsic social reward but also to avoid extrinsic social cost (disapproval), the current research did not distinguish between these two motives. Previous neuroimaging studies by Knutson et al. reported that activities in the ventral striatum increased linearly as the magnitude of the expected monetary gain increased and were not influenced by the anticipation of avoiding monetary loss, whereas the dorsal striatum showed increased activations both when more monetary gain was expected and when more monetary loss was avoided (Knutson, Fong, Bennett, Adams, & Hommer, 2003; Knutson, Adams, Fong, & Hommer, 2001). Thus, the ventral striatum activities in the present study might reflect the utility of gaining social reward rather than avoiding social cost.

Finally, the present study demonstrated that the ventral striatum is the key brain structure involved in the valuation of diverse rewards, including money and social approval, during decision-making. It is plausible that social approval is incorporated as a decision utility in many social situations. As we are social animals, being watched by others is natural for humans, and the processing of this social reward seems critical in behaving not only prosocially but also appropriately in a given social situation (such as when talking to someone you respect, dressing for a job interview, etc.). Thus, our findings shed light on the neural mechanism of how the incentive of social approval influences people's contributions to public goods, which is a well-known behavioral tendency in the social sciences, and they also provide important insight into the neural basis of everyday human social decision-making.

Acknowledgments

We thank M. J. Hayashi, Y. Mano, and A. Sasaki for their assistance in conducting the experiments. This study was supported by Grant-in-Aid for Scientific Research S#17100003 to N. S. from the Japan Society for the Promotion of Science.

Reprint requests should be sent to Norihiro Sadato, Section of Cerebral Research, Division of Cerebral Integration, National Institute for Physiological Sciences (NIPS), Okazaki, Aichi, 444-8585, Japan, or via e-mail: sadato@nips.ac.jp.

REFERENCES

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biological Letters, 2,* 412–414.

Benabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review, 96,* 1652–1678.

Berns, G. S., McClure, S. M., Pagnoni, G., & Montague, P. R. (2001). Predictability modulates human brain response to reward. *Journal of Neuroscience, 21,* 2793–2798.

Blau, P. (1964). *Exchange and power in social life*. New York: John Wiley & Sons.

Cromwell, H. C., Hassani, O. K., & Schultz, W. (2005). Relative reward processing in primate striatum. *Experimental Brain Research, 162,* 520–525.

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Sciences, 1104,* 70–88.

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology, 84,* 3072–3077.

Elliott, R., Friston, K. J., & Dolan, R. J. (2000). Dissociable neural responses in human reward systems. *Journal of Neuroscience, 20,* 6159–6165.

Fehr, E., & Camerer, C. F. (2007). Social neuroeconomics: The neural circuitry of social preferences. *Trends in Cognitive Sciences, 11,* 419–427.

Friston, K. J., Holmes, A. P., & Worsley, K. J. (1999). How many subjects constitute a study? *Neuroimage, 10,* 1–5.

Glimcher, P. W., Dorris, M. C., & Bayer, H. M. (2005). Physiological utility theory and the neuroeconomics of choice. *Games and Economic Behavior, 52,* 213–256.

Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior, 26,* 245–256.

Harbaugh, W. T. (1998a). The prestige motive for making charitable transfers. *American Economic Review, 88,* 277–282.

Harbaugh, W. T. (1998b). What do donations buy? A model of philanthropy based on prestige and warm glow. *Journal of Public Economics, 67,* 269–284.

Harbaugh, W. T., Mayr, U., & Burghart, D. R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science, 316,* 1622–1625.

Holländer, H. (1990). A social exchange approach to voluntary cooperation. *American Economic Review, 80,* 1157–1167.

Homans, G. C. (1961). *Social behavior: Its elementary forms*. New York: Harcourt Brace Jovanovich.

Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron, 58,* 284–294.

Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics, 112,* 375–405.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience, 21,* RC159.

Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., & Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: Characterization with rapid event-related fMRI. *Neuroimage, 18,* 263–272.

Knutson, B., Westdorp, A., Kaiser, E., & Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. *Neuroimage, 12,* 20–27.

Kurzban, R., DeScioli, P., & O'Brien, E. (2007). Audience effects on moralistic punishment. *Evolution and Human Behavior, 28,* 75–84.

Lau, B., & Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron, 58,* 451–463.

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage, 19,* 1233–1239.

Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souza, R., & Grafman, J. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 15623–15628.

Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron, 36,* 265–284.

Redoute, J., Stoleru, S., Gregoire, M. C., Costes, N., Cinotti, L., Lavenne, F., et al. (2000). Brain processing of visual sexual stimuli in human males. *Human Brain Mapping, 11,* 162–177.

Rege, M., & Telle, K. (2004). The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics, 88,* 1625–1644.

Samejima, K., Ueda, Y., Doya, K., & Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science, 310,* 1337–1340.

Schultz, W., Tremblay, L., & Hollerman, J. R. (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex, 10,* 272–284.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: John Wiley & Sons.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science, 315,* 515–518.