

## On Weight and Waiting: Delay Discounting in Anorexia Nervosa Pre- and Post-Treatment

### *Supplemental Information*

#### Supplemental Methods

##### Delay Discounting Task Design

Participants made choices between amounts of money available at various delays: smaller-sooner (SS) and larger-later (LL). SS options were available “today” (NOW) or “in 2 weeks” (NOT-NOW). The delay to the LL was either 2 or 4 weeks after the SS option. The relative difference in dollar amounts between SS and LL (i.e.,  $(LL-SS)/SS$ ) was either 1, 3, 5, 10, 15, 20, 25, 35, or 50%. SS amounts ranged from \$15 to \$85 dollars.

This factorial design results in 36 trials—2 (NOW or NOT-NOW) by 2 (2 or 4 week delay) by 9 (relative percentage difference). Two sets of trials were used during the fMRI scan. These sets were duplicated, with one duplicate presenting the LL option as the default (ACCELERATE) and the other presenting the SS as the default (DELAY), for a possible total of 144 trials. Participants outside the scanner used only one duplicate set, for 72 trials. The order and frame (ACCELERATE or DELAY) was counterbalanced within and between participants. For half of the participants, ACCELERATE trials were paired with amazon gift cards and DELAY trials with cash, also counterbalanced between participants.

In the scanner, participants had 10 seconds to indicate their choice, and received feedback for 2 seconds, indicating that their choice was recorded. For the feedback, the triangle below the chosen option turned green while the triangle below the alternative option disappeared. Feedback was followed by a variable intertrial interval ranging from 7 to 8 seconds. The task was presented in 4 runs of 4 minutes each, allowing participants to rest between runs. Runs were presented in one of two counterbalanced orders, either (A = Acceleration; D = Delay) A-D-D-A or D-A-A-D. Outside of the scanner there was no time limit.

##### Fitting Discount Models

We estimated the probability of making a larger-later choice given the choice subjective values on a given trial.

$$p = \frac{1}{1 + e^{(-\beta(SV_{LL} - SV_{SS}))}}$$

Where  $p$  is probability,  $\beta$  is the slope parameter, and  $SV$  is the reward (SS or LL) subjective value.  $SV$  was determined using three different discounting models (1-3):

$$SV = \frac{A}{1 + kT} \quad \text{hyperbolic}$$

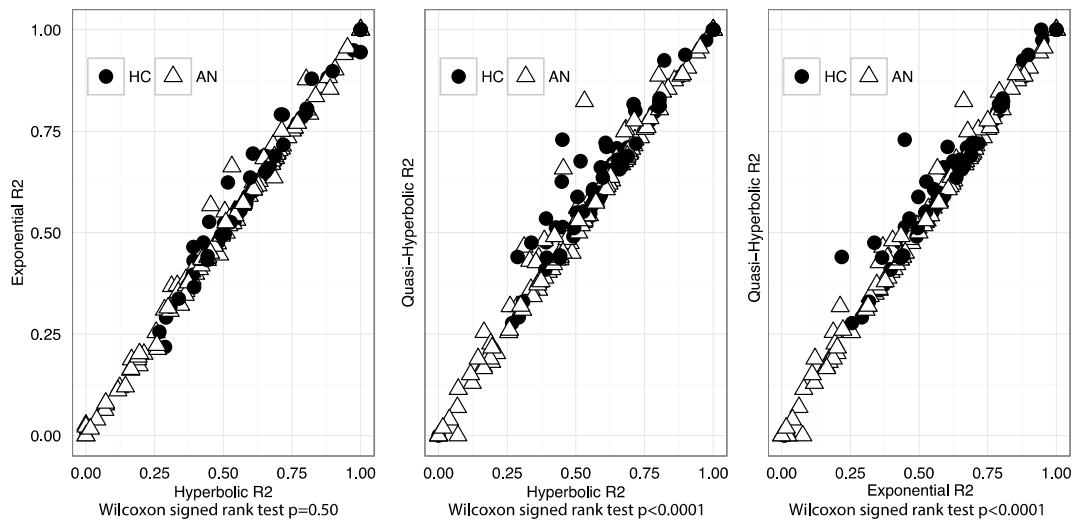
$$SV = Ae^{-kT} \quad \text{exponential}$$

$$SV = \begin{cases} A & \text{if } T = 0 \\ Abe^{-dT} & \text{if } T > 0 \end{cases} \quad \text{quasihyperbolic}$$

Where  $A$  is the offered amount,  $k$  is a 1-parameter discount rate,  $b$  and  $d$  are the 2-parameter discount rates, and  $T$  is the time to the reward in years.

These parameters were estimated to minimize the negative log-likelihood of individual choice probability using MATLAB's `fmincon` minimizing function (4).

A pseudo- $R^2$  was generated by comparing the fit against the fit of all random choices  $p = 0.5$  (5, Figure S1). Though the quasi-hyperbolic model had a significantly better fit than the other two models using the signed Wilcoxon rank sum test, there was no difference between models using an unsigned test (hyperbolic  $p = 0.25$ , exponential  $p = 0.33$ ). As such, the hyperbolic model was chosen to describe the results as a single discount rate is simpler to interpret, and is the most commonly described discount model in the literature. The exclusion of participants' data for inadequately performing the task is standard (6; 7), though it is interesting that all but one excluded behavioral session are from individuals with AN. The value was set at 0.15 to exclude as few imaging participants as possible, though behavioral and imaging results were unchanged using a slightly stricter cutoff 0.2. Other exclusion criteria are summarized further below.



**Figure S1.** Pseudo- $R^2$  Comparisons of Each Discounting Model. All participants who completed the task are included. A pseudo- $R^2$  below 0.15 for the hyperbolic model was selected as being insufficiently different from random to use in the analysis.

### Discount Rate Analysis

As mentioned in the main text, the natural log-transformed values  $\log(k)$  were computed for group comparisons. Lower values of  $\log(k)$  indicate a greater tendency to select the larger, delayed reward.

Log-transformed discount rates were compared using the lme4, afex, and lsmeans packages in the R-statistics language (8-11). *P*-values were determined using conditional *F*-tests with Kenward-Roger correction of degrees-of-freedom, as implemented in the Anova function (with Type III *F*-tests) from the package car (12; this function calls the KRmodcomp function from package pbkrtest; 13)

discount.model.1<-mixed(logK~Dx\*Session + (1 + Session | Subject))

**Table S1.** Discount Model 1 Parameters

Effect	Estimate	<i>F</i> -Stat	ndf	ddf	<i>p</i> -value
Intercept	0.68	482.23	1	95.85	<0.0001
Dx	0.06	.45	1	95.85	0.50
Session	-0.14	12.01	1	71.43	0.0009
Dx:Session	0.17	19.45	1	71.43	<0.0001

A similar pattern of behavior was seen just within the fMRI subset or behavioral subsets (Table S3). The Diagnosis by Session interaction effect remained significant when age and IQ were included in the model, scaled such they had a mean of zero and a standard deviation of one.

discount.model.2<-mixed(log(K) ~ Dx\*Session\*zAge\*zIQ + (1+Session| Subject))

**Table S2.** Discount Model 2 Parameters

Effect	Estimate	<i>F</i> -Stat	ndf	ddf	<i>p</i> -value
Intercept	0.63	15.52	1	84.99	<0.0001
Dx	0.14	.72	1	84.99	0.40
Session	-0.21	14.52	1	61.46	0.0003
zAge	0.31	2.70	1	88.38	0.10
zIQ	-0.38	4.89	1	83.66	0.030
Dx:Session	0.28	25.76	1	61.46	<0.0001
Dx:zAge	0.14	0.53	1	88.38	0.47
Session:zAge	-0.12	3.16	1	62.28	0.080
Dx:zIQ	-0.14	0.67	1	83.66	0.42
Session:zIQ	0.09	1.89	1	60.38	0.17
zAge:zIQ	-0.00	0.00	1	83.00	0.98
Dx:Session:zAge	0.18	6.45	1	62.28	0.014
Dx:Session:zIQ	-0.02	0.17	1	60.38	0.68
Dx:zAge:zIQ	-0.01	0	1	83.00	0.97
Session:zAge:zIQ	0.11	4.27	1	59.84	0.043
Dx:Session:zAge:zIQ	-0.05	0.90	1	59.84	0.35

**Table S3.** Summary of Delay Discounting Task Behavior and Statistical Analysis

Log(k) <sup>a</sup>	All Participants				Scanned Subset					
	Session 1		Session 2		Session 1		Session 2			
Diagnosis	<i>n</i>	Mean±SD	<i>n</i>	Mean±SD	<i>n</i>	Mean±SD	<i>n</i>	Mean±SD		
HC	39	0.768±1.497	31	0.589±1.524	22	0.602±1.453	17	0.463±1.468		
AN	54	0.052±1.523	43	0.890±1.500	24	0.239±1.543	19	0.969±1.131		
Scanned Subset Significance Testing <sup>b</sup>					Statistic	p-value				
Diagnosis by Session Interaction					$F_{1,34.2} = 7.39$	$p = 0.010$				
Diagnosis at Session 1					$t_{44} = 0.82$	$p = 0.42$				
Diagnosis at Session 2					$t_{34} = -1.17$	$p = 0.25$				
Session:AN paired					$t_{16} = -2.58$	$p = 0.020$				
Session:HC paired					$t_{16} = 0.6$	$p = 0.56$				
Behavioral Subset Significance Testing <sup>b</sup>					Statistic	p-value				
Diagnosis by Session Interaction					$F_{1,34.0} = 11.10$	$p = 0.0021$				
Diagnosis at Session 1					$t_{45} = 2.32$	$p = 0.025$				
Diagnosis at Session 2					$t_{36} = -0.15$	$p = 0.88$				
Session:AN paired					$t_{19} = -3.83$	$p = 0.0011$				
Session:HC paired					$t_{12} = 1.06$	$p = 0.31$				
Log(k) correlations	Session 1			Session 2			Change in Log(k)			
	t-test	<i>r</i>	<i>p</i>	t-test	<i>r</i>	<i>p</i>	t-test	<i>r</i>	<i>p</i>	
Age	HC	2.19	0.34	0.035	1.57	0.28	0.13	0.33	0.06	0.75
	AN	-0.38	-0.05	0.71	1.53	0.23	0.13	2.79	0.42	0.0084
IQ	HC	-2.33	-0.37	0.026	-2.59	-0.45	0.016	-1.82	-0.33	0.08
	AN	-0.31	-0.04	0.75	-0.26	-0.04	0.80	0.28	0.05	0.78
BMI	HC	-0.38	-0.06	0.71	0.13	0.02	0.90	-1.73	-0.31	0.094
	AN	0.27	0.04	0.79	0.55	0.09	0.58	0.88	0.15	0.38
Duration of Illness	AN	-0.95	-0.13	0.35	0.55	0.09	0.59	1.75	0.28	0.09
EDE Score	AN	0.72	0.10	0.48	0.56	0.09	0.58	0.24	0.04	0.81
Time to Discharge	AN				-0.14	-0.02	0.89	0.38	0.06	0.71
Household Income	Level	<i>n</i>	(%)	<i>n</i>	(%)	$W^c$		<i>p</i>		
< \$10,000	1	5	13.2	9	16.7	1222		0.12		
\$10,000-19,999	2	4	10.5	7	13					
\$20,000-34,999	3	7	18.4	13	24.1					
\$35,000-49,999	4	4	10.5	6	11.1					
\$50,000-99,999	5	4	10.5	12	22.2					
\$100,000-199,999	6	10	26.3	4	7.4					
> \$200,000	7	4	10.5	3	5.6					
Employment	Level	<i>n</i>	(%)	<i>n</i>	(%)	$W^c$		<i>p</i>		
	None	14	35.9	36	61	1354		0.10		
	Part-time	17	43.6	10	16.9					
	Full-Time	8	20.5	13	22					

<sup>a</sup> Log-transformed values of the discount rate (in years), *k*.

<sup>b</sup> *p*-values shown are uncorrected.

<sup>c</sup> Mann-Whitney-Wilcoxon test.

AN, anorexia nervosa patient; BMI, body mass index; EDE, eating disorder examination; HC, healthy controls.

To test whether inpatient participants changed their behavior because they felt they were leaving the institute, we did a number of post-hoc analyses. First, we entered payment-type (Amazon or Cash) into the linear mixed-effects model, and this failed to reveal a change in effects.

discount.model.3<-mixed(log(K) ~ Dx\*Session\*PaymentType + (1+Session\*PaymentType | Subject)

**Table S4.** Discount Model 3 Parameters

<b>Effect</b>	<b>Estimate</b>	<b>F-Stat</b>	<b>ndf</b>	<b>ddf</b>	<b>p-value</b>
Intercept	0.65	19.06	1	94.94	<0.0001
Dx	0.11	0.60	1	94.94	0.44
Session	-0.19	12.56	1	70.97	0.0007
PaymentType	0.04	1.31	1	90.49	0.26
Dx:Session	0.26	22.56	1	70.97	<0.0001
Dx:PaymentType	0.00	0.00	1	90.49	0.98
Session:PaymentType	-0.00	0.00	1	81.11	0.97
Dx:Session:PaymentType	-0.04	3.77	1	81.11	0.056

Secondly, we examined correlations between discount rates, and change in discount rates, with patient time to discharge (Table S3). Discharge dates were not fully determined by reaching the required weight, but this may have indicated to some individuals that they may be leaving soon. No significant correlation was found, suggesting that the change in discount rate seen in the AN group was not due to their proximity to discharge.

### **Correlation Analysis**

In addition to the time to discharge analysis mentioned above, we also examined correlations of discount rate with clinical characteristics. Among the healthy controls, there were significant associations between discount rate and age, such that older participants were more impatient, as well as associations between discount rate and IQ, with higher IQ being more patient. These associations were not present among individuals with AN (Table S3). As discussed above, including age and IQ in the discount rate analysis did not alter the results. We tested the association of discount rate with body mass index, duration of illness, and eating disorder examination scores—measures of illness severity (Table S3)—and found no significant correlations.

### **Choice Analysis**

Another approach to analyzing the delay discounting behavior is to use a generalized linear mixed-effects model for choice behavior from each trial. The model was similar to the one for discounting behavior, but additional predictors were added that model Immediacy (NowNotnow), smaller-sooner amount (SS\_Amount), time difference between options (TimeDiff), and relative difference between reward

(RelDiff100) for the given trials. The continuous predictors were scaled such that they had a mean of zero and a standard deviation of 1. We followed the advice of Barr *et al.* (14), and used a maximal random-effects structure: the repeated-measures nature of the data was accordingly modeled by including a per-participant random adjustment to the fixed intercept ("random intercept"), as well as per-participant random adjustments to the Session, NowNotnow, SS\_amount, TimeDiff, RelDiff100, and four interaction (Session:NowNotnow, Session:SS\_amount, Session:TimeDiff, and Session:RelDiff100) slopes ("random slopes"); in addition, we included all possible random correlation terms among the random effects. *P*-values were determined using the Likelihood Ratio Tests as implemented in the mixed function of the afex package.

```
choice.model.1<-mixed(Choice ~ Dx*Session* (NowNotnow+zSS_Amount+zRelDiff100+zTimeDiff) +
(1+ Session*(NowNotnow+zSS_Amount+zRelDiff+zTimeDiff)|Subject),method="LRT")
```

**Table S5.** Choice Model 1 Parameters

Effect	Estimate	df.large	df.small	Chi-sq	<i>p</i> -value
Intercept	1.17	75	74	8.17	0.0043
Dx	-0.25	75	74	0.32	0.57
Session	-0.72	75	74	14.07	0.0002
NowNotnow	0.05	75	74	0.52	0.47
zSS_Amount	1.19	75	74	119.38	<0.0001
zRelDiff100	3.58	75	74	111.88	<0.0001
zTimeDiff	-0.36	75	74	41.31	<0.0001
Dx:Session	0.83	75	74	20.34	<0.0001
Dx:zNowNotnow	-0.09	75	74	2.12	0.15
Dx:zSS_Amount	0.14	75	74	3.67	0.055
Dx:zRelDiff100	0.23	75	74	0.93	0.34
Dx1:zTimeDiff	-0.06	75	74	1.45	0.23
Session:NowNotnow	0.05	75	74	1.44	0.23
Session:zSS_Amount	-0.01	75	74	0.06	0.81
Session:zRelDiff100	-0.44	75	74	9.66	0.0019
Session:zTimeDiff	0.02	75	74	0.42	0.52
Dx:Session:NowNotnow	0.09	75	74	3.52	0.061
Dx:Session:zSS_Amount	0.04	75	74	0.77	0.38
Dx:Session:zRelDiff100	0.33	75	74	6.68	0.0098
Dx:Session:zTimeDiff	0.05	75	74	1.78	0.18

There are no group differences in immediacy ( $p = 0.15$ ), time difference ( $p = 0.23$ ), or relative difference ( $p = 0.34$ ), but a marginal effect of SS amount ( $p = 0.055$ ). This suggests that there are no systematic session-independent differences between AN and HC in how variation in the timing and in the amounts of the rewards affects their intertemporal choices, or to put it differently, there is no evidence that AN might neglect the time or amount information. The marginally significant effect of SS amount suggests that AN tend to show an attenuated magnitude effect (which is the effect that, everything else being

equal, larger amounts of money lead to increased patience). This latter result might suggest that AN tend to be less sensitive to increasing outcome magnitudes; but given that this effect is only marginally significant, we are hesitant to make strong conclusions. However, there is a significant Diagnosis-by-Session-by-Relative Difference interaction ( $p = 0.0098$ ). This term similarly suggests that that the change in preference for LL that occurs in the AN group across sessions might depend on being more responsive to changes in relative difference once weight restored. This would suggest that the AN group are somewhat less aware or responsive to the relative differences in outcome magnitudes when underweight.

In post-hoc testing using lsmeans, uncorrected, there were significant differences between groups, the same as seen with the discount rate analysis.

**Table S6.** Comparison of Linear Estimate of the Mean

<b>Contrast</b>	<b>Odds-ratio</b>	<b>SE</b>	<b>z-ratio</b>	<b>p-value</b>
HC, Session 1 – AN, Session 1	0.12	0.10	-2.38	0.017
HC, Session 2 – HC, Session 2	1.25	0.64	0.44	0.66
AN, Session 2 – AN, Session 1	0.05	0.02	-6.74	<0.0001
HC, Session 2 – AN ,Session 2	3.19	2.74	1.35	0.18

Though these two approaches are not independent from one another, it is encouraging that both show the same effects. One possible advantage of using the Choice analysis described here is that it could provide insight into whether there are differential effects of smaller-sooner amount, time difference, or relative difference. Though there was no three-way interaction with time difference between choices, future studies that include a broader range of delay intervals may better detect this effect.

### **Response Time Analysis**

Response times may give some insight into the type of strategy being used by individuals. We used a linear mixed-effects model from trial data using Diagnosis, Session, Choice, their interactions, and including a full random-effects structure. This revealed an overall quickening across sessions that did not differ by diagnosis or by choice made. There was additionally a Diagnosis by Session by Choice interaction, which was further examined in the main text. This revealed that the AN group went from being slower on SS relative to LL trials at Session 1 when underweight, but faster at Session 2 once weight restored, whereas the HC group showed no choice specific change. This analysis was then repeated including the absolute difference in subjective value between the two options presented (LL-SS). Absolute value was used as it is the magnitude of the difference, and not the direction, that captures the difficulty of a given trial. With this term included, the three-way interaction is marginal ( $p = 0.063$ ), but follows the same pattern, suggesting that this change is due to more than just the change in difficulty across trials.

response.model.1<-mixed(RT ~ Dx\*Session\*Choice+ (1+Session\* Choice | Subject),method="LRT")

**Table S7.** Response Time Model 1 Parameters

Effect	Estimate	df.large	df.small	Chi-square	p-value
Intercept	3055.3	19	18	221.28	0
Dx	-63.2	19	18	0.37	0.54
Session	177.3	19	18	12.67	0.0004
Choice	-8.1	19	18	0.02	0.88
Dx:Session	-19.9	19	18	0.18	0.67
Dx:Choice	-41.4	19	18	0.66	0.42
Session:Choice	2.9	19	18	0.02	0.90
Dx:Session:Choice	-59.8	19	18	5.86	0.016

response.model.2<-mixed(RT ~ Dx\*Session\*Choice+zAbsPresDiff + (1+Session\*Choice|Subject), method="LRT")

**Table S8.** Response Time Model 2 Parameters

Effect	Estimate	df.large	df.small	Chi-square	p-value
Intercept	2988.5	25	24	221.64	<0.0001
Dx	-83.5	25	24	0.72	0.40
Session	158.0	25	24	10.3	0.0013
Choice	-73.0	25	24	2.72	0.099
zAbsPresDiff	-233.0	25	24	81.64	<0.0001
Dx:Session	-7.8	25	24	0.03	0.86
Dx:Choice	-35.3	25	24	0.65	0.42
Session:Choice	-11.6	25	24	0.2	0.65
Dx:Session:Choice	-44.8	25	24	3.46	0.063

### Exclusion Criteria and Missing Data

Ninety-eight participants had at least partial behavioral data (59 AN). At Session 1, 3 AN had uninterpretable data. At Session 2, 2 AN and 1 HC had uninterpretable data, and 1 AN had data lost. Two AN were tested only at Session 2; 13 AN and 7 HC were not available for Session 2 testing. This left 93 participants from Session 1 (54 AN), 74 from Session 2 (43 AN), and 67 from both (37 AN). Forty-eight of the 98 participants performed the task during the fMRI scan (26 AN). Forty-six had behavioral data at Session 1 (24 AN), and 36 at Session 2 (19 AN), with 34 participants having behavioral data at both Sessions (17 AN). Mixed-effects modeling is robust to some amount of missing data, and the degrees of freedom are adjusted according to the specific effect being tested.

Imaging data were excluded for movement (greater than 2 mm movement in more than 10% of trials) and fewer than 12 smaller-sooner or 12 larger-later choices, but behavioral data was still used. Two AN from Session 1, 1 AN from Session 2, and 1 HC from both Sessions were excluded due to movement. Two AN (from Session 1) and 1 HC (from Session 2) had too few SS trials to use their SS regressors. One AN and 1 HC had too few LL trials to use their regressors from both Sessions. From these 47 participants, 78 scans could be used in the analysis. Participants that had both SS and LL estimates:

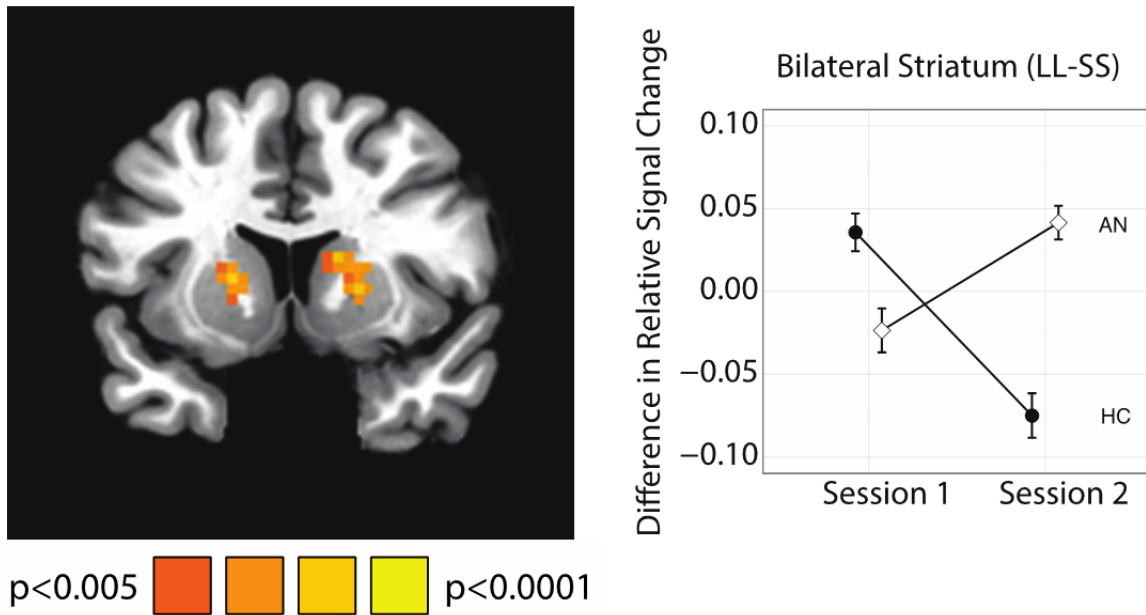


Session 1 (20 AN, 20 HC), Session 2 (17 AN, 14 HC), both Sessions (14 AN, 14 HC). While the final sample size was reduced, the sample sizes remained large for a study of this kind.

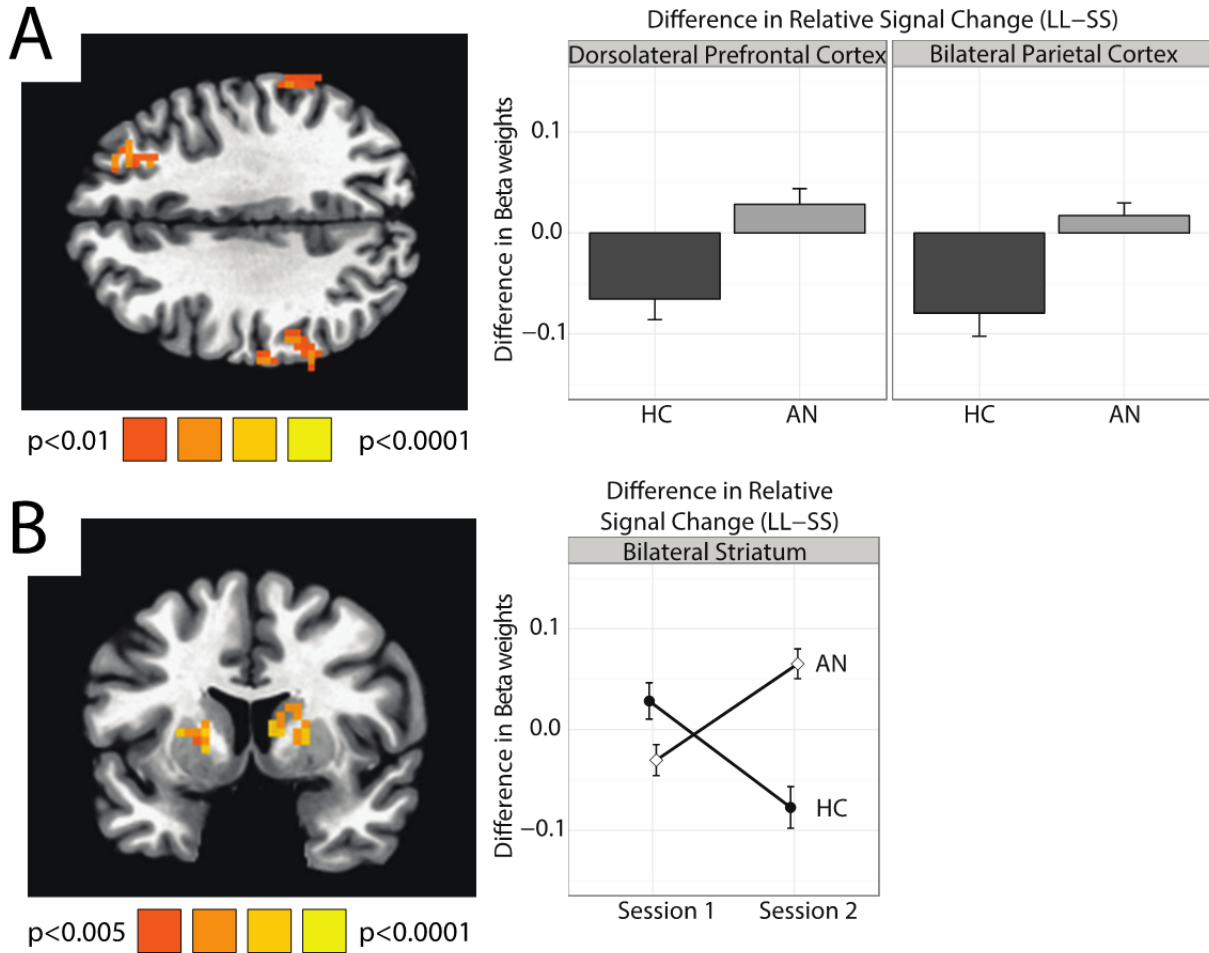
### **Difference in Larger-Later and Smaller-Sooner Subjective Value Regressor for Single-Subject General Linear Model Analysis**

In an attempt to account for the change in subjective value, we added an additional regressor to the one used in the main text, giving us 23 regressors: baseline, trend, and quadratic signal regressors to capture shifts in signal change for each of the 4 runs (12), motion parameters (6), and 5 trial-specific regressors. These included a trial regressor (1 for each trial), a Choice regressor (-1 for smaller-sooner (SS) choices and +1 for larger-later (LL) choices), an Immediacy regressor (-1 for NOW trials and +1 for NOT-NOW trials), an interaction (Choice x NOW/NOT-NOW) regressor, and an amplitude-modulated regressor for the difference in subjective value between LL and SS options of a given trial according to the subject-specific, behaviorally determined discount rate (z-normalized). We did this analysis two ways, once with the absolute value of the difference in subjective value (capturing choice difficulty), and once with true difference (capturing choice specific differences in difficulty). The group analysis for both of these first level regressions were identical to the main text, and the results are presented below. It is important to note that the main text analysis likely already accounted for some of these differences in group difficulty, due to the inclusion of the duration modulated regressor of response time. Response times were negatively correlated with trial absolute difference in subjective value ( $r = -0.231$ ,  $p < 0.0001$ ). When using directional difference in subjective value, the response times correlated negatively for SS trials ( $r = -0.226$ ,  $p < 0.0001$ ), and positively for LL trials ( $r = 0.299$ ,  $p < 0.0001$ ).

When comparing the group analyses using either of these subjective value regressors (absolute or true difference) to the main text analysis with no subjective value regressor, only the pattern in the bilateral striatum remained significant in the diagnosis-by-session-by-choice interaction term (Figure S2, Table S10). This suggests that the anterior cingulate, dorsolateral prefrontal, and parietal cortex activity differences seen in the main text are due to the differences in subjective value that the groups experienced and each time point, and perhaps therefore differences in difficulty across sessions. An additional significant interaction term, diagnosis-by-choice, was revealed in the dorsolateral prefrontal cortex and bilateral parietal cortex in the group analysis using the true, but not the absolute, subjective value regressor (Figure S3, Table S10). The group differences here suggest that this frontoparietal system is recruited more for SS choices compared to LL choices in the healthy controls, whereas the AN group showed no difference in activity between choices. These results point to the potential role of the frontoparietal network in AN disease process, where there is no differential activity seen between SS and LL choices as is observed in the HC group.



**Figure S2.** Group Analysis when the Absolute Difference in Subjective Value was Included in the Single Subject General Linear Model. The diagnosis-by-session-by-choice interaction seen in the bilateral striatum reported in the main text is also present when including the regressor of absolute difference in subjective value. The dorsolateral prefrontal cortex, parietal cortex, and anterior cingulate no longer pass threshold correction.



**Figure S3.** Group Analysis when the True Direction Difference in Subjective Value was Included in the Single Subject General Linear Model. **(A)** Dorsolateral prefrontal cortex and bilateral parietal cortex show greater activity for SS than for LL choices in healthy controls, whereas no difference is seen in the AN group. **(B)** The diagnosis-by-session-by-choice interaction seen in the main text is the same when including the regressor of true difference in subjective value.

**Table S9.** Activation Maps of Individuals with Anorexia Nervosa (AN) and Healthy Controls (HC) During the Delay Discount Task

Region	x	y	z	voxels	volume	F-stat
<b>Interaction Effect of Choice (SS vs LL), Diagnosis (AN vs HC), and Session (1 vs 2)</b>						
Cluster FWE-corrected $p < 0.01$ , individual voxel threshold 0.01, size $\geq 41$ voxels						
Left Putamen	+25.5	+19.5	+8.5	89	2.4 cm <sup>3</sup>	20.3
Left Striatum	+7.5	-1.5	+11.5	75	2.0 cm <sup>3</sup>	20.6
Right Striatum	-16.5	-7.5	+11.5	69	1.9 cm <sup>3</sup>	21.7
Dorsal Anterior Cingulate	+7.5	-19.5	+29.5	66	1.8 cm <sup>3</sup>	13.2
Right Parietal Cortex	-55.5	+40.5	+35.5	60	1.6 cm <sup>3</sup>	19.6
Right Dorsolateral Prefrontal Cortex	-31.5	+25.5	+32.5	44	1.2 cm <sup>3</sup>	14.2
<b>Absolute Difference in Subjective Value</b>						
<b>Interaction Effect of Choice (SS vs LL), Diagnosis (AN vs HC), and Session (1 vs 2)</b>						
Cluster FWE-correct $p < 0.01$ , individual voxel threshold 0.005, size $\geq 31$ voxels						
Left Striatum	+16.5	-4.5	+8.5	64	1.7 cm <sup>3</sup>	17.2
Right Striatum	-16.5	-13.5	+5.5	52	1.4 cm <sup>3</sup>	15.6
Left Putamen	+31.5	+16.5	-3.5	37	1.0 cm <sup>3</sup>	14.6
<b>True Difference in Subjective Value</b>						
<b>Interaction Effect of Choice (SS vs LL) and Diagnosis (AN vs HC)</b>						
Cluster FWE-corrected $p < 0.02$ , individual voxel threshold 0.01, size $\geq 41$ voxels						
Left Parietal Cortex	+54.6	+33.1	+29.7	73	2.0 cm <sup>3</sup>	14.0
Right Parietal Cortex	-59.0	+36.4	+24.7	49	1.3 cm <sup>3</sup>	16.8
Right Dorsolateral Prefrontal Cortex	-27.9	-35.6	+30.4	64	1.1 cm <sup>3</sup>	16.4
<b>Interaction Effect of Choice (SS vs LL), Diagnosis (AN vs HC), and Session (1 vs 2)</b>						
Cluster FWE-corrected $p < 0.02$ , individual voxel threshold 0.005, size $\geq 28$ voxels						
Right Striatum	+54.6	+33.1	+29.7	73	2.0 cm <sup>3</sup>	14.0
Left Striatum	-59.0	+36.4	+24.7	49	1.3 cm <sup>3</sup>	16.8

Talairach-Tournoux coordinates.

FWE, family-wise error.

**Table S10.** Comparisons of the Difference in Neural Activity between Larger-Later and Smaller-Sooner Choices

Region	Test	Statistic	<i>p</i> -value
<b>Clusters Identifying Differences by Diagnosis and Session (main test)</b>			
Bilateral Anterior Caudate	Diagnosis:Session	$F_{1,30.9} = 48.0$	<0.0001
	Diagnosis:Session 1	$t_{38} = 2.45$	0.019
	Diagnosis:Session 2	$t_{29} = -3.66$	0.0010
	Session:AN paired	$t_{13} = -5.51$	<0.0001
	Session:HC paired	$t_{13} = 4.13$	0.0012
	AN Session 1	$t_{19} = -2.97$	0.0079
	AN Session 2	$t_{16} = 3.41$	0.0036
	HC Session 1	$t_{19} = 1.00$	ns
	HC Session 2	$t_{13} = -2.00$	0.066
Dorsal Anterior Cingulate (dACC)	Diagnosis:Session	$F_{1,31.3} = 29.7$	<0.0001
	Diagnosis:Session 1	$t_{38} = 1.92$	0.062
	Diagnosis:Session 2	$t_{29} = -4.01$	0.0004
	Session:AN paired	$t_{13} = -4.5$	0.0006
	Session:HC paired	$t_{13} = 2.51$	0.026
	AN Session 1	$t_{19} = -2.76$	0.012
	AN Session 2	$t_{16} = 2.92$	0.010
	HC Session 1	$t_{19} = 0.28$	Ns
	HC Session 2	$t_{13} = -2.73$	0.017
Right Dorsolateral Prefrontal Cortex (rdIPFC)	Diagnosis:Session	$F_{1,31.4} = 15.0$	0.0005
	Diagnosis:Session 1	$t_{38} = 0.41$	Ns
	Diagnosis:Session 2	$t_{29} = -3.71$	0.0009
	Session:AN paired	$t_{13} = -3.22$	0.0067
	Session:HC paired	$t_{13} = 2.08$	0.058
	AN Session 1	$t_{19} = -1.12$	Ns
	AN Session 2	$t_{16} = 2.31$	0.035
	HC Session 1	$t_{19} = -0.33$	Ns
	HC Session 2	$t_{13} = -2.78$	0.016
Right Inferior Parietal Cortex (rPar)	Diagnosis:Session	$F_{1,30.3} = 21.2$	<0.0001
	Diagnosis:Session 1	$t_{38} = -0.15$	Ns
	Diagnosis:Session 2	$t_{29} = -3.53$	0.0014
	Session:AN paired	$t_{13} = -2.88$	0.013
	Session:HC paired	$t_{13} = 3.27$	0.0061
	AN Session 1	$t_{19} = -1.97$	0.063
	AN Session 2	$t_{16} = 2.09$	0.053
	HC Session 1	$t_{19} = -1.32$	Ns
	HC Session 2	$t_{13} = -2.71$	0.018

Region	Test	Statistic	p-value
<b>Clusters Identifying Differences by Diagnosis and Session with Absolute Difference in Subjective Value</b>			
Right Striatum	Diagnosis:Session	$F_{1,35.8} = 46.8$	<0.0001
	Diagnosis:Session 1	$t_{38} = 3.00$	0.0047
	Diagnosis:Session 2	$t_{29} = -6.95$	<0.0001
	Session:AN paired	$t_{13} = -2.69$	0.019
	Session:HC paired	$t_{13} = 8.64$	<0.0001
	AN Session 1	$t_{19} = -1.48$	0.16
	AN Session 2	$t_{16} = 3.79$	0.0016
	HC Session 1	$t_{19} = 3.21$	0.0046
	HC Session 2	$t_{13} = -6.04$	<0.0001
Left Striatum	Diagnosis:Session	$F_{1,36.2} = 47.1$	<0.0001
	Diagnosis:Session 1	$t_{38} = 3.40$	0.0016
	Diagnosis:Session 2	$t_{29} = -5.7$	<0.0001
	Session:AN paired	$t_{13} = -2.94$	0.012
	Session:HC paired	$t_{13} = 5.78$	<0.0001
	AN Session 1	$t_{19} = -2.01$	0.058
	AN Session 2	$t_{16} = 3.24$	0.0051
	HC Session 1	$t_{19} = 2.76$	0.013
	HC Session 2	$t_{13} = -4.52$	0.0006
Left Putamen	Diagnosis:Session	$F_{1,37.1} = 38.9$	<0.0001
	Diagnosis:Session 1	$t_{38} = 2.82$	0.0076
	Diagnosis:Session 2	$t_{29} = -6.31$	<0.0001
	Session:AN paired	$t_{13} = -2.39$	0.032
	Session:HC paired	$t_{13} = 6.2$	<0.0001
	AN Session 1	$t_{19} = -1.38$	0.18
	AN Session 2	$t_{16} = 3.88$	0.0013
	HC Session 1	$t_{19} = 2.71$	0.014
	HC Session 2	$t_{13} = -5.29$	0.0002
<b>Clusters Identifying Differences by Diagnosis and with True Difference in Subjective Value</b>			
Right Dorsolateral Prefrontal Cortex	HC-AN	$t_{34.7} = -3.64$	0.0009
	AN	$t_{21} = 1.83$	0.081
	HC	$t_{18} = -3.18$	0.0052
Left Parietal Cortex	HC-AN	$t_{25.0} = -3.26$	0.0032
	AN	$t_{21} = 1.01$	0.32
	HC	$t_{18} = -3.12$	0.0059
Left Parietal Cortex	HC-AN	$t_{32.0} = -3.80$	0.0006
	AN	$t_{21} = 1.54$	0.14
	HC	$t_{18} = -3.54$	0.0024

Region	Test	Statistic	<i>p</i> -value
<b>Clusters Identifying Differences by Session and Diagnosis with True Difference in Subjective Value</b>			
Right Striatum	Diagnosis:Session	$F_{1,31.8} = 60.0$	<0.0001
	Diagnosis:Session 1	$t_{38} = 2.37$	0.023
	Diagnosis:Session 2	$t_{29} = -4.8$	<0.0001
	Session:AN paired	$t_{13} = -4.63$	0.0005
	Session:HC paired	$t_{13} = 6.63$	<0.0001
	AN Session 1	$t_{19} = -2.05$	0.055
	AN Session 2	$t_{16} = 3.95$	0.0011
	HC Session 1	$t_{19} = 1.48$	0.15
	HC Session 2	$t_{13} = -2.96$	0.011
Left Striatum	Diagnosis:Session	$F_{1,36.2} = 47.1$	<0.0001
	Diagnosis:Session 1	$t_{38} = 3.40$	0.0016
	Diagnosis:Session 2	$t_{29} = -5.7$	<0.0001
	Session:AN paired	$t_{13} = -2.94$	0.012
	Session:HC paired	$t_{13} = 5.78$	<0.0001
	AN Session 1	$t_{19} = -2.01$	0.058
	AN Session 2	$t_{16} = 3.24$	0.0051
	HC Session 1	$t_{19} = 2.76$	0.013
	HC Session 2	$t_{13} = -4.52$	0.0006

## Supplemental References

1. Mazur JE, Nevin J, Rachlin H (1987): An adjusting procedure for studying delayed reinforcement. *Quantitative analyses of behavior: The effect of delay and intervening events on reinforcement value*. Vol 5. Hillsdale, NJ: Erlbaum.
2. McKerchar TL, Green L, Myerson J, Pickford TS, Hill JC, Stout SC (2009): A comparison of four models of delay discounting in humans. *Behavioural processes* 81(2): 256-9.
3. Benhabib J, Bisin A, Schotter A (2010): Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economics Behavior* 69(2): 205-223.
4. MATLAB: version 8.0.0.783 (R2012b). Natick, Massachusetts: The MathWorks Inc.
5. Doll BB, Jacobs WJ, Sanfey AG, Frank MJ (2009): Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. *Brain Res* 1299:74-94.
6. Green L, Fry AF, Myerson J (1994): Discounting of Delayed Rewards. A Life-Span Comparison. *Psychol Sci* 5:33-36.
7. Reynolds B, Schiffbauer R (2004): Measuring state changes in human delay discounting: an experiential discounting task. *Behav Processes* 67:343-56.
8. R Core Team (2013): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
9. Bates D, Maechler M, Bolker B, Walker S (2013): lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-5. <http://CRAN.R-project.org/package=lme4>.
10. Singmann H (2013): afex: Analysis of Factorial Experiments. R package version 0.6-82. <http://cran.r-project.org/package=afex>.
11. Lenth RV (2014): lsmeans: Least-Squares Means. R package version 2.00-1. <http://cran.r-project.org/package=lsmeans>.
12. Fox J, Weisberg S (2011): *An {R} Companion to Applied Regression, Second Edition*. Thousand Oaks CA: Sage. URL: <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>.
13. Højsgaard UHS (2013): pbkrtest: Parametric bootstrap and Kenward Roger based methods for mixed model comparison. R package version 0.3-8. <http://CRAN.R-project.org/package=pbkrtest>.
14. Barr DJ, Levy R, Scheepers C, Tily HJ (2013): Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang* 68:255-278.