

## **An investigation of habit learning in Anorexia Nervosa**

### **Abstract**

Anorexia Nervosa (AN) is a disorder characterised by compulsive behaviour, such as self-starvation and excessive exercise, which develop in the pursuit of weight-loss. Recent theory suggests that once established, compulsive weight-loss behaviours in AN may become habitual. In two parallel studies, we measured whether individuals with AN showed a bias toward habits using two outcome-devaluation tasks. In Study 1, 23 women with AN (restrictive and binge/purge subtypes), and 18 healthy controls (HC) completed the slips-of-action paradigm, designed to assess reward-based habits. In Study 2, 13 women with restrictive AN, 14 women recovered from restrictive AN, and 17 female HC participants completed the slips-of-action paradigm, and an avoidance paradigm, designed to assess aversive habits. AN participants showed no deficit relative to HCs in the ability to use feedback to respond correctly to stimuli. Following devaluation of outcomes, all groups in both studies were equally able to withhold inappropriate responses, suggesting no deficit in the balance between goal-directed and habitual control of behaviour in these tasks in AN. These results suggest that individuals with AN do not show a generalised tendency to rely on habits in two outcome-devaluation tasks. Future research is needed to investigate the potential role of disorder-specific habits in the maintenance of behaviour in AN.

**Key Words:** Anorexia Nervosa; Habit Formation; Compulsivity; Goal-directed learning; Eating Disorders.

## 1. **Introduction**

Anorexia nervosa (AN) is a severely debilitating psychiatric disorder characterized by an intense fear of weight gain or becoming fat, despite significantly low body weight (American Psychiatric Association, 2013). Individuals with AN place extreme over-importance on the control of weight and shape, and often have disturbed body image perception (Fairburn et al., 2003). These distorted beliefs and perceptions are accompanied by a perpetual drive for thinness and continuous lowering of weight goals (Barbarich-Marsteller et al., 2011). The characteristic behaviours seen in AN to achieve weight-loss goals, such as extreme dietary restriction and over-exercise, have been described as evidence of the compulsive nature of the disorder (Godier and Park, 2014a; Park et al., 2014). Furthermore, individuals with AN show cognitive inflexibility (Tchanturia et al., 2004), a rigid cognitive style suggested to contribute to compulsivity (Fineberg et al., 2010).

Compulsivity can be defined as a trait in which actions are persistently repeated despite adverse consequences (Robbins et al., 2012). This can be seen in repetitive, and highly ritualised behaviours of OCD, which impair patients ability to engage in normal daily activities (American Psychiatric Association, 2013), and in the lack of control felt over drug-seeking behaviour in substance dependence, despite the adverse consequences (Kalivas and Volkow, 2005). Compulsive behaviour in AN has been compared to both OCD (Steinglass and Walsh, 2006), and addiction (Barbarich-Marsteller et al., 2011; Godier and Park, 2014a, b, 2015; Kaye et al., 2013; Park et al., 2014; Scheurink et al., 2010; Zink and Weinberger, 2010). Indeed, studies using the Iowa Gambling task in participants with AN, OCD and substance dependence, suggest in all three disorders a tendency to make disadvantageous decisions when choosing between immediate or long terms gains (Lawrence et al., 2006; Tchanturia et al., 2007; Verdejo-Garcia et al., 2007), which may be linked to the compulsive,

self-destructive and sometimes impulsive behaviours seen across these disorders (Tchanturia et al., 2007). Impulsivity, defined as the tendency to perform actions prematurely without foresight (Dalley et al., 2011), has already been directly associated with engaging in binge-purge behaviours compared to restrictive behaviour (Claes et al., 2005; Favaro et al., 2005; Rosval et al., 2006; Waxman, 2009). The present study aimed to assess compulsivity more directly in AN, using tasks for which poor performance has been associated with compulsive behavior in disorders such as OCD and addiction (Gillan et al., 2015; Gillan et al., 2013; Gillan et al., 2011; Sjoerds et al., 2013).

Emergent evidence suggests that compulsivity may arise, at least in part, as a result of over-reliance on habit-learning, at the expense of more considered modes of action selection. Habits are learnt (instrumental) behaviours that have been engaged in repeatedly and consequentially become fixed, occur without conscious effort, and can be elicited by external stimuli (Graybiel, 2008). Habit ('stimulus-response') learning can be contrasted with goal-directed ('action-outcome') control (Robbins et al., 2012). Goal-directed behaviours are purposeful actions driven by anticipation and evaluation of a rewarding outcome. As such, goal-directed actions are less likely to be performed if the value of their associated outcomes is lessened (Balleine and O'Doherty, 2010). However, if these new actions are engaged in repeatedly (over-trained), this may lead to the formation of stimulus-response associations, such that external stimuli can trigger habitual responses even when the consequences are no longer rewarding (Dickinson, 1985).

A shift in balance away from goal-directed control and towards excessive habit learning has been shown in substance dependence (Sjoerds et al., 2013; Voon et al., 2014), OCD (Gillan et al., 2015; Gillan et al., 2013; Gillan et al., 2011; Voon et al., 2014), binge eating disorder (BED) (Voon et al., 2014), and Tourette's syndrome (Delorme et al., 2015). Walsh (2013) outlines the mechanisms by which aberrant habit formation may contribute to

the maintenance of dietary restriction in AN. Restrictive eating may begin as the result of goal-directed weight-loss behaviour, in which behaviour is associated with a rewarding outcome (weight loss). If restrictive eating behaviour is repeated enough it may become relatively independent of reward, such that weight loss as a rewarding outcome may be needed only intermittently, or even no longer necessary for this behaviour to continue. Habitual behaviour, as measured by the persistence of a devalued action may be reflective of the treatment resistance often observed in individuals with AN (Walsh, 2013).

The two studies presented here were exploratory in nature, and aimed to begin to test the hypothesis that a generalised reliance on habits, as seen in other compulsive disorders, may contribute to the development of the compulsive weight-loss behavior, within a small group of individuals with AN. These studies were carried out in parallel at the New York State Psychiatric institute (Study 1) and at the University of Oxford (Study 2). In Study 1, we studied individuals with restrictive and binge/purge subtype AN, and compared them to healthy controls. We used a simplified version of the outcome-devaluation task previously used to provide evidence for reliance on habits in OCD (Gillan et al., 2011), namely the Slips-of-Action paradigm (for simplified version of the task see Worbe et al, 2015). In Study 2, we compared individuals both currently ill and recovered from restrictive AN (as starvation alone is associated with severe alternations in cognitive and physiological systems (Cowdrey et al., 2011; Kaye et al., 2009; Wagner et al., 2008), to healthy controls on the Slips-of-Action paradigm, replacing the fruit pictures with pictures of animals, in order to avoid the confound of food stimuli in the AN participants. In addition, an adapted version of an avoidance habit task used previously by Gillan et al (2013, 2015) was employed to further explore habit bias in AN, and whether this is modulated by valence, i.e. appetitive versus aversive learning (Gillan et al., 2015; Gillan et al., 2013). This is an important consideration as AN features both avoidance behaviour; i.e. an aversion to energy-dense foods (Cowdrey et

al., 2013), which are experienced as anxiogenic (Bailer et al., 2012; Zink and Weinberger, 2010), as well as the appetitive behaviour; i.e pursuit of reward in the form of weight-loss (Godier and Park, 2014a). We hypothesised that individuals with current and past AN would show enhanced habit formation in each of the tasks, in both studies, evidenced by a persistence of previously learned responses despite devaluation of the outcome.

## **2. Methods**

### **2.1 Study 1**

#### **2.1.1 Participants**

Forty-one participants were recruited for two groups: women with a current diagnosis of AN (AN group,  $n=23$ ), and healthy control subjects (HC group,  $n=18$ ). A power analysis indicated a 98% chance of detecting a significant effect based on a Cohen's  $d$  of 1.32 calculated from a previous study using the Slips-of-Action paradigm (Gillan et al, 2011). Subjects were recruited by advertisements, the clinic website, clinician referral, and word of mouth. HC participants had no current or lifetime Axis I or II diagnoses and no exposure to psychotropic medications or psychotherapy. See supplementary materials for full inclusion and exclusion criteria. This study was approved by the New York State Psychiatric Institute Institutional Review Board.

#### **2.1.2. Procedures**

Individuals with AN were tested within 3 weeks of hospital admission. All participants provided written informed consent to partake in this study. First, they were administered semi-structured psychiatric interviews by trained research staff and completed self-report questionnaires (see below), followed by the Slips-of-Action paradigm. Height and weight were measured to calculate BMI.

#### **2.1.3. Measures**

The Structured Clinical Interview for the Diagnostic Statistical Manual IV (SCID) (Spitzer et al., 2006) was used to screen for DSM-IV Axis-I disorders. Eating disorder symptoms were measured using the global mean scores on the Eating Disorder Examination (EDE) (Fairburn et al., 2008a) and Eating Disorder Examination Questionnaire (EDE-Q) (Fairburn and Beglin,

2008). Depressive symptoms were measured using the Beck Depression Inventory (BDI-II) (Beck et al., 1996). Anxiety symptoms were measured using the State Trait Anxiety Scale (STAI) (Spielberger et al., 1983a). Impulsivity was measured using the Barratt Impulsiveness Scale (BIS) (Patton et al., 1995). Internal consistency values for the measures used can be found in the Supplementary materials.

#### **2.1.4. Slips-of-Action paradigm**

A validated and shortened version of the original ‘Fabulous Fruit Task’ designed to investigate goal-directed behaviour and habit learning was used (de Wit et al., 2012; Gillan et al., 2011) (see Worbe et al, 2015 for the simplified version). The task involves 3 stages: instrumental discrimination training, simple outcome devaluation choice test, baseline test vs. Slips-of-Action test (see Figure 1). In the initial instrumental training stage of this paradigm, participants learned by trial-and-error which responses led to rewarding outcomes in the presence of different discriminative stimuli. In the outcome-devaluation test, some of the outcomes were devalued, and participants had to use their knowledge of the response–outcome (R-O) relationships to respond only to still-valuable outcomes. In the subsequent slips-of-action test, participants were shown the stimuli from the training stage, and were asked to selectively respond to stimuli that signaled the availability of still-valuable outcomes. A reliance on habits was indicated by a perseverance of responses to stimuli that signaled devalued outcomes. Finally, a baseline test was identical in all respects except that stimuli were devalued instead of outcomes, and was included in order to control for general cognitive control functioning. For a full description of the task stages see supplementary materials.

After the task participants were asked to complete questionnaires assessing their knowledge of the stimulus-response-outcome contingencies learnt during the instrumental training stage.

They were also asked to rate their confidence in their answers on a 1 to 100 scale.

[Figure 1]

## **2.2 Study 2**

### **2.2.1 Participants**

Forty-four female participants were recruited for 3 experimental groups: current restrictive AN (AN-C group, n=13), recovered from restrictive AN (AN-R group, n=14), and healthy controls (HC, n=17). The effects of starvation are associated with severe alterations in cognitive and physiological systems (Cowdrey et al., 2011; Kaye et al., 2009; Wagner et al., 2008), therefore we included a group of individuals fully recovered from AN to separate any impairments associated with a history of AN from possible starvation effects on task performance. A power analysis was calculated based on this sample size and a Cohen's d of 1.32 from a previous study using the Slips-of-Action paradigm (Gillan et al, 2011). Power analysis indicated a 93% chance of detecting a significant effect between the AN-C and HC group and a 94% chance between the AN-R and HC groups. Participants were recruited via email, internet and poster advertisement. In addition, a number of participants in the AN-C and AN-R group were recruited from the Oxford Research List for Anorexia Nervosa, which is maintained by the research team in Oxford. General exclusion criteria included age <18 or >60, insufficient English language skills, male sex, and left-handedness. See supplementary materials for full inclusion and exclusion criteria for each experimental group. Ethical permission for this study was obtained from the South Central – Oxford A Research Ethics Committee.

### **2.2.2 Procedures**

After obtaining informed consent, trained researchers administered psychiatric interviews,



and participants completed a battery of self-report questionnaires (described below). Height and weight were taken to calculate BMI. Participants subsequently attended one or two further sessions to complete the two experimental tasks. The order of the tasks was counterbalanced across participants.

### **2.2.3 Measures**

Participants completed the same interview and questionnaire measures as those outlined in Study 1. In addition, compulsivity was indexed using Obsessive-Compulsive Inventory Revised (OCI-r) (Foa et al., 2002), and the Yale-Brown-Cornell Eating Disorder Scale Self-Report Questionnaire (YBC-EDs-SRQ) (Bellace et al., 2012). Clinical impairment was indexed using the Clinical Impairment Assessment (CIA) (Bohn and Fairburn, 2008). Verbal IQ was assessed using the National Adult Reading Test (NART) (Nelson, 1982). Internal consistency values for these additional measures can be found in the Supplementary materials.

### **2.2.4 The Slips-of-Action Paradigm**

This task was identical to that employed in Study 1, except that fruit pictures were replaced with non-food stimuli (e.g., cartoon animals) to avoid the confound of food stimuli in those with AN.

### **2.2.5 Noise Avoidance Task**

In addition to a thorough investigation of appetitive habits using the Slips-of-Action Paradigm in both studies, we also carried out an exploratory investigation of avoidance habits in AN. Avoidance habits were assessed using the Noise Avoidance Task (see Figure 2), an adapted version of a shock avoidance task used by Gillan et al (2013, 2015) in patients with OCD (Gillan et al., 2015; Gillan et al., 2013). The task consisted of four stages: a brief

training session, followed by a devaluation test in extinction (devaluation sensitivity test), an extended training session, and a final devaluation test in extinction (habit test). In the training sessions participants were instructed to make a right or left response to two different coloured rectangles to avoid hearing an unpleasant noise in their left or right ear. In the subsequent devaluation stages, one of the outcomes was devalued, and participants were instructed to continue to respond only to the still-valuable stimulus. Unpleasant noise outcomes were devalued by disconnecting the headphone from one of the ears in full view of the participants. A reliance on habits was indicated by a perseverance of responses to devalued stimuli. See supplementary materials for a full description of the task.

Following the task, participants were tested on their explicit knowledge of stimulus-action-outcome associations experienced during training. Participants also retrospectively rated VAS scales from 0 to 100 probing 1) their level of expectancy that a shock would follow the devalued CS; 2) the extent to which they experienced an urge to continue responding in spite of the devaluation; and 3) the extent to which they actively attempted to suppress this urge during the extinction test.

[Figure 2]

## **2.3 Data Analysis**

Analysis of both studies was carried out in SPSS. Group comparisons of demographic and clinical measures were carried out using one-way ANOVAs, with Bonferroni corrected pairwise comparisons. The threshold for significance was set at  $p=0.05$  for all analyses.

### **2.3.1 Slips-of-Action Paradigm.**

In both studies, repeated-measures ANOVAs were used to assess performance on the 8 training blocks in the instrumental training phase. Performance on the instructed outcome devaluation test was analyzed using an independent-samples t-test to determine differences between the AN and HC groups in Study 1, and a one-way ANOVA to determine differences between the three groups (AN-C, AN-R and HC) in Study 2. In both studies, data from the baseline and slips-of-action test were combined for a repeated-measures ANOVA, with the within-subject factors test type (slips-of-action versus baseline), and devaluation (valued versus devalued), and the between-subjects factor group (HC vs AN in Study 1, comparisons between AN-C, AN-R and HC and Study 2).

### **2.3.2 Noise Avoidance Task.**

A one-way ANOVA was used to assess performance during the training stages of the Noise Avoidance task in Study 2. Following the extended training stage, a repeated-measures ANOVA was carried out to assess differences in response to valued and devalued stimuli between the three groups (AN-C, AN-R and HC).

### **2.3.3 Covariates and additional analyses**

As age significantly differed between groups in both studies, and years of education differed between groups in Study 1 (See Table 1), analyses were repeated including age and years of education as covariates. To control for the potential effect of recent weight gain during

treatment in the AN group in Study 1, change in BMI during treatment was included as a covariate for the analysis in Study 1. In order to investigate any effect of AN subtype, analysis of Study 1 was repeated using the subtypes of AN as the grouping variable. Additionally, in Study 2, as a number of participants in the AN group were taking antidepressant and/or antipsychotic medication, the analysis was repeated using medication status as the grouping variable. As some of the participants in the AN group in Study 2 had diagnoses of depression, BDI scores were also included as a covariate.

Ability to recall explicit stimulus-response-outcome contingencies was assessed using independent samples t-tests in Study 1, and one-way ANOVAs in Study 2. Spearman's Rho with bonferroni correction was used to assess correlations between task performance and clinical/questionnaire measures.

#### **2.3.4 Bayesian Analyses.**

In addition to the above analyses, we carried out a Bayesian statistical analysis on the data from both studies. In Study 1, a Bayesian independent-samples t-test was performed on responses to devalued stimuli and the difference scores (valued minus devalued responses) in the slips-of-action test. In Study 2, Bayesian ANOVAs were carried out on responses to devalued stimuli and the difference scores for both tasks. An advantage of Bayesian analysis is that it allows a comparative approach of the probability of the null and alternative hypotheses given the observed data, whereas frequentist methods only provide information regarding the null hypothesis (Jarosz and Wiley, 2014). The analysis was carried out in JASP (Love et al., 2015). We used the program's default option of a Cauchy prior, and defined the width as .707. Bayes Factor ( $BF_{01}$ ) values provide an indicator of how many times more likely the null hypothesis is to the alternative hypothesis (Jarosz and Wiley, 2014). For example, a  $BF_{01}$  of 3-10 is considered substantial evidence for the null hypothesis, and

suggests this is 3-10 times more likely than the alternative. A value of above 10 would be considered strong evidence in favour of the null hypothesis, whereas a value below 3 would provide only weak evidence that the null hypothesis is more likely than the alternative (Jeffreys, 1961).

### **3. Results**

#### **3.1 Study 1**

##### **3.1.1. Demographic results**

Table 1 summarizes the demographic and psychological characteristics of the two experimental groups.

[Table 1]

##### **3.1.2 Slips-of-Action Paradigm**

###### ***3.1.2.1 Instrumental Discrimination Training and Instructed Outcome-Devaluation Test.***

The groups acquired the instrumental discriminations at the same rate as reflected both in accuracy,  $F(2,173)=1.384$ ,  $p=0.212$ , and in reaction times,  $F(7,273)=1.573$ ,  $p=0.143$ .

Subsequently, both groups performed at the same level on the instructed outcome-devaluation test,  $t(39)=-0.402$ ,  $p=0.690$ , suggesting they were equally able to direct responses towards a still-valuable outcome, and away from a devalued outcome.

###### ***3.1.2.2 Slips-of-action versus baseline test.***

The analysis failed to show a Group\*Devaluation\*Test type interaction, or a Group\*Devaluation interaction,  $F_s < 1$ . Therefore, individuals in the AN group showed no difference in performance compared to HC on either task, and neither group were impaired at withholding previously rewarded responses when the associated outcome was devalued (see Figure 3). Subsequent analysis confirmed there were no group differences in the response to the valued ( $t[39]=-0.279$ ,  $p=0.782$ , Cohen's  $d=0.09$ ) or devalued stimuli ( $t[39]=-0.935$ ,  $p=0.356$ , Cohen's  $d=0.29$ ) on the slips-of-action test. In addition there was no significant group effect in the difference scores (i.e. valued-devalued responses;  $t[39]=-0.389$ ,  $p=0.699$ , Cohen's

$d=0.260$ ). A significant main effect of devaluation was found across the slips-of-action and baselines tests,  $F(1,39) = 396.217$ ,  $p < 0.001$ , indicating participants were able to withhold responses associated with devalued events, whilst continuing to perform the valued responses. There were no significant correlations between performance on the slips-of-action task and any clinical/questionnaire measures.

[Figure 3]

### ***3.1.2.3 Additional Analyses***

Including age, years of education and change in BMI during treatment as covariates yielded the same pattern of results with, no significant Group\*Devaluation\*Test type interaction, nor a Group\*Devaluation interaction,  $F_s < 1$ .

Comparison of AN subtypes revealed a near-significant Subgroup\*Devaluation\*Test type interaction,  $F(1,21) = 4.338$ ,  $p = 0.050$ . A trend towards poorer discrimination between valued and devalued responses was observed in the restrictive subgroup (mean percentages of responding: valued=85.8%, devalued=39.5%) relative to the binge-purge subgroup (valued=91.5%, devalued=20.4%). Separate analysis of the slips-of-action test failed to confirm this, as the Group\*Devaluation interaction was not significant,  $F(21) = 2.944$ ,  $p = 0.101$ .

The AN and HC groups performed equally in the test of explicit contingency knowledge,  $p > 0.050$  (note: data from one AN participant were missing from this analysis).

A number of the task outcomes were found to be non-normal, however, non-parametric tests confirmed the main results (see supplementary materials).

## **3.2 Study 2**

### **3.2.1 Demographic characteristics**

Table 2 summarises the demographic and psychological characteristics of the three groups.

[Table 2]

### **3.2.2 Slips-of-Action Paradigm**

#### ***3.2.2.1 Instrumental Discrimination Training and Instructed Outcome-Devaluation Test***

The groups acquired the instrumental discriminations at the same rate, as reflected both in accuracy,  $F(8.3, 166.1)=0.834$ ,  $p=0.577$ , and reaction times,  $F(6.6, 131.6)=1.171$ ,  $p=0.325$ .

Both groups also performed at the same level on the instructed outcome-devaluation test,  $F(2,40)=.044$ ,  $p=0.957$ , suggesting they were equally able to base their choices on the relative value of two outcomes.

#### ***3.2.2.2 Slips-of Action and Baseline test***

There was no significant Group\*Devaluation\*Task interaction, nor Group\*Devaluation interaction found ( $F_s < 1$ ), indicating there was no difference between the three groups in either task in the ability to withhold previously rewarded responses when the outcome is devalued (see Figure 4). Subsequent analysis confirmed that the groups did not differ in their response to the still-valuable ( $F[2,40]=0.391$ ,  $p=0.679$ , Cohen's  $d=0.30$ ) or devalued stimuli ( $F[2,40]=0.689$ ,  $p=0.508$ , Cohen's  $d=0.19$ ) on the slips-of-action test. In addition there was no significant difference in the difference scores between groups,  $F(2,40)=0.467$ ,  $p=0.630$ .

There was a significant main effect of devaluation,  $F(1,40)=291.822$ ,  $p<0.001$ , indicating all participants were able to withhold responses associated with devalued events, and continue to respond to valuable events in both the slips-of-action and baseline tests. There were no



significant correlations between performance on the slips-of-action task and any clinical/questionnaire measures.

[Figure 4]

### 3.2.3 Noise Avoidance Task

#### 3.2.3.1 Training Stage and devaluation sensitivity

All groups were equally proficient in learning the contingencies during the initial training stage of the task,  $F(2,40)=0.611$ ,  $p=0.548$ . Prior to the extended training session, there was no Group\*Devaluation interaction,  $F(2,40)=1.058$ ,  $p=0.357$ , suggesting all groups were equally able to withhold unnecessary responses to stimuli with a devalued outcome.

#### 3.2.3.2 Habit Test

Following the extended training session, there was a near significant Group\*Devaluation interaction,  $F(2,40)=3.135$ ,  $p=0.054$ . However, post-hoc comparisons revealed a significant group difference in response to the *still-valuable stimuli* only ( $F[2,40]=3.767$ ,  $p=0.032$ , Cohen's  $d=0.77$ ), with increased responding to the still-valuable stimuli in the AN-R (mean:3.42; SD: 0.55) compared to the AN-C (mean:2.51; SD: 1.38) group ( $p=0.033$ ). There were no differences in the response to still-valuable stimuli between the HC group and the AN-C ( $p=0.141$ ) or AN-R group ( $p=0.264$ ). A near-significant group-effect was also found in the difference scores,  $F(2,40)=3.135$ ,  $p=0.054$ , however this was again driven by a near significant difference between the AN-R and AN-C group ( $p=0.058$ ) only, and likely reflects the difference in response to the still-valuable stimuli. Importantly, there were no significant differences in response to the devalued stimuli between groups ( $F[2,40]=0.044$ ,  $p=0.957$ , Cohen's  $d = 0.28$ ), indicating no difference in the perseverance of response to devalued stimuli. This is supported by a strong devaluation effect across all groups,  $F(1,40)=404.414$ ,

$p < 0.001$ , indicating that all participants again responded more to the still-valuable stimulus than the stimulus with a devalued outcome (see Figure 5). There were no significant correlations between performance on the habit test and any clinical/questionnaire measures.

[Figure 5]

### **3.2.3.3 Additional Analyses**

Including age, years of education and depression (BDI scores) as covariates for both tasks yielded the same pattern of results, with no significant Group\*Devaluation interaction (Slips-Of-Action Paradigm:  $F[2,33]=0.543$ ,  $p=0.586$ ; Noise Avoidance Task:  $F[2,33]=1.235$ ,  $p=0.304$ ).

When comparing the medicated and unmedicated participants there was again no Group\*Devaluation interaction in either task (Slips-Of-Action Paradigm:  $F[1,25]=0.529$ ,  $p=0.474$ ; Noise Avoidance Task:  $F[1,24]=1.337$ ,  $p=0.259$ ).

For both tasks, there were no differences between groups in explicit knowledge of task contingencies ( $F_s < 1$ ).

A number of the task outcomes were found to be non-normal, however, non-parametric tests confirmed the main results (see supplementary materials), and in contrast to the parametric tests no difference was observed in responses to still valuable stimuli between groups.

### **3.3 Bayesian Analyses**

Bayes Factor ( $BF_{01}$ ) values are provided in Table 3 for both studies. With the exception of the difference scores in the Noise Avoidance Paradigm in Study 2, the  $BF_{01}$  values suggest substantial evidence in favour of the null hypothesis (Jeffreys, 1961), suggesting there are no differences between groups in responses to devalued stimuli in these tasks. The  $BF_{01}$  value for the difference scores in the Noise Avoidance Paradigm suggests only anecdotal evidence for the null hypothesis (Jeffreys, 1961), likely due to the increased response in the AN-R group to the still valuable stimuli reported above. All other values suggest the null hypothesis is between 3 and 5.7 times more likely than the alternative hypothesis.

[Table 3]

#### 4. **Discussion**

To our knowledge these studies are the first to directly investigate the balance between goal-directed and habitual control of behaviour in two established outcome-devaluation tasks in AN. Using two participant cohorts we found no evidence for a bias towards developing appetitive habits in patients with current or past AN. All participants were equally successful at using feedback to learn the correct stimulus-response contingencies in the initial training phase of the tasks, and explicit knowledge of contingencies was equal across groups. Following instructed devaluation of outcomes, participants with current or past AN, and healthy controls, were equally able to withhold responses for devalued outcomes in all tasks, suggesting intact goal-directed learning. In addition, no correlations were found between performance on the tasks and any of the clinical measures or questionnaires in either study. The addition of an avoidance based habit task yielded the same results, suggesting a tendency towards habits is not dependent on valence in AN.

It is emphasized that this study was exploratory and the limited sample size means the results can only be considered preliminary, and require replication in larger sample sizes. However, our results were further supported by non-parametric tests, and a Bayesian analysis, which suggested the null hypothesis (e.g. no group differences in responses to devalued stimuli) was between 3 and 5.7 times more likely than the alternative hypothesis (e.g. between-group difference in responses to devalued stimuli).

The findings from these combined studies suggest that AN may not involve a generalized vulnerability to habitual responding in these paradigms. Our findings in this population differ from studies involving individuals with OCD and addictive disorders, in which excessive habitual responding has been found using these tasks (Gillan et al., 2015;

Gillan et al., 2013; Gillan et al., 2011; Sjoerds et al., 2013). This is surprising considering suggestions of both behavioural and neurobiological parallels between AN and these disorders (Barbarich-Marsteller et al., 2011; Godier and Park, 2014a, b, 2015; Kaye et al., 2013; Park et al., 2014; Scheurink et al., 2010; Zink and Weinberger, 2010). However, whilst these studies suggest that AN may not involve a *generalised* vulnerability to forming habits, disorder-specific habits in AN warrant further investigation. Individuals with AN experience intense reward from the pursuit of thinness (Park et al., 2014), reflected by increased salience and neural response to disorder-related stimuli in reward and habit-related regions in AN (Cowdrey et al., 2011; Fladung et al., 2010; Fladung et al., 2013; Foerde et al., 2015; Giel et al., 2013). Food restriction is linked to an upregulation of reward (Fulton et al., 2004), and as such this may further increase the rewarding value of weight-loss behaviour in individuals with AN. Thus, the increased reward value associated with weight-loss may lead to the repetition of weight-loss behaviour over time, and the development of habits in a way that is not captured by these tasks.

Whilst the small sample size in these studies limits the generalizability of our findings to the wider AN population, the combination of Study 1 and Study 2, which were conducted entirely separately, strengthens the interpretation of these null findings. Study 1 suggested no food-cue specific impairment in goal-directed responding in inpatients with AN. This study included both subtypes of AN, and suggested a possible impairment in the restrictive compared to binge/purge subtype; however using a purely restrictive AN sample in Study 2 did not indicate any impairment in this group compared to healthy controls. Study 2 also extended the findings of Study 1, using non-food-related stimuli, and the addition of an avoidance-based paradigm, making it unlikely that the null results in Study 1 were task-specific.

Whilst the studies were conducted in parallel, they were designed separately, and as such some important differences between studies should be considered. Measures of obsessive-compulsive symptoms and clinical impairment were only included in Study 2, limiting the potential for comparison across the sample. Furthermore, whilst the AN sample in Study 1 were current inpatients, the AN-C sample in Study 2 were not. However, a number of the AN-C participants in Study 2 had previously been treated as inpatients, outpatients and day patients. Importantly, no differences were found in any of the clinical measures of severity, BMI or duration of illness, suggesting that despite this, the two samples appear to be comparable. A difference in age was found between the acute AN samples in each study, and between experimental groups in both Study 1 and 2. In Study 1 the HC group were significantly older than the AN group, and this pattern was reversed in Study 2. Healthy ageing is associated with impairments on the Slips-of-Action paradigm (de Wit et al., 2014), and so could in theory have masked habit biases in the AN group in Study 1. However, including age as a covariate in both analyses had no effect on the pattern of results.

A number of participants in the AN groups had diagnoses of depression, and were taking serotonergic antidepressants (SSRI's and SNRI's) or antipsychotics. Animal research has shown decreased sensitivity to outcome-devaluation as a result of serotonin receptor antagonism, or serotonin depletion (Altman and Normile, 1986; Clarke et al., 2007). Furthermore, a recent study investigating tryptophan (a precursor to serotonin) depletion in humans suggested this promoted a reliance on habits in the Slips-of-Action paradigm (Worbe et al., 2015), and so an effect of serotonergic medication cannot be ruled out. However, previous research suggests no difference in the persistence of devalued responses between medicated and unmedicated OCD patients (Gillan et al., 2013). In addition, including depression as a covariate, and comparison of medicated and un-medicated participants in Study 2, indicated medication status did not have an effect.

It is important to emphasize the exploratory nature of these studies, which limited the power to detect differences in the tasks used. However, the consistency in findings across the two studies, and the addition of a Bayesian analysis providing further evidence for the null hypothesis, and adds weight to the robustness of our observation that individuals with AN do not display an aberrancy in habit formation two established outcome-devaluation tasks. Replication in larger samples will be needed to support the preliminary conclusions drawn in this study. Future research may also benefit from the use of alternative experimental tasks used to assess deficits in goal-directed control of behaviour in AN, such as those that require participants to track changing contingencies and outcomes on a trial-by-trial basis (Voon et al., 2014).

In sum, this is the first reported study to assess the balance between goal-directed and habitual control of behaviour in individuals with current or past AN, and provided no evidence for a reliance on either appetitive or avoidance habits in AN in these tasks. Whilst the data from these tasks suggest AN does not involve a generalised reliance on habits in learning, it remains to be seen whether habit formation is engaged when it comes to disorder-specific behaviours. Replication of these preliminary results require replication with larger sample sizes to be confident of our conclusions.

## **Acknowledgements**

This work (Study 1) was supported by National Institute of Mental Health K23MH076195 (Principal Investigator: Steinglass), Klarman Family Foundation, and by New York State. The research reported in Study 2 was funded by a MRC Confidence in Concept grant (ref: CiC12) awarded to Dr Park and was conducted at the Department of Psychiatry, University of Oxford. Lauren Godier is funded by an MRC Ph.D. Studentship. Dr Gillan is supported by a Sir Henry Wellcome Postdoctoral Fellowship (101521/Z/12/Z). Dr. de Wit is supported by a Vidi grant 452-13-006 of the Netherlands Organization for Scientific Research (NWO). The authors would like to thank Angelos Krypotos for his help with the Bayesian Analyses, and Ariana Dichiara for her assistance with data management in Study 1.

## **Financial Disclosures**

Dr. Steinglass receives funding from the Global Foundation for Eating Disorders and the National Eating Disorder Association. Dr Godier, Dr de Wit, Dr Pinto, Dr Gillan, Dr Greene, Dr Scaife, Dr Simpson, Professor Walsh, and Dr Park report no biomedical financial interests or potential conflicts of interest.



## **References**

- Altman, H.J., Normile, H.J., 1986. Enhancement of the memory of a previously learned aversive habit following pre-test administration of a variety of serotonergic antagonists in mice. *Psychopharmacology (Berl)* 90 (1), 24-27.
- American Psychiatric Association, 2013. *Diagnostic and Statistical Manual of Mental Disorders (Fifth Ed)*, Fifth ed. American Psychiatric Publishing, Arlington, VA.
- Bailer, U.F., Narendran, R., Frankle, W.G., Himes, M.L., Duvvuri, V., Mathis, C.A., Kaye, W.H., 2012. Amphetamine induced dopamine release increases anxiety in individuals recovered from anorexia nervosa. *International Journal of Eating Disorders* 45 (2), 263-271.
- Balleine, B.W., O'Doherty, J.P., 2010. Human and rodent homologues in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35 (1), 48-69.
- Barbarich-Marsteller, N.C., Foltin, R.W., Walsh, B.T., 2011. Does anorexia nervosa resemble an addiction? *Current drug abuse reviews* 4 (3), 197-200.
- Beck, A.T., Steer, R.A., Brown, G., 1996. *Manual for the Beck Depression Inventory-II*. Psychological Corporation, San Antonio, TX.
- Belin, D., Economidou, D., Pelloux, Y., Everitt, B.J., 2011. Habit Formation and Compulsion, in: Olmstead, M.C. (Ed.), *Animal Models of Drug Addiction*. Humana Press, pp. 337-378.
- Bellace, D.L., Tesser, R., Berthod, S., Wisotzke, K., Crosby, R.D., Crow, S.J., Engel, S.G., Le Grange, D., Mitchell, J.E., Peterson, C.B., Simonich, H.K., Wonderlich, S.A., Halmi, K.A., 2012. The Yale-Brown-Cornell eating disorders scale self-report questionnaire: A new, efficient tool for clinicians and researchers. *International Journal of Eating Disorders* 45 (7), 856-860.
- Bohn, K., Fairburn, C.G., 2008. The Clinical Impairment Assessment Questionnaire (CIA 3.0), in: Fairburn, C.G. (Ed.), *Cognitive behaviour therapy and eating disorders*. Guildford Press, New York.
- Claes, L., Vandereycken, W., Vertommen, H., 2005. Impulsivity-related traits in eating disorder patients. *Personality and Individual Differences* 39 (4), 739-749.
- Clarke, H.F., Walker, S.C., Dalley, J.W., Robbins, T.W., Roberts, A.C., 2007. Cognitive inflexibility after prefrontal serotonin depletion is behaviorally and neurochemically specific. *Cerebral Cortex* 17 (1), 18-27.
- Cowdrey, F.A., Finlayson, G., Park, R.J., 2013. Liking compared with wanting for high- and low-calorie foods in anorexia nervosa: aberrant food reward even after weight restoration. *American Journal of Clinical Nutrition* 97 (3), 463-470.
- Cowdrey, F.A., Park, R.J., Harmer, C.J., McCabe, C., 2011. Increased Neural Processing of Rewarding and Aversive Food Stimuli in Recovered Anorexia Nervosa. *Biol Psychiatry* 70 (8), 736-743.
- Dalley, J.W., Everitt, B.J., Robbins, T.W., 2011. Impulsivity, Compulsivity, and Top-Down Cognitive Control. *Neuron* 69 (4), 680-694.
- de Wit, S., van de Vijver, I., Ridderinkhof, K.R., 2014. Impaired acquisition of goal-directed action in healthy aging. *Cognitive, Affective, & Behavioral Neuroscience* 14 (2), 647-658.
- de Wit, S., Watson, P., Harsay, H.A., Cohen, M.X., van de Vijver, I., Ridderinkhof, K.R., 2012. Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control. *Journal of Neuroscience* 32 (35), 12066-12075.

- Delorme, C., Salvador, A., Valabrègue, R., Roze, E., Palminteri, S., Vidailhet, M., de Wit, S., Robbins, T., Hartmann, A., Worbe, Y., 2015. Enhanced habit formation in Gilles de la Tourette syndrome. *Brain*.
- Dickinson, A., 1985. Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society B: Biological Sciences* 308 (1135), 67-78.
- Fairburn, C.G., Beglin, S., 2008. Eating disorder examination questionnaire (EDE-Q 6.0), in: Fairburn, C. (Ed.), *Cognitive behaviour therapy and eating disorders*. Guildford Press, New York, pp. 309-313.
- Fairburn, C.G., Cooper, Z., O'Connor, M., 2008a. Eating Disorder Examination (16.0D). in: C.G., F. (Ed.), *Cognitive Behavior Therapy and Eating Disorders*. Guildford Press, New York.
- Fairburn, C.G., Cooper, Z., O'Connor, M., 2008b. Eating Disorder Examination (Edition 16.0D), *Cognitive Behaviour Therapy and Eating Disorders*. Guildford Press, New York.
- Fairburn, C.G., Cooper, Z., Shafran, R., 2003. Cognitive behaviour therapy for eating disorders: a "transdiagnostic" theory and treatment. *Behaviour research and therapy* 41 (5), 509-528.
- Favaro, A., Zanetti, T., Tenconi, E., Degortes, D., Ronzan, A., Veronese, A., Santonastaso, P., 2005. The relationship between temperament and impulsive behaviors in eating disordered subjects. *Eat Disord* 13 (1), 85-92.
- Fineberg, N.A., Potenza, M.N., Chamberlain, S.R., Berlin, H.A., Menzies, L., Bechara, A., Sahakian, B.J., Robbins, T.W., Bullmore, E.T., Hollander, E., 2010. Probing compulsive and impulsive behaviors, from animal models to endophenotypes: a narrative review. *Neuropsychopharmacology* 35 (3), 591-604.
- Fladung, A.K., Gron, G., Grammer, K., Herrnberger, B., Schilly, E., Grasteit, S., Wolf, R.C., Walter, H., von Wietersheim, J., 2010. A neural signature of anorexia nervosa in the ventral striatal reward system. *Am J Psychiatry* 167 (2), 206-212.
- Fladung, A.K., Schulze, U.M.E., Schöll, F., Bauer, K., Grön, G., 2013. Role of the ventral striatum in developing anorexia nervosa. *Translational psychiatry* 3 (10), e315.
- Foa, E.B., Huppert, J.D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., Salkovskis, P.M., 2002. The Obsessive-Compulsive Inventory: development and validation of a short version. *Psychol Assess* 14 (4), 485-496.
- Foerde, K., Steinglass, J.E., Shohamy, D., Walsh, B.T., 2015. Neural mechanisms supporting maladaptive food choices in anorexia nervosa. *Nat Neurosci* advance online publication.
- Fulton, S., Richard, D., Woodside, B., Shizgal, P., 2004. Food restriction and leptin impact brain reward circuitry in lean and obese Zucker rats. *Behavioural brain research* 155 (2), 319-329.
- Giel, K.E., Kullmann, S., Preißl, H., Bischoff, S.C., Thiel, A., Schmidt, U., Zipfel, S., Teufel, M., 2013. Understanding the reward system functioning in anorexia nervosa: Crucial role of physical activity. *Biological psychology* 94 (3), 575-581.
- Gillan, C.M., Apergis-Schoute, A.M., Morein-Zamir, S., Urcelay, G.P., Sule, A., Fineberg, N.A., Sahakian, B.J., Robbins, T.W., 2015. Functional Neuroimaging of Avoidance Habits in Obsessive-Compulsive Disorder. *American Journal of Psychiatry* 172 (3), 284-293.
- Gillan, C.M., Morein-Zamir, S., Urcelay, G.P., Sule, A., Voon, V., Apergis-Schoute, A.M., Fineberg, N.A., Sahakian, B.J., Robbins, T.W., 2013. Enhanced Avoidance Habits in Obsessive-Compulsive Disorder. *Biol Psychiatry* 3223 (13), 145-145.

- Gillan, C.M., Papmeyer, M., Morein-Zamir, S., Sahakian, B.J., Fineberg, N.A., Robbins, T.W., de Wit, S., 2011. Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *American Journal of Psychiatry* 168 (7), 718-726.
- Godier, L.R., Park, R.J., 2014a. Compulsivity in anorexia nervosa: a transdiagnostic concept. *Frontiers in Psychology* 5, 778.
- Godier, L.R., Park, R.J., 2014b. A novel measure of compulsive food restriction in anorexia nervosa: Validation of the Self-Starvation Scale (SS). *Eating Behaviors* 17c, 10-13.
- Godier, L.R., Park, R.J., 2015. Does compulsive behaviour in Anorexia Nervosa resemble an addiction? A qualitative investigation. *Frontiers in Psychology* 6.
- Graybiel, A.M., 2008. Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience* 31, 359-387.
- Jarosz, A.F., Wiley, J., 2014. What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. *The Journal of Problem Solving* 7 (1).
- Jeffreys, S.H., 1961. *Theory of Probability* (3rd Edition). Oxford University Press, Oxford, UK.
- Kalivas, P.W., Volkow, N.D., 2005. The neural basis of addiction: a pathology of motivation and choice. *American Journal of Psychiatry* 162 (8), 1403-1413.
- Kaye, W.H., Fudge, J.L., Paulus, M., 2009. New insights into symptoms and neurocircuit function of anorexia nervosa. *Nature Reviews Neuroscience* 10 (8), 573-584.
- Kaye, W.H., Wierenga, C.E., Bailer, U.F., Simmons, A.N., Wagner, A., Bischoff-Grethe, A., 2013. Does a Shared Neurobiology for Foods and Drugs of Abuse Contribute to Extremes of Food Ingestion in Anorexia and Bulimia Nervosa? *Biol Psychiatry* 73 (9), 836-842.
- Keys, A., Brozek, J., Henschel, A., Mickelsen, O., Taylor, H.L., 1950. *The biology of human starvation*. University of Minnesota Press, Oxford.
- Lawrence, N.S., Wooderson, S., Mataix-Cols, D., David, R., Speckens, A., Phillips, M.L., 2006. Decision making and set shifting impairments are associated with distinct symptom dimensions in obsessive-compulsive disorder. *Neuropsychology* 20 (4), 409-419.
- Love, J., Selker, R., Marsman, M., Jamil, T., Verhagen, A.J., Ly, A., Gronau, Q.F., Smira, M., Epskamp, S., Matzke, D., Wild, A., Rouder, J.N., Morey, R.D., Wagenmakers, E.-J., 2015. JASP (Version 0.6.6).
- Nelson, H.E., 1982. *National Adult Reading Test (NART): Test Manual*. NFER-NELSON, Windsor.
- Park, R.J., Godier, L.R., Cowdrey, F.A., 2014. Hungry for reward: How can neuroscience inform the development of treatment for Anorexia Nervosa? *Behavior and Research Therapy*.
- Patton, J.H., Stanford, M.S., Barratt, E.S., 1995. Factor structure of the Barratt Impulsiveness Scale. *J Clin Psychol* 51 (6), 768-774.
- Pinto, A., Steinglass, J.E., Greene, A.L., Weber, E.U., Simpson, H.B., 2014. Capacity to delay reward differentiates obsessive-compulsive disorder and obsessive-compulsive personality disorder. *Biol Psychiatry* 75 (8), 653-659.
- Robbins, T.W., Gillan, C.M., Smith, D.G., de Wit, S., Ersche, K.D., 2012. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends Cogn Sci* 16 (1), 81-91.
- Rosval, L., Steiger, H., Bruce, K., Israel, M., Richardson, J., Aubut, M., 2006. Impulsivity in women with eating disorders: problem of response inhibition, planning, or attention? *International Journal of Eating Disorders* 39 (7), 590-593.

- Scheurink, A.J., Boersma, G.J., Nergardh, R., Sodersten, P., 2010. Neurobiology of hyperactivity and reward: agreeable restlessness in anorexia nervosa. *Physiology & Behaviour* 100 (5), 490-495.
- Sjoerds, Z., de Wit, S., van den Brink, W., Robbins, T.W., Beekman, A.T.F., Penninx, B.W.J.H., Veltman, D.J., 2013. Behavioral and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients. *Translational Psychiatry* 3 (12), e337.
- Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R., Jacobs, G.A., 1983a. Manual for the State-Trait Anxiety Inventory. . Consulting Psychologists Press, Palo Alto, CA.
- Spielberger, C.D., Gorsuch, R.L., Vagg, P.R., Jacobs, G.A., 1983b. Manual for the State-Trait Anxiety Inventory. Consulting Psychologists Press, Palo Alto, CA.
- Spitzer, R.L., Kroenke, K., Williams, J.B.W., Lowe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine* 166 (10), 1092.
- Spitzer, R.L., Williams, J.B., Gibbons, M., First, M.B., 2004. Structured clinical interview for the DSM-IV (SCID-I/P). American Psychiatric Press, Arlington, VA.
- Steinglass, J.E., Walsh, T., 2006. Habit learning and anorexia nervosa: A cognitive neuroscience hypothesis. *International Journal of Eating Disorders* 39 (4), 267-275.
- Tchanturia, K., Anderluh, M., Morris, R., Rabe-Hesketh, S., Collier, D., Sanchez, P., Treasure, J., 2004. Cognitive flexibility in anorexia nervosa and bulimia nervosa. *Journal of the International Neuropsychological Society* 10 (4), 513-520.
- Tchanturia, K., Liao, P.C., Uher, R., Lawrence, N., Treasure, J., Campbell, I.C., 2007. An investigation of decision making in anorexia nervosa using the Iowa Gambling Task and skin conductance measurements. *Journal of the International Neuropsychological Society* 13 (4), 635-641.
- Verdejo-Garcia, A., Benbrook, A., Funderburk, F., David, P., Cadet, J.L., Bolla, K.I., 2007. The differential relationship between cocaine use and marijuana use on decision-making performance over repeat testing with the Iowa Gambling Task. *Drug Alcohol Depend* 90 (1), 2-11.
- Voon, V., Derbyshire, K., Rück, C., Irvine, M., Worbe, Y., Enander, J., Schreiber, L., Gillan, C., Fineberg, N., Sahakian, B., 2014. Disorders of compulsivity: a common bias towards learning habits. *Molecular psychiatry*.
- Wagner, A., Aizenstein, H., Mazurkewicz, L., Fudge, J., Frank, G.K., Putnam, K., Bailer, U.F., Fischer, L., Kaye, W.H., 2008. Altered insula response to taste stimuli in individuals recovered from restricting-type anorexia nervosa. *Neuropsychopharmacology* 33 (3), 513-523.
- Walsh, B.T., 2013. The Enigmatic Persistence of Anorexia Nervosa. *American Journal of Psychiatry* 170 (5), 477-484.
- Waxman, S.E., 2009. A Systematic Review of Impulsivity in Eating Disorders. *European Eating Disorders Review* 17 (6), 408-425.
- Worbe, Y., Savulich, G., de Wit, S., Fernandez-Egea, E., Robbins, T.W., 2015. Tryptophan Depletion Promotes Habitual over Goal-Directed Control of Appetitive Responding in Humans. *International Journal of Neuropsychopharmacology*.
- Zink, C.F., Weinberger, D.R., 2010. Cracking the moody brain: The rewards of self starvation. *Nature Medicine* 16 (12), 1382-1383.

**Table 1.** Demographic and Clinical Characteristics of Study 1

	AN (n=23)		HC (n=18)		Significance
					( <i>p</i> two -tailed)
	Mean	SD	Mean	SD	
<b>Age</b>	25.65	6.40	31.78	7.45	0.009
<b>BMI</b>	16.79	1.89	24.42	2.78	<0.001
<b>Years of Education</b>	14.83	2.57	16.44	2.09	0.030
<b>Duration of illness (yrs)</b>	11.28	6.93	---	---	---
<b>EDE</b>	2.61	1.05	---	---	---
<b>EDE-Q</b>	3.78	1.15	---	---	---
<b>BIS</b>	59.64	13.60	48.61	8.82	0.005
<b>BDI</b>	28.74	12.6	---	---	---
<b>STAI-State</b>	51.35	11.51	25.56	9.04	<0.001
<b>STAI-Trait</b>	62.77	9.19	27.72	8.13	<0.001

Note: BMI = Body Mass Index; EDE= Eating Disorder Examination; EDE-Q= Eating Disorder Examination Questionnaire; BIS = Barratt Impulsiveness Scale; BDI = Beck Depression Inventory; STAI = State Trait Anxiety Inventory

\*BMI's were only available for 11 out of 18 HCs.

“---“ indicates that data are not available

**Table 2. Demographics and Clinical Characteristics of Study 2**

	AN-C (n=13)		AN-R (n=14)		HC (n=17)		Significance ( <i>p</i> two-tailed)	Pairwise post-hoc group comparisons
	Mean	SD	Mean	SD	Mean	SD		
<b>Age</b>	31.15	7.96	27.07	6.49	24.06	5.60	0.011	AN-C>HC
<b>BMI</b>	15.79	1.91	20.94	1.63	21.24	1.90	<0.001	HC, AN-R>AN-C
<b>Years of Education</b>	18.17	4.26	18.64	2.66	16.81	2.23	0.286	--
<b>Duration of illness (yrs)</b>	10.25	5.2	5.75	4.15	---	---	0.022	AN-C>AN-R
<b>EDE</b>	3.10	1.44	0.74	0.61	0.22	0.20	<0.001	AN-C>AN-R, HC
<b>EDE-Q</b>	3.36	1.74	0.95	0.85	0.46	0.32	<0.001	AN-C>AN-R, HC
<b>BIS</b>	56.46	12.53	61.57	9.72	54.94	11.74	0.264	--
<b>BDI</b>	29.62	18.12	6.36	6.43	3.06	4.41	<0.001	AN-C>AN-R, HC
<b>STAI-State</b>	50.08	13.04	34.43	5.45	26.53	6.56	<0.001	AN-C>AN->HC
<b>STAI-Trait</b>	60.69	12.91	45.21	10.56	32.59	9.51	<0.001	AN-C>AN-R> HC
<b>NART</b>	113.84	7.54	114.57	6.77	109.29	7.43	0.100	--
<b>CIA</b>	29.46	12.43	7.93	6.5	1.76	2.25	<0.001	AN-C>AN-R, HC
<b>OCI-r</b>	20.85	15.78	10.79	6.45	6.06	6.94	0.006	AN-C> HC
<b>YBC-EDS – SRQ Current</b>	16.38	8.21	3.79	3.95	0	0	<0.001	AN-C>AN-R, HC
<b>YBC-EDS- SRQ Past</b>	24.62	5.30	23.82	5.30	0	0	<0.001	AN-C, AN-R> HC

Note: BMI = Body Mass Index; EDE= Eating Disorder Examination; EDE-Q= Eating Disorder Examination Questionnaire; BIS = Barratt Impulsiveness Scale; BDI = Beck Depression Inventory; STAI = State Trait Anxiety Inventory; NART = National Adult

Reading Test; CIA = Clinical Impairment Assessment; OCI-r = Obsessive Compulsive Inventory Revised; YBC-EDS-SRQ = Yale-Brown-Cornell Eating Disorder Scale Self-Report Questionnaire. -- indicates no significant difference between groups

**Table 3.  $BF_{01}$  values for the Devalued and Difference Scores in Study 1 and Study 2.**

	$BF_{01}$	
	Devalued	Difference
<b>Study 1</b> – Slips of Action Paradigm	2.998	3.036
<b>Study 2</b> – Avoidance Paradigm	5.674	0.742
<b>Study 2</b> – Slips of Action Paradigm	3.630	4.284

Note: BF = Bayes Factor

## **Figure Legends**

### **Figure 1. The four stages of the Slips of Action Paradigm**

### **Figure 2. The Noise Avoidance Task**

### **Figure 3. Slips-of-Action and Baseline Test Study 1**

The left panel shows responses for valuable versus devalued outcomes during the slips-of-action test; the right panel responses for valuable versus devalued stimuli during the baseline test.

### **Figure 4. Slips-of-Action and Baseline Test Study 2**

The left panel shows responses for valuable versus devalued outcomes during the slips-of-action test; the right panel responses for valuable versus devalued stimuli during the baseline test.

### **Figure 5. Noise Avoidance Habit Test Study 2**

Responses to valued and devalued stimuli following overtraining across experimental groups