Supporting Information

Izuma et al. 10.1073/pnas.1011879108

SI Results

Preference Change as the "Spreading of Alternatives." In past studies using the free-choice paradigm, preference change has typically been reported as the "spreading of alternatives," that is, the sum of the increase in rating for the chosen item and the decrease in rating for the rejected item between the first and second preference ratings. Our data show that the spreading of alternatives in the Post-Experimental Choice (PostEx-Choice) condition [mean (M) = 0.83] was significantly greater than that in the Self-Easy [$M = -0.44$, $t(19) = 7.94$, $P < 0.001$] and Computer conditions [$M = 0.16$, $t(19) = 4.90$, $P < 0.001$], suggesting that preferences for items rejected by subjects themselves decreases even in the absence of choice-induced dissonance. Importantly, however, our critical Self-Difficult condition ($M = 1.01$) showed a still greater spread compared with the PostEx-Choice condition [$t(19) = 1.80$, $P = 0.044$].

Hunger Rating. The mean hunger ratings were 6.00 at the beginning of the experiment and 6.95 at the end of the experiments using an eight-point scale [the difference was significant: $t(19) = 4.05$, $P < 0.001$], suggesting that our subjects placed high values on food stimuli throughout the experiment.

Reaction Time. The mean (± SD) reaction times (RTs) were 2,435 (± 560) ms during Preference task 1 and 2,105 (± 444) ms during Preference task 2 [the difference was significant; $t(19) = 4.62$, $P < 0.001$]. Fig. S1*A* plots the mean RTs of each condition in Preference task 2. The result of a 2 (subjects' past decision; Chosen or Rejected) × 3 (trial type; Self-Difficult, Self-Easy, or Computer) repeated-measures ANOVA revealed significant main effects of both subjects' past decision [$F(1,19) = 6.65$, $P = 0.018$] and trial type [$F(2,38) = 3.27$, $P = 0.049$]. However, the interaction was not significant [$F(2,38) = 1.56$, $P = 0.22$, n.s.]. Posthoc tests (with Bonferroni correction) revealed that subjects were faster to respond when they were presented with chosen items ($P = 0.018$), regardless of who chose them (Self or Computer). Although RTs in our critical condition (the Rejected-Self-Difficult condition, where the highest dissonance was expected) tended to be longer, suggesting a high degree of response conflict in this condition, there were no significant difference between this condition and the other two conditions in which rejected foods were also presented (Rejected-Self-Easy and Rejected-Computer conditions; both $P > 0.11$, n.s.) (Fig. S1*A*).

Correlational analyses revealed that in Preference task 2, subjects' RTs were weakly positively correlated with the degree of cognitive dissonance [the Cognitive Dissonance Index (CDI); *Methods*), and this correlation was significant at the group level [mean of 20 within-subject Pearson correlation coefficient, $r = 0.05$, $t(19) = 2.44$, $P = 0.025$, range -0.05 to 0.29]. In contrast, the same index was not correlated with RTs in Preference task 1 [$r = -0.01$, $t(19) = -0.28$, $P = 0.78$, n.s., range -0.16 to 0.15].

The RT data for the Choice task is presented in Fig. S1*B*. RTs for both the Self-Difficult and Self-Easy trials were longer than those in the Computer trials [vs. the Self-Difficult condition, $t(19) = 12.80$, $P < 0.001$; vs. the Self-Easy condition, $t(19) = 16.63$, $P < 0.001$]. During the Self trials, subjects took significantly longer to choose between two alternatives when they were close in preference (Self-Difficult trial) than when they were far apart [Self-Easy trial; $t(19) = 3.74$, $P = 0.001$].

Correlation Between Choices. For the food items presented in the Computer trials, choices were made once by a computer (in the Choice task) and once by the subject (in the PostEx-Choice task). However, these two choices were not correlated with each other [mean within-subject correlation: $r = 0.02$, $t(19) = 0.33$, $P = 0.74$, n.s.].

Brain-Behavior Correlations. We investigated the relationship between brain activation and self-report preference changes in our critical condition (i.e., the Rejected-Self-Difficult condition). We extracted each subject's β-value for this condition [general linear model (GLM) 2] from the peak voxel in the dorsal anterior cingulate cortex (dACC) ($x = 12$, $y = 30$, $z = 46$) and found that those who reduced their preference more showed less activation in the dACC ($r = 0.60$, $P = 0.005$, two-tailed) (Fig. S5). However, as Chen and Risen (1) pointed out, self-report preference changes seen in the Rejected-Self-Difficult condition reflect not only dissonance (choice)-induced preference changes but also changes explained by the preference information revealed by subjects' choices. Therefore, we used preference change in the Rejected-Self-Difficult condition minus that in the Rejected-PostEx-Choice condition as each subject's measure of dissonance-induced preference change. With this individual self-report measure of preference change, we still found significant positive correlation with dACC activation ($r = 0.44$, $P = 0.05$, two-tailed). These correlations remain significant even when controlling for mean preference rating in Preference task 1 and mean reaction time (all $P < 0.024$). Among areas correlated with the CDIs during Preference task 2 (Table S3), the dACC was the only region showing significant positive correlation in both of these self-report measures of preference change (Table S4).

Interestingly, whereas a high degree of cognitive dissonance leads to more preference change (i.e., reduced preference) and higher dACC activation on a within-subject basis, we found the opposite relationship between preference change and dACC activation on a between-subject basis. Therefore, while the dACC activation reflects perceived cognitive dissonance on a within-subject basis, this positive between-subject correlation suggests that individual differences in dACC activation reflect how much each subject resolved their dissonance by changing their preference. In other words, when subjects were faced with inconsistency between their past behavior (e.g., "I rejected it") and cognition (e.g., "I like it"), those who resolved the dissonance by reducing their preference showed less dACC activation. Interestingly, this correlation between dACC activations and preference (or attitude) changes is the opposite of what van Veen et al. (2) reported. In their study, the more dACC is activated, the more subjects change their attitude. In contrast, in the present study, those who exhibited more preference change showed less dACC activity.

This difference in the direction of correlation seems to be attributable to the important difference in the experimental paradigm in these two studies. In the previous study (2), dACC activity was observed when subjects were making counter attitudinal arguments, and attitude change was measured after the fMRI experiment. Therefore, subjects had no explicit opportunity to reduce their perceived dissonance during the fMRI scanning. On the other hand, in the present study, dissonance was induced by presenting subjects with food items that they preferred but had rejected in the past, and at the same time, they were asked to rate their preferences for the foods. Therefore, subjects were given the opportunity to reduce the dissonance they just felt by explicitly reporting their new preference. In other words, it might be the case that while van Veen et al. (2)

focused on passive emotional reaction to perceived cognitive dissonance, we focused on the active dissonance reduction process following dissonance perception [this idea also seems to be consistent with the stronger dorsolateral prefrontal cortex (DLPFC) activation in the present study because the DLPFC is known to be involved in conflict resolution]. Together with the previous report (2), these brain-behavior correlations suggest that when subjects' brains are scanned during the induction of dissonance in isolation, the dACC activity predicts how much subjects change their attitude later on. In contrast, when subjects were given the chance to reduce dissonance at the same time, individual differences in dACC activation reflect how much it was reduced by bringing their attitude in line with their past behaviors. However, the exact nature of how individual differences in ACC activity explain preference change should be further investigated in future research.

SI Methods

Subjects. A total of 24 healthy right-handed subjects participated in the study. The reported analyses were based on 20 subjects (10 male; age range, 18–24 y). Two subjects were excluded from the analyses due to excessive head motion (one of them also did not follow the instructions correctly during the PostEx-Choice task). One subject was excluded because responses (and reaction time) were not recorded due to a button malfunction. One additional subject was not included in the analysis because after the experiment, he reported losing his appetite due to repeated exposure to sweet food items that he disliked; his hunger ratings went down by 5 (i.e., he felt less hungry after the experiment), which is >5 SD lower than the mean change in the hunger ratings for the remaining subjects.

Experimental Stimuli. A total of 160 foods were used in the experiments (e.g., chips, chocolate, ice cream, cake, etc.). Throughout the experiment, subjects were presented with color pictures of these food stimuli and their names. All of the foods were available at local grocery stores, and subjects were fairly familiar with the most of the items (the mean familiarity rating was 5.10 on an eight-point scale: 1 = not familiar at all, 8 = highly familiar).

Experimental Tasks and Procedures. Subjects were asked not to eat for at least 3 h before the experiment to increase their preference for the food items presented during the experiment, and thus make their decisions during the Choice task more salient. At the beginning of the experiment, subjects were first asked to rate their hunger on an eight-point scale (1 = not hungry at all, 8 = very hungry).

Following the hunger rating, subjects performed Preference task 1, Choice task, and Preference task 2 inside the fMRI scanner. Then, after these tasks, they also performed the PostEx-Choice task outside the fMRI scanner. During Preference task 1 (Fig. 1*A*), subjects were instructed to rate their preference for each food on an eight-point scale. Responses were made using two button boxes with four buttons each, and subjects used their left hand for ratings one to four and right hand for ratings five to eight. Each food was presented for 3 s, and the order in which food stimuli were presented was random. The intertrial interval (ITI) was jittered between 5 and 9 s.

In the Choice task (Fig. 1*B*), subjects were instructed to use the right index or middle finger, which spatially corresponded to the location of the foods. At the beginning of each of the Self and Computer trials, a cue indicating which of two trials to be performed was presented for 1.5 s and, after a brief delay of 500 ms, two foods were presented for 5 s. The Self and Computer trials were pseudorandomly intermixed with the ITI jittered between 5 and 9 s. Because food pairs in the Choice task depended on how each subject rated 160 foods in Preference task 1, the number of trials for each of Self-Difficult, Self-Easy, and Computer trial varied for each subject. The maximum number of trials in each condition was set to 20 trials; on average, there were 14.9 Self-Difficult trials, 16.2 Self-Easy trials, and 14.9 Computer trials (the number of trials in Self-Difficult and Computer conditions was matched within each subject).

In Preference task 2 (Fig. 1*C*), subjects were asked to rate their preference for each food again using the same eight-point scale. While subjects' (or computer's) past decisions (e.g., "You chose it," "You rejected it," "Computer chose it," or "Computer rejected it") during the Choice task were presented under each food, they were instructed not to pay attention to this information and to concentrate on reporting their preference for each food.

After Preference task 2, subjects took part in the PostEx-Choice task outside the fMRI scanner. During this task, the same pairs of foods that had appeared during the Computer trials of the Choice task were presented on the screen of a desktop computer, and subjects were asked to choose the one they preferred by pressing one of two buttons on the keyboard. Subjects were told that after the experiment, one trial from the Choice task or the PostEx-Choice task would be randomly selected, and the food chosen by themselves (or a computer if the chosen trial was a Computer trial) in that trial would be given to them at the end of the experiment. Finally, subjects rated the familiarity of each of 160 foods, which were presented in random order on the same computer screen, using an eight-point scale, and they again rated how hungry they felt at that time.

Functional MRI Data Acquisition. The functional imaging was conducted using a 3 Tesla Siemens Trio A Tim MRI scanner. For functional imaging during the sessions in both experiments, interleaved T2*-weighted gradient-echo echo-planar imaging (EPI) sequences were used to produce 42 continuous 3-mm thick transaxial slices covering nearly the entire cerebrum (repetition time = 2,500 ms; echo time = 25 ms; flip angle = 90°; field of view = 192 mm^2; 64 × 64 matrix; voxel dimensions = 3.0 × 3.0 × 3.0 mm). A high-resolution anatomical T1-weighted image was also acquired for each subject.

Functional MRI Data Preprocessing. Before data processing and statistical analysis, we discarded the first four volumes to allow for stabilization of the magnetization. The data were analyzed using Statistical Parametric Mapping 8 (SPM8, Wellcome Department of Cognitive Neurology, London, United Kingdom) software implemented in Matlab 7.8 (Mathworks). After correcting for differences in slice timing within each image volume, head motion was corrected using the realignment program of SPM8. Following realignment, the volumes were normalized to the Montreal Neurological Institute (MNI) 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 8 mm (full-width at half-maximum) in the x, y, and z axes.

Functional MRI Data Analysis. To define the striatum region of interest (ROI), an anatomical mask was used to limit our ROI to be inside the anterior part (y coordinate ≥ 0) of the striatum. The anatomical mask of the striatum (caudate and putamen bilaterally with a dilation factor of 2) was first generated and then its posterior part (areas whose y coordinate was < 0) was excluded by using the WFU PickAtlas toolbox for SPM (3).

For the ROI analysis of fMRI data, as well as the analysis of behavioral data for preference change, we used one-tailed paired t tests because of our a priori hypotheses about the specific direction of preference changes. For all other analyses, including RT data and the brain-behavior correlations, two-tailed P values are reported.

1. Chen MK, Risen JL (2010) How choice affects and reflects preferences: Revisiting the free-choice paradigm. *J Pers Soc Psychol* 99:573–594.
2. van Veen V, Krug MK, Schooler JW, Carter CS (2009) Neural activity predicts attitude change in cognitive dissonance. *Nat Neurosci* 12:1469–1474.
3. Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19:1233–1239.

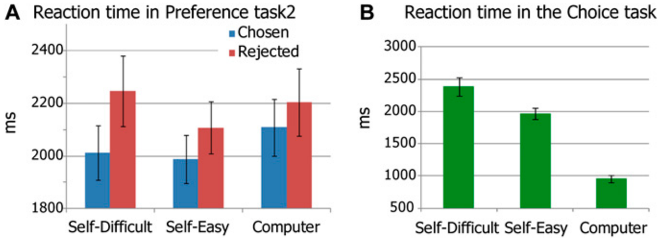

Fig. S1. Reaction time data in Preference task 2 (*A*) and the Choice task (*B*). Error bars depict the SEM.

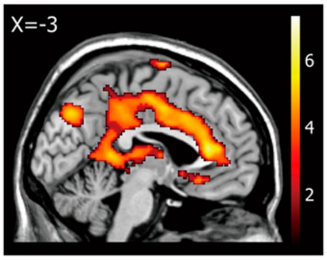

Fig. S2. Areas correlated with subjects' reported preference during Preference task 1. Almost the entire cingulate cortex (from anterior to posterior) was activated. All activations are listed in Table S1. A statistical threshold was set at $P < 0.001$ for height (uncorrected) and cluster $P < 0.05$ (corrected for multiple comparisons). The scale shows the *t* values.

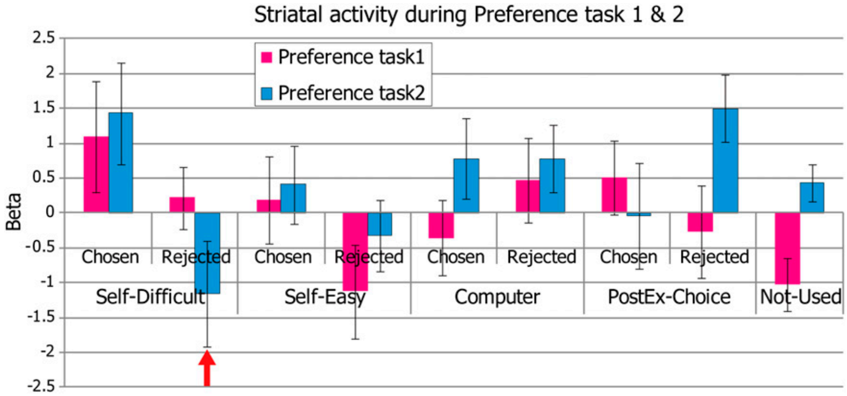

Fig. S3. The striatal ROI (Fig. 3*B*) activity in all conditions during Preference task 1 and 2. Note that in almost all conditions except for the Rejected-Self-Difficult condition (red arrow), activations increased from Preference task 1 to Preference task 2, most likely because of increased hunger levels (*SI Results*). When compared with the Not-Used condition, which controlled for a general change in striatal sensitivity to food stimuli, our critical condition (the Rejected-Self-Difficult condition) was the only one that showed a significant 2 (condition) × 2 (session; preference task 1 vs. 2) interaction [$F(1,19) = 7.48$, $P = 0.013$]. Error bars depict the SEM.

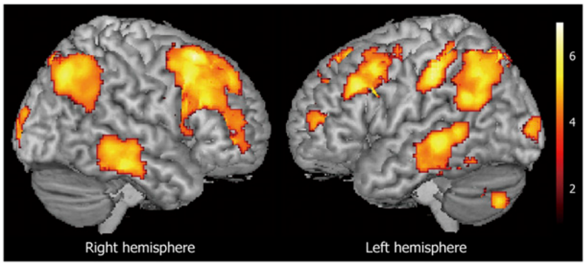

Fig. S4. Areas positively correlated with the degree of cognitive dissonance (CDI) during Preference task 2. Note that this contrast was explored within the areas not correlated with CDIs during Preference task 1 at $P < 0.05$ (uncorrected) level. See also Table S3. A statistical threshold was set at $P < 0.001$ for height (uncorrected) and cluster $P < 0.05$ (corrected for multiple comparisons). The scale shows the *t* values.

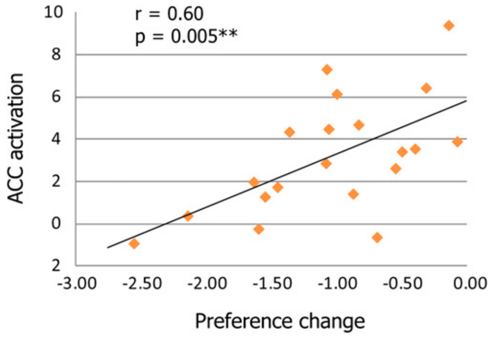

Fig. S5. A brain-behavior correlation plot between self-report preference changes (the mean preference ratings in Preference task 2 minus those in Preference task 1) and ACC activations in the Rejected-Self-Difficult condition.

Table S1. Areas correlated with self-reported preference during preference task 1

		MNI coordinate				
Location	BA	*x*	*y*	*z*	Z	Cluster size
Right vmPFC	11	8	38	−16	3.68	12,875
Left vmPFC	11/25	−8	22	−12	4.22	
Left anterior striatum		−6	22	−10	3.63	
Rostral ACC	24/32	0	18	30	4.30	
Left posterior putamen		−24	−10	−8	5.29	
Left precentral gyrus	4/6	−32	−16	70	6.73	
Left thalamus		−14	−20	2	5.02	
Left superior temporal gyrus	22	−48	−26	20	4.64	
Paracentral lobule/PCC	23/31	−8	−30	52	5.06	
Left postcentral gyrus	1/2/3	−34	−36	70	5.16	
Left SPL	7	−6	−74	38	4.65	
Right anterior striatum 1		8	14	0	3.58	53
Right anterior striatum 2		10	2	−8	3.72	44
Left superior frontal gyrus	8/9	−24	38	46	4.08	300
Right cerebellum		26	−54	−22	5.70	2,218

ACC, anterior cingulate cortex; BA, Brodmann area; PCC, posterior cingulate cortex; SPL, superior parietal lobule; vmPFC, ventromedial prefrontal cortex. Activations in the striatum ROI are in italics. Within the striatum, the statistical threshold was set at $P < 0.001$ (uncorrected for multiple corrections) with an extent threshold of 20 contiguous voxels. Outside the striatum, the statistical threshold was $P < 0.001$ for height (uncorrected) and cluster $P < 0.05$ (corrected for multiple comparisons).

Table S2. The CDI used in GLMs 4 and 5

Condition	Preference rating	CDI
Chosen-Self-Difficult and Chosen-Self-Easy condition	8	1
	7	2
	6	3
	5	4
Rejected-Self-Difficult condition	8	8
	7	7
	6	6
	5	5
Rejected-Self-Easy condition	4	4
	3	3
	2	2
	1	1
Chosen-Computer condition	5–8	0
Rejected-Computer condition	5–8	0
Not-Used condition	1–8	0

Note that preference ratings in Preference task 1 were used to compute the CDI.

Table S3. Areas positively correlated with CDIs during Preference Task 2

Location	BA	MNI coordinate *x*	*y*	*z*	Z	Cluster size
Left IFG	10	−34	50	12	3.78	348
Dorsal ACC	32	12	30	46	4.81	6,131
Right IFG	10	26	46	−2	4.34	
Right DLPFC	8/9/46	40	26	24	4.68	
Left DLPFC	9/46	−30	10	40	4.69	1,978
Left middle temporal gyrus	21	−52	−32	−4	4.92	1,371
Right middle temporal gyrus	21	64	−36	−10	4.05	984
Left cerebellum		−32	−58	−38	4.15	369
Right SPL	7/31	10	−74	50	5.08	9,608
PCC	23	0	−30	30	4.36	
Left IPL	39/40	−34	−54	36	4.93	
Right IPL	39/40	44	−56	44	4.09	
Left SPL	7/31	−14	−62	44	5.07	
Left primary visual cortex	17	−8	−100	4	4.03	912
Right primary visual cortex	17	8	−102	8	3.87	

Note that this contrast was explored outside of areas correlated with CDIs during Preference task 1 with a threshold of $P < 0.05$ (uncorrected). ACC, anterior cingulate cortex; BA, Brodmann area; DLPFC, dorsolateral prefrontal cortex; IFG, inferior frontal gyrus; IPL, inferior parietal lobule; PCC, posterior cingulate cortex; SPL, superior parietal lobule. Activation in the ACC ROI is in italics. Within the ACC, the statistical threshold was set at $P < 0.001$ (uncorrected for multiple corrections) with an extent threshold of 20 contiguous voxels. Outside the ACC, the statistical threshold was $P < 0.001$ for height (uncorrected) and cluster $P < 0.05$ (corrected for multiple comparisons).

Table S4. Brain-behavior correlations between preference changes and brain activities in the Rejected-Self-Difficult condition

Location	Preference change in self-report measures: Rejected-Self-Difficult	Rejected-Self-Difficult minus Rejected-PostEx-Choice
Right IFG	0.07	−0.20
Left IFG	−0.05	0.03
Dorsal ACC	0.60**	0.44*
Right DLPFC	−0.03	−0.25
Left DLPFC	0.26	0.42
PCC	−0.45*	−0.17
Right middle temporal gyrus	0.04	−0.09
Left middle temporal gyrus	−0.11	0.29
Right IPL	−0.09	−0.14
Left IPL	0.13	0.03
Left cerebellum	−0.38	−0.21
Right SPL	0.15	−0.08
Left SPL	0.22	0.44†
Right primary visual cortex	0.20	0.23
Left primary visual cortex	0.13	0.10

Brain activations (β values) for each region were extracted from the peak coordinates reported in Table S3. The ACC ROI is in italics.
$^{\dagger}P < 0.1$; $*P < 0.05$; $**P < 0.01$ (two-tailed).