Behavioral/Cognitive

Neural Mechanisms of Gain–Loss Asymmetry in Temporal Discounting

Saori C. Tanaka,[1,2] **Katsunori Yamada,**[2,3] **Hiroyasu Yoneda,**[4] and **Fumio Ohtake**[2]

[1]Department of Cognitive Neuroscience, ATR Brain Information Communication Research Laboratory Group, Kyoto 619-0288, Japan, [2]Institute of Social and Economic Research, Osaka University, Osaka 567-0047, Japan, [3]Faculty of Economics, Kindai University, Higashi-Osaka 577-0813, Japan, and [4]Pharmaceutical Policy & Health Economics, Kyoto University, Kyoto 606-8501, Japan

Humans typically discount future gains more than losses. This phenomenon is referred to as the "sign effect" in experimental and behavioral economics. Although recent studies have reported associations between the sign effect and important social problems, such as obesity and incurring multiple debts, the biological basis for this phenomenon remains poorly understood. Here, we hypothesized that enhanced loss-related neural processing in magnitude and/or delay representation are causes of the sign effect. We examined participants performing intertemporal choice tasks involving future gains or losses and compared the brain activity of those who exhibited the sign effect and those who did not. When predicting future losses, significant differences were apparent between the two participant groups in terms of striatal activity representing delay length and in insular activity representing sensitivity to magnitude. Furthermore, participants with the sign effect exhibited a greater insular response to the magnitude of loss than to that of gain, and also a greater striatal response to the delay of loss than to that of gain. These findings may provide a new biological perspective for the development of novel treatments and preventive measures for social problems associated with the sign effect.

Key words: delay discounting; fMRI; insula; loss aversion; sign effect; striatum

Introduction

Humans exhibit a temporal preference known as "delay discounting," whereby a sooner reward is typically preferred to a future reward of equal magnitude (Ainslie, 1975). Value that is discounted to account for a time delay is expressed as *discounted value = u* (outcome) $\times$ *g* (delay), where *u* is the utility function and *g* is the discount function. Previous studies in experimental and behavioral economics have reported an asymmetry in the discounted value assigned to gains and losses, whereby the value of a future loss appears to be less steeply discounted than the value of a future gain (Thaler, 1981; Loewenstein, 1987, 1988; Frederick et al., 2002). This gain–loss asymmetry in delay discounting is referred to as the "sign effect." Recent economic studies have reported that the lack of the sign effect was associated with social problems, such as smoking (Odum et al., 2002), debt, and obesity (Ikeda et al., 2010).

Theoretical work has suggested that a sign effect could result from a value function that is more elastic for losses than that for

gains, and a combination of loss aversion and the magnitude effect (Loewenstein and Prelec, 1992). Here, we focused on loss aversion as one potential source of the sign effect. It has been suggested in theoretical studies of loss aversion that the loss of utility from a specific magnitude of loss is greater than the gain of utility from the same magnitude of gain, when they are taken at the same timing (Kahneman and Tversky, 1984). The greater utility can generate greater discounted value, which is correlated with more gradual discounting. Hence, loss aversion can cause the sign effect. At a neural level, loss has been shown to be neurally coded in the insula and striatum. The insula is activated when predicting monetary loss (Paulus et al., 2003; Pessiglione et al., 2006; Palminteri et al., 2012), and a recent study reported that individual insular activity responding to monetary losses reflected the individual degree of loss aversion in economic gambling tasks (Canessa et al., 2013). In the striatum (commonly activated by primary and monetary rewards), a positive relationship between striatal BOLD response and monetary loss has also been observed. For instance, in tasks with both gains and losses in primary and monetary outcomes, a correlation between striatal activity and negative outcomes has been observed when participants are forced to receive negative outcomes while the potential to receive positive outcomes is present (Seymour et al., 2004, 2007b). These findings on gain and loss coding indicate the enhanced responses to loss in the insula and striatum.

Another potential source of the sign effect is that salient delayed losses may create a "dread effect." People sometimes prefer to receive aversive outcomes, such as pain and monetary loss sooner rather than later. This type of preference has been termed "dread" (Loewenstein, 1987; Loewenstein and Prelec, 1991, 1993;

Story et al., 2013). According to economic theory, this preference can be explained by an increasing discount function with time, whereby future outcomes are weighted more than immediate outcomes because anticipation itself harbors disutility. This pattern of time discounting can be a cause of the sign effect (Loewenstein, 1987). At a neural level, delay is suggested in some studies to correlate negatively striatal activity; for example, striatal activity correlates negatively with length of delay (Gregorios-Pippas et al., 2009) and correlates positively with hyperbolic discount functions that decrease with increasing length of the delay (Pine et al., 2009; Pine et al., 2010). These studies suggest downward sloping discount functions where future outcomes are discounted more. However, one study showed that subjects under the influence of dread exhibited a time discounting pattern where they weighted future outcomes more; that is, they exhibited a positive correlation between the length of delay to an electric shock and neural activities in anticipation of the shock in the insula (Berns et al., 2006).

Together, economic theory suggests two possible explanations for the sign effect: loss aversion and dread. The loss aversion hypothesis would suggest that the sign effect is related to an enhanced neural response to the magnitude of losses relative to gains. The dread hypothesis would suggest that the sign effect is related to an enhanced neural response to delay for losses relative to gains. Here, we tested these hypotheses by examining activity in insula and striatum with an intertemporal choice task involving losses and gains (Fig. 1). We then quantitatively compared brain activity representing magnitude and delay information at the time of decision making between those participants who exhibited the sign effect and those who did not. In doing so, we managed to identify the effects of magnitude and delay components of future gains and losses in a single experimental setting.

Materials and Methods

Participants. We initially conducted an fMRI experiment with 32 participants and identified those who exhibited the sign effect and those who did not. Of these 32 volunteer participants (18 males and 14 females), 3 did not exhibit the sign effect. To obtain a greater number of participants without the sign effect, we ran an additional behavioral experiment with 60 additional volunteers (30 males and 30 females) using the same methods as in the first fMRI experiment. We found 9 did not exhibit the sign effect, and these 9 were added to the original 32 volunteers for a total of 41 participants who went through to our fMRI experiment. These 41 healthy, right-handed volunteers (23 males and 18 females; age 26.1 ± 7.0 years; age range 20–52 years) consented to participate in the experiment, which was conducted with approval from the Review Board Ethics and Safety Committees for Functional Resonance Imaging Research of the Advanced Telecommunication Research Institute International and Osaka University. From these 41, we excluded 3 participants before behavioral data analysis because of problems acquiring imaging or behav-
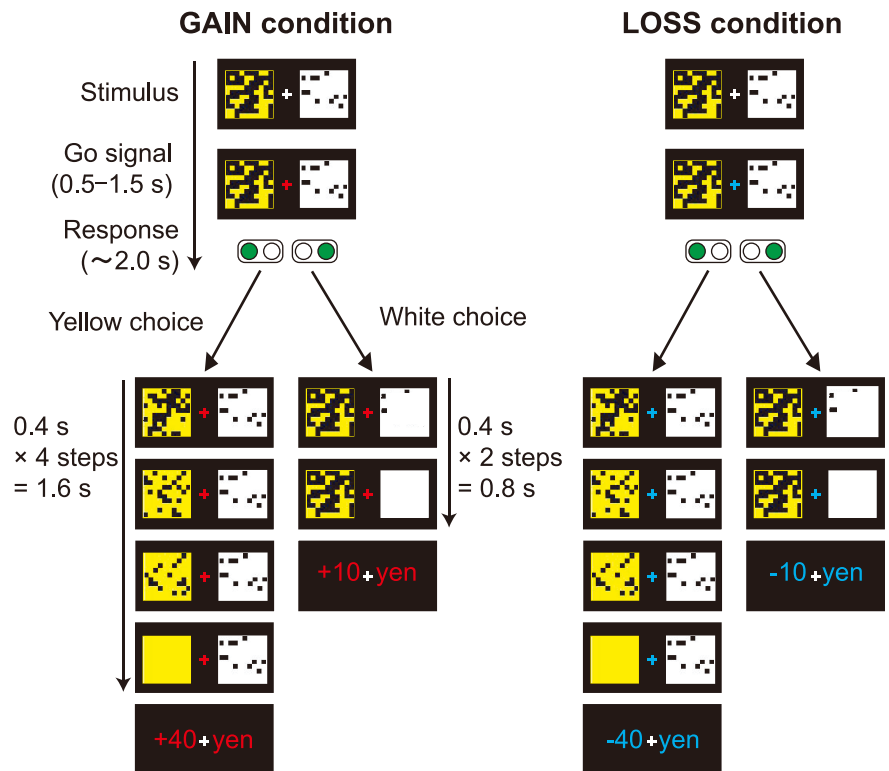

Figure 1. Experimental task. At the beginning of each trial, a white square (small-sooner option) and a yellow square (large-later option) were displayed side by side on a screen for 2000 ms. The squares were occluded by a variable number of black patches corresponding to the length of monetary payoff delays. In the GAIN condition, participants chose between a white square indicating a small gain (10 yen) with a shorter delay (1–7 s) and a yellow square indicating a large gain (40 yen) with a longer delay (6–26 s). In the LOSS condition, participants chose between a small loss (−10 yen) with a shorter delay and a large loss (−40 yen) with a longer delay. Once a square had been selected by pressing a button, the black patches were removed from the chosen square at 400 ms intervals. When the square was completely exposed, the monetary payoff was displayed on the screen for 1000 ms. In the next trial, white and yellow squares were displayed with novel mosaic patterns. In the example shown, choosing a white square led to a small monetary gain (10 yen) or a small monetary loss (−10 yen), which was delivered in two steps (0.4 × 2 = 0.8 s) in the GAIN or LOSS condition, respectively. When choosing a yellow square, participants waited four steps (0.4 × 4 = 1.6 s) to obtain a larger monetary gain (40 yen) or a larger loss (−40 yen) in the GAIN or LOSS condition, respectively.

ioral data during the experimental trials. We also excluded 6 participants before imaging data analysis because they exclusively chose the small-sooner or larger-later option in the experimental task; because we used a method of indifference points (see Behavioral data analysis), we were unable to estimate the discounting parameter of these participants when they had consistently chosen a single option. Thus, in total, behavioral and imaging data from 32 participants (19 males and 13 females; age 26.6 ± 7.4 years; age range 20–52 years) were analyzed in this study: 23 participants with the sign effect and 9 participants without.

Experimental task. We used a repeated intertemporal choice task (Fig. 1) in which participants were asked to choose between a smaller-sooner and a large-later option. In the original task used in our previous studies (Schweighofer et al., 2006; Tanaka et al., 2007), only options with potential gains were presented. In the present study, we added a novel condition in which participants chose between small-sooner and large-later losses. At the beginning of each trial, a white square (small-sooner option) and a yellow square (large-later option), which were occluded by a variable number of black patches corresponding to the length of monetary payoff delays, were displayed side by side on a screen for 2000 ms. In the GAIN condition, participants chose between a white square indicating a small gain (10 yen) with a short delay and a yellow square indicating a large gain (40 yen) with a longer delay. In the LOSS condition, participants chose between a small loss (−10 yen) with a short delay and a large loss (−40 yen) with a longer delay. Once a square had been selected by pressing a button, the black patches started being removed sequentially. The number of patches taken away at each time interval (400 ms) was a random variable S, taking the value of 3, 4, 5, 6, or 7. S is uniformly
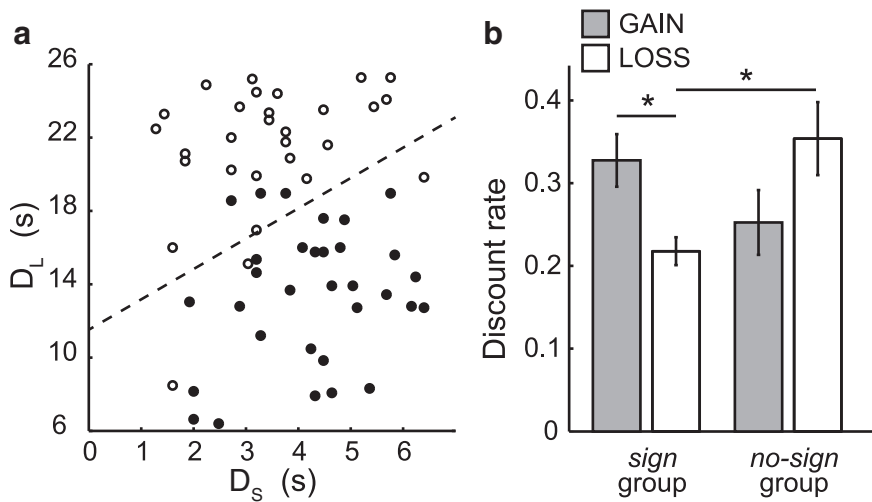
Figure 2. **a**, Choice pattern of a representative participant in the GAIN condition. The delay length of the large-later option (D_L) is plotted against the delay length of the small-sooner option (D_S) for each choice. Filled circles represent selection of the large-later option (yellow square: 40 yen); open circles represent selection of the small-sooner option (white square: 10 yen). The dotted line indicates the estimated indifference line of the participant. The intercept of the indifference line is the point where a small-sooner option with delay = 0 is equivalent to a large-later option with the delay of the intercept. **b**, The *sign* group exhibited a smaller discount rate (slower discounting) in the LOSS condition than in the GAIN condition. *$p < 0.05$ (multiple comparison with Tukey–Kramer method). In the LOSS condition (white bars), there was a significantly smaller discount rate in the *sign* group than in the *no-sign* group. *$p < 0.05$ (multiple comparison with Tukey–Kramer method). Error bars indicate SE.

distributed. When the white or yellow square image was completely exposed, the monetary payoff was displayed on the screen for 1000 ms. In the next trial, the white and yellow squares were displayed with novel mosaic patterns of black patches.

The initial numbers of black patches against white and yellow squares, termed M_S for small-sooner options and M_L for large-later options, were random variables and followed uniform distributions. M_S had an average value of 48, a minimum value of 16, and a maximum value of 80; M_L had an average value of 200, a minimum value of 80, and a maximum value of 320. The expected delays to small outcomes and large outcomes, termed D_S and D_L, respectively, are given by the following:

$$D_S = \tau^* M_S / mean(S) \text{ and } D_L = \tau^* M_L / mean(S),$$

where τ is the time interval of 400 ms. Thus, the ranges of expected delays for the average values of D_S and D_L were as follows: $(D_S, D_L) = (1.2–10.8$ s, $4.8–42.8$ s). The delay time associated with a small reward/loss was typically shorter than the delay length associated with a large reward/loss. Thus, at the beginning of each trial, participants were required to choose between a sooner but smaller reward/loss (white) and a later but larger reward/loss (yellow) by comparing the patterns of black patches on the two squares. The participants were required to respond within 2000 ms, so that decision making was based on the visual impression of darkness on the squares rather than on an explicit count of the black patches. Each GAIN and LOSS condition lasted ∼20 min. The order of the conditions was counterbalanced across participants.

Procedure. Before the fMRI experiment, the 32 participants initially recruited practiced the identical experimental task they performed in the fMRI scanner, except that the initial number of black patches was different in each trial. The second 9 participants who did not exhibit the sign effect in the additional behavioral experiment were exempt from this practice on the day of fMRI experiment because they had learned the task. All participants received a cash honorarium immediately after finishing the tasks in both the GAIN and LOSS conditions. Participants who completed the fMRI experiment received a participation fee (10,000 yen, equivalent to ∼100 U.S. dollars) as well as the total amount earned in the GAIN and LOSS conditions, minus withholding tax (10%). Participants earned an average of 1559.7 yen (in the GAIN condition) and lost an average of 1134.4 yen (in the LOSS condition). Therefore, the participants in the fMRI experiments received, in total, an average of 9382.8

yen. When we recruited the second group of participants in the behavioral tasks, participants received an average of 2000.3 yen. We informed participants approximately the honorarium while describing the experimental methods.

Behavioral data analysis. The degree of discounting was calculated for each participant using his or her individual choice patterns with a method of indifference points (Ho et al., 1999). The delay length of the large-later option (D_L) was plotted against the delay length of the small-sooner option (D_S) for each choice (for the GAIN condition, see Fig. 2a). The probability of choosing the large-later option (P_L) was fitted using the following logistic regression analysis (Schweighofer et al., 2006; Tanaka et al., 2007):

$$P_L = 1/(1 + exp[-(\beta_L D_L + \beta_S D_S + \beta_0)]).$$

$$(1)$$

When $P_L = 0.5$, this equation was transformed to:

$$D_L = -\beta_S/\beta_L {}^* D_S - \beta_0/\beta_L. \qquad (2)$$

For any point on this "indifference line," the discounted values of the small-sooner and large-later options were equal. The intercept of the indifference line $-\beta_0/\beta_L$ is the point where a small-sooner option with delay = 0 is equivalent to a large-later option with a delay value equivalent to the intercept.

By applying the method of indifference points to the specific discount model, we can compute an individual discount rate from the estimated intercept. The theoretical model of discounting in the hyperbolic style $g(D) = 1/(1 + k_h D)$ can be transformed to the same form as the indifference line equation in the following way:

$$1/(1 + k_h D_L) {}^* u(40) = 1/(1 + k_h D_S) {}^* u(10)$$

$$D_L = bD_S + c/k_h \text{ in the hyperbolic model}, \qquad (3)$$

$$b = u(40)/u(10), \ c = (u(40) - u(10))/u(10).$$

$u(40)$ and $u(10)$ is utility for 40 yen and 10 yen, respectively. We adopted a linear utility function here; thus, $u(40) = 40$, and $u(10) = 10$. k_h is the discount rate and is included in the intercept: the intercept in Equation 2, $-\beta_0/\beta_L$, reflects the "inverse discount rate" c/k_h in Equation 3. To categorize the participants according to the existence of the sign effect, we computed the individual discount rate as $k_h = -\beta_0/c\beta_L$. To investigate the detailed relationship between neural activity and behavioral parameters (Fig. 3e,f), we defined the degree of the sign effect as the difference between the intercepts of the indifference line in the GAIN and LOSS conditions. We suggest that the obtained intercept parameters accurately reflected individual neural features because the abovementioned method of indifference points avoids potential data deterioration that can be caused by using specific discounting functions (Gregorios-Pippas et al., 2009).

Imaging acquisition. A 3.0-Tesla scanner (Magnetom Trio, A Tim System; Siemens Japan) was used to acquire structural T1-weighted images and T2*-weighted echo planar images (repetition time = 2000 ms; echo time = 30 ms; flip angle = 80°; matrix = 64 × 64; field of view = 192 mm; slice thickness = 4 mm; slice gap = 0 mm) with BOLD contrast. SPM8 software (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London) was used for preprocessing and statistical analysis. The first six image volumes were discarded to avoid T1 equilibrium effects. The images were realigned to the first image as a reference, spatially normalized with respect to the MNI echo planar imaging template, and spatially smoothed with a Gaussian kernel (8 mm, full width at half maximum).
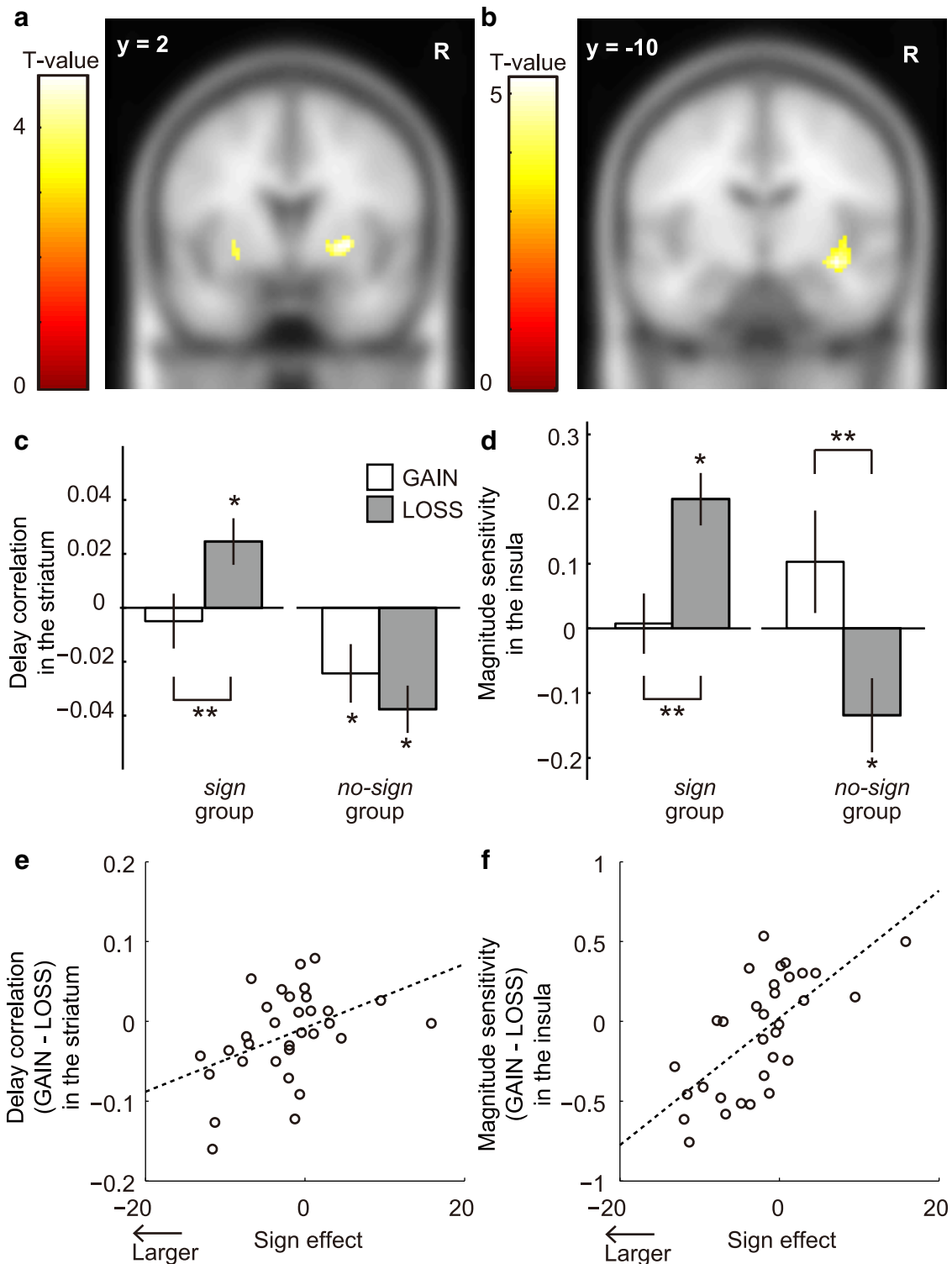
Figure 3. ***a***, In the striatum, we found a significant difference between the *sign* and *no-sign* groups in the LOSS condition in terms of the correlation between BOLD signals and delay length. ***b***, In the insula, we found a significant group difference in the LOSS condition in terms of magnitude sensitivity. ***c***, Delay correlation in the striatum in the GAIN (white bars) and LOSS (gray bars) conditions. ***d***, Magnitude sensitivity in the insula in the GAIN (white bars) and LOSS (gray bars) conditions (*sign*: $n = 23$; *no-sign*: $n = 9$). ***p < 0.05 (multiple comparison with Tukey–Kramer method). *p < 0.05 (one-sample t test). ***e***, Differences between the behaviorally measured discount rates in the GAIN and LOSS conditions and between the striatal delay correlations in the GAIN and LOSS conditions were significantly correlated ($r = 0.437$, $p < 0.05$). ***f***, Differences between the behaviorally measured discount rates in the GAIN and LOSS conditions and between the left insular magnitude sensitivity in the GAIN and LOSS conditions were significantly correlated ($r = 0.673$, $p < 0.05$). Error bars indicate SE.

Imaging data analyses. In the fMRI data analyses, we prepared an event-related design matrix for the GAIN and LOSS sessions. The design matrix contained the following five regressors for each of the GAIN and LOSS conditions: (1) an event regressor (δ-function) at the timing of stimulus display; (2) a parametric modulation of event regressor 1 by the delay length of the chosen option; (3) a parametric modulation of event regressor 1 by the outcome magnitude (10 or 40) of the chosen option; (4) an event regressor at the time the black patches were removed (every

Table 1. Results of two-sample *t* test for delay correlation between the *sign* and *no-sign* groups[a]

Contrast	Area	Coordinates	Voxels	*t* value
GAIN condition				
(*Sign*) − (*no-sign*)	None	—	—	—
(*No-sign*) − (*sign*)	None	—	—	—
LOSS condition				
(*Sign*) − (*no-sign*)	Striatum	28, 2, −2	130	4.77
(*No-sign*) − (*sign*)	None	—	—	—

[a]Corrected threshold of $p < 0.05$ (small-volume correction) for voxels with activation exceeding the uncorrected height threshold $p < 0.001$ and no extent threshold. All *xyz* coordinates correspond to the MNI coordinates system. —, Not applicable.

Table 2. Results of two-sample *t* test for magnitude sensitivity between the *sign* and *no-sign* groups[a]

Contrast	Area	Coordinates	Voxels	*t* value
GAIN condition				
(*Sign*) − (*no-sign*)	None	—	—	—
(*No-sign*) − (*sign*)	None	—	—	—
LOSS condition				
(*Sign*) − (*no-sign*)	Right insula	36, −10, −10	376	5.01
(*No-sign*) − (*sign*)	None	—	—	—

[a]Corrected threshold of $p < 0.05$ (small-volume correction) for voxels with activation exceeding the uncorrected height threshold of $p < 0.001$ and no extent threshold. All *xyz* coordinates correspond to the MNI coordinates system. —, Not applicable.

400 ms from the stimulus display until the outcome display); and (5) an event regressor at the timing of the outcome display. We excluded trials in which participants failed to press a button within the required duration (2000 ms). All regressors were convolved with a canonical hemodynamic response function and were entered into a general linear model. First, we created parameter estimates for the parametric modulation of delay length ("delay correlation," regressor 2). These parameter estimates were then used for the second-level group analyses (random-effects analysis) using a two-sample *t* test (Table 1). Next, we created parameter estimates for the parametric modulation of outcome magnitude ("magnitude sensitivity," regressor 3). These parameter estimates were then used for the second-level group analyses using a two-sample *t* test (Table 2).

ROIs and correction for multiple comparisons. Our hypothesis was that differential neural response in the striatum or insula between gains and losses to magnitude and/or delay may be related to the sign effect. We therefore report brain responses corrected for multiple comparisons in *a priori* ROIs in the striatum and insula. We referred to the Automatic Anatomical Labeling (AAL) Atlas (Tzourio-Mazoyer et al., 2002) to define each of the clusters of the striatum and insula regions anatomically. We used a combined image of the right putamen and left putamen from AAL as the region of the striatum (size = 16,584 mm³) and a combined image of the right insula and left insula from AAL as the region of the insula (size = 29,024 mm³). For inference of second-level group analyses (Tables 1 and 2), we applied a family-wise error correction of $p < 0.05$ by an anatomically defined small-volume correction. For visualization of the second-level group analyses (Fig. 3a,b), we used an uncorrected height threshold of $p < 0.001$ and no extent threshold.

To evaluate the brain responses associated with delay and magnitude information, we defined the functional ROIs as the peak voxel of clustered voxels where significant between-group differences were observed in terms of delay correlation and magnitude sensitivity in either the GAIN or LOSS condition (Tables 1 and 2). For the peak voxels in the striatum (x, y, z = 28, 2, −2) and insula (x, y, z = 36, −10, −10), we plotted the value of the parameter estimates for delay correlation and magnitude sensitivity averaged across participants for the GAIN and LOSS conditions (Fig. 3c–f). We used the MarsBaR toolbox (Brett et al., 2002) for these analyses. We used Pearson's correlation coefficient in the correlation analysis of the size of behavioral and neural sign effects (Fig. 3e,f).

Results

Behavioral results

We estimated the discount rate of each participant based on their individual choice patterns in the GAIN and LOSS conditions (Fig. 2a). The behavioral results showed that 23 participants exhibited the sign effect (*sign* group), as indicated by a smaller discount rate in the LOSS condition than in the GAIN condition. The other 9 participants did not exhibit the sign effect (*no-sign* group). Figure 2b shows the discount rate of the *sign* group and *no-sign* group in the GAIN and LOSS conditions. Discount rates for gains and losses were subjected to a 2 × 2 mixed ANOVA with condition (GAIN/LOSS) as the within-subject factor and group (*sign/no-sign*) as the between-subject factor. No significant effect of condition ($F_{(1,30)} = 0.037$, $p > 0.1$) or group ($F_{(1,30)} = 0.502$, $p > 0.1$) was found on discount rates, although we did find a significant interaction between condition and group ($F_{(1,30)} = 23.401$, $p < 0.05$). A multiple comparison analysis of discount rate revealed that the *sign* group exhibited a significantly smaller discount rate in the LOSS condition than the *no-sign* group ($p < 0.05$, Tukey–Kramer method). However, we observed no significant differences between the two groups in the GAIN condition ($p > 0.05$). Hence, our hypothetical focus on loss as a cause of the sign effect was confirmed by the behavioral patterns of the participants.

We did not find any demographic attributes that could account for the differences between the two groups. Although there was a higher male-to-female ratio in the *no-sign* group (77.8%) compared with the *sign* group (52.2%), this difference was not significant ($p > 0.1$, χ^2 distribution). Additionally, we did not observe any significant differences in age between the two groups ($p > 0.1$, two-sample *t* test).

Imaging results

To examine our hypothesis that the sign effect is caused by enhanced loss-related neural processing in magnitude and/or delay representation, we measured the brain activity of participants during decision making in the GAIN and LOSS conditions. We then assessed event-related BOLD signals at the time of decision-making (display stimulus). To evaluate the effect of delay information at the neural level, we assessed the correlation between event-related BOLD signals and delay length for the chosen options (termed "delay correlation"). To evaluate the effect of magnitude information at the neural level, we focused on the correlation between event-related BOLD signal and the binary variable (i.e., 10 or 40) of outcome magnitude for the chosen outcome magnitude (termed "magnitude sensitivity"); a small correlation would indicate weak magnitude sensitivity at the neural level.

We compared the delay correlation and magnitude sensitivity between the *sign* and *no-sign* groups in the LOSS condition (Tables 1 and 2). We observed a significant group difference in delay correlation in the striatum in the LOSS condition (Fig. 3a; two-sample *t* test with a corrected height threshold, $p < 0.05$). We also observed a significant group difference in magnitude sensitivity in the insula in the LOSS condition (Fig. 3b; two-sample *t* test with a corrected height threshold, $p < 0.05$). Both in the striatum and insula, a significant group difference was not observed in the GAIN condition, supporting our hypothesis.

Respective neural responses of the striatum and the insula to delay/magnitude information among the *sign* group and *no-sign* group were plotted (Fig. 3c,d). We found two major findings. The first was contrasting correlations in the LOSS condition between the *sign* group and *no-sign* group. In the striatum (Fig. 3c), delay

correlation was significantly positive for the *sign* group, but negative for the *no-sign* group ($p < 0.05$, one-sample t test). Similarly, in the insula (Fig. 3d), size sensitivity was significantly positive for the *sign* group, but negative and significant for participants in the *no-sign* group ($p < 0.05$, one-sample t test).

The second finding was an across-condition difference in correlations for the *sign* group. In the striatum (Fig. 3c), delay correlation was significantly positive and greater in the LOSS condition than in the GAIN condition ($p < 0.05$, Tukey–Kramer method). Similarly, in the insula (Fig. 3d), we found significantly positive and greater magnitude sensitivity in the LOSS condition than in the GAIN condition ($p < 0.05$, Tukey–Kramer method).

To complement the above analyses, we also checked whether there was any relationship established between magnitude and striatal activation or between delay and insular activation. We found no significant delay correlation in the insula or significant magnitude sensitivity in the striatum in all conditions and groups ($p > 0.05$, one-sample t test). Moreover, we found no significant across-condition differences in delay correlation in the insula or in magnitude sensitivity in the striatum in both groups ($p > 0.05$, Tukey–Kramer method). In summary, our results revealed magnitude-related responses in the insula and delay-related responses in the striatum.

The size of the behaviorally defined sign effect was significantly correlated with the size of the neural effects. The larger the degree of the sign effect (i.e., difference between discount rates for the GAIN and LOSS conditions), the larger the difference in striatal delay correlations ($r = 0.437$, $p < 0.05$; Fig. 3e) and insular magnitude sensitivity ($r = 0.673$, $p < 0.05$; Fig. 3f) for the GAIN and LOSS conditions.

Discussion

For the *sign* group relative to the *no-sign* group, we found enhanced neural responses in the striatum to a delay in loss outcomes, and enhanced neural responses in the insula to loss magnitude. In addition, participants showing greater difference in striatal and insular activities between the gain and loss outcomes exhibited a larger sign effect seen in the behavioral data.

Many investigations with human participants in the field of neuroscience have demonstrated that specific brain areas, such as the prefrontal cortex, insula, and striatum, are involved in temporal discounting of future appetitive outcomes, such as a drink and monetary rewards (McClure et al., 2004, 2007; Hariri et al., 2006; Kable and Glimcher, 2007; Tanaka et al., 2007; Wittmann et al., 2007; Ballard and Knutson, 2009; Gregorios-Pippas et al., 2009; Pine et al., 2009, 2010). In contrast, only a few studies have examined the neural differences between delay discounting of gains and losses (Bickel et al., 2009; Xu et al., 2009), and little is known about the neural substrates of the sign effect. A previous study (Bickel et al., 2009) indeed found no behavioral or neural differences in intertemporal choice between gains and losses. Such a result may have been obtained because the study considered differences in discounting for loss and gain without considering individual level variation. However, individual level differences have been reported in other studies (Kable and Glimcher, 2007; Gregorios-Pippas et al., 2009). Here, we addressed the important issue of individual variation in discounting, focusing on its potential relationship with the sign effect. Accordingly, we compared the brain activity of participants who exhibited the sign effect with that of participants who did not. Another study by Xu et al. (2009) found that the frequency of selecting smaller-sooner options was greater in a gain task than in a loss task. This choice pattern is also explained by the sign effect, in line with the

present findings. Brain activity was also measured during decision-making periods in their study, and greater activation was found when discounting losses than when discounting gains in some brain regions, such as the prefrontal area, insula, and striatum. Although our findings are consistent with those of Xu et al. (2009), they could not distinguish the effect of magnitude from that of delay for the sign effect. Bypassing the complicated distinction between decision making with positive and negative reinforcers (Seymour et al., 2007a) and focusing instead on delay-related or magnitude-related responses enabled the present study to better investigate the neural substrate underlying the sign effect.

We found greater insular responses to the magnitude of loss than gain among participants with the sign effect. This result could imply that the sign effect might be related to loss aversion, as proposed by economic theory (Loewenstein and Prelec, 1992), and that the insula forms the neural substrate of this effect. In the present study, participants with the sign effect exhibited a smaller discount rate for future losses than for future gains, as determined from the behavioral data. This observationally smaller discount rate for future loses could be the result of a stronger sensitivity for future losses in the insula, independent of the delay component. Consistent with this, participants who exhibited the sign effect had stronger sensitivity to losses in the insula. A recent study by Canessa et al. (2013), who showed that individual insular activity responding to monetary losses reflected individual degree of loss aversion in economic gambling tasks, supports our conclusion.

Our other main finding is the opposite delay correlation in the striatum seen across the groups for future losses. A possible interpretation of this result (Fig. 3c) is that the striatum represents the discount function. First, the striatal responses of the *no-sign* group showed a negative correlation between BOLD signals and delay. This finding is in line with previous findings of a negative correlation between striatal BOLD signals and length of the delay (Gregorios-Pippas et al., 2009) and a correlation between striatal BOLD signals and the hyperbolic discount functions (Pine et al., 2009; Pine et al., 2010). Second, in contrast to the *no-sign* group, we found a significant positive delay correlation in the striatum of the *sign* group, which suggests that these participants exhibiting the sign effect may exhibit the "dread effect." Neural correlates of dread in the insula have been observed in a previous study involving choice between delayed aversive outcomes (Berns et al., 2006), although not previously in the striatum. We observe a negative correlation between striatal activity and delay for losses, which may reflect a neural correlate for dread in the striatum. It has been previously suggested that the striatum represents subjective value after discounting has taken place (Kable and Glimcher, 2007). This would be consistent with the striatum representing an upward-sloping discount function for losses here. The interplay between striatal and insular activity in representing dread-discount functions suggests an interesting target for future studies.

Regarding the core neural mechanism of the sign effect, it is useful to consider studies investigating the neural basis of delay discounting. Previous human fMRI studies have found neuronal activity representing discounted value in the striatum (Hariri et al., 2006; Kable and Glimcher, 2007; Tanaka et al., 2007; Gregorios-Pippas et al., 2009; Pine et al., 2009, 2010). Recent neurophysiological studies have reported neuronal activity representing delay information in the dorsolateral prefrontal cortex (Kim et al., 2012) and lateral intraparietal cortex (Louie and Glimcher, 2010). Taking these previous findings together with

our current findings, we suggest one possible mechanism of the sign effect as follows: delay information is processed in cortical regions, such as the frontal and parietal areas before being transmitted to the striatum, which then represents the upward-sloping discount function of future losses. Finally, by combining with magnitude information from the insula that emphasizes losses, the striatum computes the discounted value needed for decision making. Future studies manipulating the activities of these cortical areas by transcranial magnetic stimulation could clarify the detailed biological basis of the sign effect.

Our results suggest that, at the neural level, both the effect of magnitude and the effect of delay are important sources of the sign effect. However, we had only limited (i.e., two) variations of magnitudes to identify the effect of magnitude. In addition, we could not remove the limitation inherent to intertemporal choice tasks whereby delay and magnitude are correlated, namely, that an option with a smaller magnitude had a shorter delay, even though we randomly defined the length of delay from the uniform distribution. Importantly, however, the group difference in delay correlation was observed in the striatum (Fig. 3c) and not in the insula, and the group difference in magnitude sensitivity was observed in the insula (Fig. 3d) and not in the striatum. These results indicate that we have identified the effect of delay and magnitude on the sign effect. Future studies should develop novel paradigms that allow for a completely uncorrelated assessment of delay-related and magnitude-related neural systems by manipulating delay and magnitude independently as, for example, seen in a study using rats (Roesch et al., 2006). The other potential drawback of the study is that we set a time restriction in the intertemporal choice task. Thus, participants tried to maximize their total amount of gains as well as minimize their total amount of losses within a particular time. However, in a standard intertemporal choice task, the duration of a single trial is fixed regardless of the participants' choice, which allows them not to optimize the number of trials dependent on their choices. Our experimental task was able to capture not only the neural mechanism of discounting for each trial but also the neural mechanism of optimizing in the decision-making problem over a longer time frame (Sonuga-Barke, 2002, 2003; Plichta et al., 2009). Future studies should more clearly identify the role played by the striatum and insula for time discounting in a trial and optimize the whole session.

This study showed that the majority of participants exhibited gain–loss asymmetry at the behavioral level, which is actually an "anomaly" in standard economics. In the first and second recruitment stages of this study, 3 of the 32 participants (9.4%) and 9 of the 60 participants (15%) were identified as *no-sign* participants. This minor proportion of *no-sign* individuals is similar to that reported by a survey in Japan in which 11.5% of 2289 participants did not exhibit the sign effect (Ikeda et al., 2010). Thus, it is reasonable to expect that ~10% of individuals in a sample of healthy adults will not consistently exhibit the sign effect. This consistency in behavioral results suggests that the intuitive and relatively short-term intertemporal choice task (on the order of seconds) used in the present study was able to capture the sign effect similarly to the previous study that used intertemporal choice tasks in a questionnaire form with longer delay ranges (on the order of days). Future studies should analyze the relationships between the degree of the sign effect, striatal and insular activities, and individual biological attributes (e.g., ethnicity, sex, age, obesity, and gene polymorphisms) as well as social attributes (e.g., culture, income, job, social status, and civil status) to further enhance understanding of the prevalence of the sign effect.

References

Ainslie G (1975) Specious reward: a behavioral theory of impulsiveness and impulse control. Psychol Bull 82:463–496. CrossRef Medline

Ballard K, Knutson B (2009) Dissociable neural representations of future reward magnitude and delay during temporal discounting. Neuroimage 45:143–150. CrossRef Medline

Berns GS, Chappelow J, Cekic M, Zink CF, Pagnoni G, Martin-Skurski ME (2006) Neurobiological substrates of dread. Science 312:754–758. CrossRef Medline

Bickel WK, Pitcock JA, Yi R, Angtuaco EJ (2009) Congruence of BOLD response across intertemporal choice conditions: fictive and real money gains and losses. J Neurosci 29:8839–8846. CrossRef Medline

Brett M, Anton JL, Valabregue R, Poline JB (2002) Region of interest analysis using an SPM toolbox. Neuroimage 16:497.

Canessa N, Crespi C, Motterlini M, Baud-Bovy G, Chierchia G, Pantaleo G, Tettamanti M, Cappa SF (2013) The functional and structural neural basis of individual differences in loss aversion. J Neurosci 33:14307–14317. CrossRef Medline

Frederick S, Loewenstein G, O'Donoghue T (2002) Time discounting and time preference: a critical review. J Econ Lit 40:351–401. CrossRef

Gregorios-Pippas L, Tobler PN, Schultz W (2009) Short-term temporal discounting of reward value in human ventral striatum. J Neurophysiol 101:1507–1523. CrossRef Medline

Hariri AR, Brown SM, Williamson DE, Flory JD, de Wit H, Manuck SB (2006) Preference for immediate over delayed rewards is associated with magnitude of ventral striatal activity. J Neurosci 26:13213–13217. CrossRef Medline

Ho MY, Mobini S, Chiang TJ, Bradshaw CM, Szabadi E (1999) Theory and method in the quantitative analysis of "impulsive choice" behaviour: implications for psychopharmacology. Psychopharmacology 146:362–372. CrossRef Medline

Ikeda S, Kang MI, Ohtake F (2010) Hyperbolic discounting, the sign effect, and the body mass index. J Health Econ 29:268–284. CrossRef Medline

Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. Nat Neurosci 10:1625–1633. CrossRef Medline

Kahneman D, Tversky A (1984) Choices, values, and frames. Am Psychol 39:341–350. CrossRef Medline

Kim S, Cai X, Hwang J, Lee D (2012) Prefrontal and striatal activity related to values of objects and locations. Front Neurosci 6:108. CrossRef Medline

Loewenstein G (1987) Anticipation and the valuation of delayed consumption. Econ J 97:666–684. CrossRef

Loewenstein GF (1988) Frames of mind in intertemporal choice. Manage Sci 34:200–214. CrossRef

Loewenstein G, Prelec D (1991) Negative time preference. Am Econ Rev 81:347–352.

Loewenstein G, Prelec D (1992) Anomalies in intertemporal choice: evidence and an interpretation. Q J Econ 107:573–597. CrossRef

Loewenstein GF, Prelec D (1993) Preferences for sequences of outcomes. Psychol Rev 100:91–108. CrossRef

Louie K, Glimcher PW (2010) Separating value from choice: delay discounting activity in the lateral intraparietal area. J Neurosci 30:5498–5507. CrossRef Medline

McClure SM, Laibson DI, Loewenstein G, Cohen JD (2004) Separate neural systems value immediate and delayed monetary rewards. Science 306:503–507. CrossRef Medline

McClure SM, Ericson KM, Laibson DI, Loewenstein G, Cohen JD (2007) Time discounting for primary rewards. J Neurosci 27:5796–5804. CrossRef Medline

Odum AL, Madden GJ, Bickel WK (2002) Discounting of delayed health gains and losses by current, never- and ex-smokers of cigarettes. Nicotine Tobacco Res 4:295–303. CrossRef

Palminteri S, Justo D, Jauffret C, Pavlicek B, Dauta A, Delmaire C, Czernecki V, Karachi C, Capelle L, Durr A, Pessiglione M (2012) Critical roles for anterior insula and dorsal striatum in punishment-based avoidance learning. Neuron 76:998–1009. CrossRef Medline

Paulus MP, Rogalsky C, Simmons A, Feinstein JS, Stein MB (2003) Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. Neuroimage 19:1439–1448. CrossRef Medline

Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006)

Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442:1042–1045. CrossRef Medline

Pine A, Seymour B, Roiser JP, Bossaerts P, Friston KJ, Curran HV, Dolan RJ (2009) Encoding of marginal utility across time in the human brain. J Neurosci 29:9575–9581. CrossRef Medline

Pine A, Shiner T, Seymour B, Dolan RJ (2010) Dopamine, time, and impulsivity in humans. J Neurosci 30:8888–8896. CrossRef Medline

Plichta MM, Vasic N, Wolf RC, Lesch KP, Brummer D, Jacob C, Fallgatter AJ, Grön G (2009) Neural hyporesponsiveness and hyperresponsiveness during immediate and delayed reward processing in adult attention-deficit/hyperactivity disorder. Biol Psychiatry 65:7–14. CrossRef Medline

Roesch MR, Taylor AR, Schoenbaum G (2006) Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. Neuron 51:509–520. CrossRef Medline

Schweighofer N, Shishida K, Han CE, Okamoto Y, Tanaka SC, Yamawaki S, Doya K (2006) Humans can adopt optimal discounting strategy under real-time constraints. PLoS Comput Biol 2:e152. CrossRef Medline

Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS (2004) Temporal difference models describe higher-order learning in humans. Nature 429:664–667. CrossRef Medline

Seymour B, Singer T, Dolan R (2007a) The neurobiology of punishment. Nat Rev Neurosci 8:300–311. CrossRef Medline

Seymour B, Daw N, Dayan P, Singer T, Dolan R (2007b) Differential encoding of losses and gains in the human striatum. J Neurosci 27:4826–4831. CrossRef Medline

Sonuga-Barke EJ (2002) Psychological heterogeneity in AD/HD: a dual pathway model of behaviour and cognition. Behav Brain Res 130:29–36. CrossRef Medline

Sonuga-Barke EJ (2003) The dual pathway model of AD/HD: an elaboration of neuro-developmental characteristics. Neurosci Biobehav Rev 27:593–604. CrossRef Medline

Story GW, Vlaev I, Seymour B, Winston JS, Darzi A, Dolan RJ (2013) Dread and the disvalue of future pain. PLoS Comput Biol 9:e1003335. CrossRef Medline

Tanaka SC, Schweighofer N, Asahi S, Shishida K, Okamoto Y, Yamawaki S, Doya K (2007) Serotonin differentially regulates short- and long-term prediction of rewards in the ventral and dorsal striatum. PLoS One 2:e1333. CrossRef Medline

Thaler R (1981) Some empirical evidence on dynamic inconsistency. Econ Lett 8:201–207. CrossRef

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15:273–289. CrossRef Medline

Wittmann M, Leland DS, Paulus MP (2007) Time and decision making: differential contribution of the posterior insular cortex and the striatum during a delay discounting task. Exp Brain Res 179:643–653. CrossRef Medline

Xu L, Liang ZY, Wang K, Li S, Jiang T (2009) Neural mechanism of intertemporal choice: from discounting future gains to future losses. Brain Res 1261:65–74. CrossRef Medline