Altered Implicit Category Learning in Anorexia Nervosa

Megan E. Shott
University of Colorado Denver

J. Vincent Filoteo
University of California San Diego

Leah M. Jappe
University of Minnesota

Tamara Pryor
Eating Disorder Center of Denver, Denver, Colorado

W. Todd Maddox
University of Texas

Michael D. H. Rollin, Jennifer O. Hagman, and
Guido K. W. Frank
University of Colorado Denver

Objective: Recent research has identified specific cognitive deficits in patients with anorexia nervosa (AN), including impairment in executive functioning and attention. Another such cognitive process, implicit category learning has been less studied in AN. This study examined whether implicit category learning is impaired in AN. Method: Twenty-one women diagnosed with AN and 19 control women (CW) were administered an implicit category learning task in which they were asked to categorize simple perceptual stimuli (Gabor patches) into one of two categories. Category membership was based on a linear integration (i.e., an implicit task) of two stimulus dimensions (orientation and spatial frequency of the stimulus). Results: AN individuals were less accurate on implicit category learning relative to age-matched CW. Model-based analyses indicated that, even when AN individuals used the appropriate (i.e., implicit) strategy they were still impaired relative to CW who also used the same strategy. In addition, task performance in AN patients was worse the higher they were in self-reported novelty seeking and the lower they were in sensitivity to punishment. Conclusions: These results indicate that AN patients have implicit category learning deficits, and given this type of learning is thought to be mediated by striatal dopamine pathways, AN patients may have deficits in these neural systems. The finding of significant correlations with novelty seeking and sensitivity to punishment suggests that feedback sensitivity is related to implicit learning in AN.

Keywords: implicit, learning, reward, anorexia nervosa, temperament

Anorexia nervosa (AN) is a severe psychiatric illness characterized by restrictive eating, fears of being fat despite life threatening low weight, and can be associated with binge eating and/or purging behavior (APA, 2000). Common temperament traits of eating disorder patients include high harm avoidance, cognitive rigidity, and alterations in feedback sensitivity (Anderluh, Tchan-turia, Rabe-Hesketh, & Treasure, 2003; Bulik et al., 2003; Jappe et al., 2011; Wagner et al., 2006). AN has the highest mortality among all psychiatric disorders (Sullivan, 1995), and as such, understanding this disease from multiple perspectives is an important undertaking.

In recent years, there has been growing interest in studying cognitive functioning in patients with AN. These studies are important for a better understanding of AN given the possibility that cognitive deficits may (a) contribute to the development and persistence of AN, (b) result from neurological changes associated with the disease, or (c) influence the choice of treatment approaches. Many of these studies have found deficits primarily in executive functioning (Fassino et al., 2002; Holliday, Tchanturia, Landau, Collier, & Treasure, 2005; Jones, Duncan, Brouwers, & Mirsky, 1991; Lauer, Gorzewski, Gerlinghoff, Backmumd, & Zihl, 1999; Roberts, Tchanturia, Stahl, Southgate, & Treasure, 2007; Steinglass, Walsh, & Stern, 2006; Tchanturia, Whitney, & Treasure, 2006), visual-spatial abilities (Grunwald et al., 2001; Jones et al., 1991; Kingston, Szmukler, Andrewes, Tress, & Desmond, 1996; Palazidou, Robinson, & Lishman, 1990; Szmukler et al., 1992), and attention (Bosanac et al., 2007; Johansson, Ghaderi, & Andersson, 2005; Olhrann et al., 2004; Seed, Dixon, McCluskey, & Young, 2000; Seed, McCue, Wesnes, Dahabra, & Young, 2002; Shafran, Lee, Cooper, Palmer, & Fairburn, 2007). Although past
studies have also identified deficits in AN patients in general intellectual functioning (Chui et al., 2008; Mathias & Kent, 1998), a recent systematic review and meta-analysis found that people with AN tend to have higher IQ scores compared with healthy comparison groups (Lopez, Stahl, & Tchanturia, 2010). These results suggest that difficulties in cognitive functioning in AN may be more specific than an impairment in general intellectual functioning, and raises the possibility that AN might impact only certain neural systems. In a recent review, Lena et al. (Lena, Fiocco, & Leyenaar, 2004) stressed the need for greater exploration of these neurocognitive deficits in AN suggesting that impaired cognition may in fact contribute to the genesis of the disease.

Although there have been more recent studies examining cognition in AN, there are other potentially important processes that have received less attention, such as those associated with learning and memory (Chui et al., 2008; Jones et al., 1991; Mathias & Kent, 1998). This area of research could have important implications for a better understanding as to whether or how faulty learning and memory mechanisms contribute to the development or maintenance of AN. While a handful of studies have examined explicit memory performance in AN, few have studied implicit memory in these patients. Implicit memory denotes a class of learning and memory processes that operate outside conscious awareness and includes processes such as priming, classical conditioning, and procedural learning. The few studies that have studied implicit learning in AN have examined whether patients with an eating disorder are abnormally primed to process food- or body-related information (Hermans, Pieters, & Eelen, 1998; Johansson, Ghaderi, Hallgren, & Andersson, 2008). The results of those studies have been somewhat inconsistent. Whereas Hermans et al. (1998) found no implicit memory bias in their word stem completion task, Johansson et al. (2008) found that AN individuals exhibited attentional interference for food-related words. Other forms of implicit learning that have been examined include measures of procedural learning, which is a process primarily involved in the acquisition of skills and habits. Lawrence et al. (Lawrence et al., 2003), for example, found that AN patients were impaired relative to controls on a measure of visual discrimination learning (which is believed to be a measure of procedural learning). These same patients were not impaired on a pattern recognition task, which is a measure of explicit memory. These authors argued that the AN patients’ deficits on the visual discrimination learning task were likely related to alterations in the subcortical dopamine reward-based learning system, given that alterations in this system in primates and humans diminishes the learning of this task. In contrast, the normal performance of the AN patients on the pattern recognition task was thought to be related to intact functioning in the medial temporal lobe memory system, which is thought to mediate explicit learning and memory. The suggestion that changes in the dopamine reward-based learning system resulted in AN patients’ deficits on this task is consistent with a growing literature suggesting alterations in the dopamine system in these patients (Fladung et al.; Frieling et al., 2010; Kaye, Ebert, Gwirtsman, & Weiss, 1984).

Additional evidence that AN patients may have alterations in reward-based mechanisms comes from studies using self-reports, such as the Sensitivity to Punishment-Sensitivity to Reward Questionnaire (SPSRQ), which measures the saliency of rewarding or punishing real-world reinforcers (e.g., social or monetary reinforc-ers). Using the SPSRQ, Jappe et al. (2011) found that AN patients were more sensitive to both punishment and reward than controls, indicating alterations in feedback sensitivity in AN patients. Despite these interesting findings, the relationship between the AN patients’ altered feedback sensitivity and behavioral measures of reinforcement-based learning have yet to be examined. The potential link between learning and feedback sensitivity in AN is important in this context, since traits such as sensitivity to reward or punishment affect learning (Cavanagh, Frank, & Allen, 2011) and could be related to the ability to change behavior and thus the course of the illness. Taken together, the study of implicit learning, and how deficits in this area relate to self-report AN temperament or behaviors, may help identify specific brain alterations in AN, which could have important implications in our understanding of what AN individuals learn to find rewarding or reinforcing (Wächter, Lungu, Liu, Willingham, & Ashe, 2009).

The purpose of the present study was to examine implicit learning in women with AN. To this end, AN patients and control women (CW) were administered a category learning task that emphasized the procedural-based learning system that relies on dopamine reward-based learning (Ashby, Alfonso-Reese, Turken, & Waldron, 1998). The task is described in detail below, but briefly, participants are presented with simple perceptual stimuli, categorize the stimuli into one of two categories, and are given corrective feedback immediately following their response. Importantly, this task is believed to be mediated by the dopamine reward-based learning system since fast, corrective feedback stimulates dopamine activation and learning (Schultz, 2002), while delaying such feedback results in diminished learning (Maddox, Ashby, & Bohil, 2003; Maddox et al., 2005). Other evidence that dopamine signaling might play an important role in learning the procedural-based category learning task comes from studies showing that patients with known damage to the dopamine systems (i.e., Parkinson’s disease) are impaired on this task (Filoteo, Maddox, Salmon, & Song, 2005) as well as functional neuroimaging studies with normal participants demonstrating activation of the posterior caudate (Nomura et al., 2007), a brain region believed to make use of dopamine signaling during learning (Miyachi, Hikosaka, Miyashita, Karadi, & Rand, 1997; Nakamura & Hikosaka, 2006).

Given the previously identified alterations in dopamine systems in AN and the proposed involvement of such systems in implicit category learning, we predicted that AN patients would be significantly impaired in implicit learning relative to controls as demonstrated by lower accuracy on the category learning task. To assess whether implicit category learning was associated with temperament and behavior variables, we also examined the relationship between task performance and self-report measures of feedback sensitivity such as reward and punishment. We predicted that poorer performance in category learning would be associated with altered levels of feedback sensitivity in the AN group.

Method

Participants

Twenty-one female patients with AN and 19 control women (CW), all Caucasian, participated in the study. Ten of the AN patients were of the binge/purge subtype, while the remaining 11 AN patients were restricting subtype. Table 1 shows the mean age
and body mass index (BMI) of the AN patients and CW participants, and the lowest BMI, duration of illness, and age of disease onset for the AN patients. The mean age of the AN group and the CW group did not differ (p = .28). As expected, the mean BMI for the AN patients was significantly lower than for CW participants (p < .001). A χ² test revealed that there were no educational level differences between the AN and CW groups (p = .27). Patients were recruited from the Partial Hospitalization Program at the Eating Disorder Center of Denver (a facility that treats adult eating disorder patients, n = 17) or the inpatient eating disorders unit at Children’s Hospital Colorado (a facility that treats children and adolescents with eating disorders through the age of 21, n = 4). While the mean age and duration of illness differed in the two treatment settings, AN patients had similar current or lowest BMI, and age of disease onset regardless of the type of treatment facility. Each patient completed the study within a few weeks of starting treatment and the subject’s height and weight were taken the day of testing. Patients completed the Structural Clinical Interview for DSM–IV Disorders (SCID-IV) with a trained doctoral level clinician to ensure they met criteria for AN. Control participants also completed the SCID to screen for abnormal eating behaviors or other DSM–IV disorders. All CW participants had a BMI between 19 and 25 and had no personal or family history of mental or neurological disorders. This study was approved by the Colorado Multiple Institutional Review Board and participants gave written informed consent.

Clinical Measures

The Beck Depression Inventory (BDI-2; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) is a 21-item self-report that measures depression symptom severity. The Eating Disorder Inventory 2 (EDI-2; Garner, 1991) contains 91 items organized onto 12 primary scales. Of the 12 scales, four were used: Drive for Thinness, Body Dissatisfaction, Bulimia, and Perfectionism. The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ; Torrubia, Avila, Molto, & Caseras, 2001) is a 48-item self assessment that is based on Reinforcement Sensitivity Theory (RST; Gray, 1970). The scale provides separate scores for sensitivity to punishment (SP) and sensitivity to reward (SR).

Stimuli and Stimulus Generation

The specific version of the category learning task used in this study required participants to learn to categorize simple perceptual stimuli (Gabor patches; see Figure 1) into one of two categories (Category A or Category B) when the rule that dictated category membership was nonverbal (implicit learning) and required a linear integration of two stimulus dimensions (orientation and spatial frequency of the gratings of the Gabor patch; see below for details). The study used the randomization technique introduced by Ashby and Gott (Ashby & Gott, 1988). Stimuli consisted of Gabor patches that varied in spatial frequency and orientation (see Figure 1). For each experiment, an equal number of Category A and Category B stimuli were generated by sampling randomly from two bivariate normal distributions. Each category distribution is specified by a mean and a variance on each dimension, and by a covariance between dimensions. For both category structures it was always the case that the covariance matrix for Category A was identical to the covariance matrix for Category B. The categories differed only in the location of their means. The exact parameter values for each experiment are listed in Table 2 (Note: these are in arbitrary units that were transformed into the stimulus units using the methods described below). The category structure for the experiment is displayed in Figures 2. Each filled square in these figures denotes the spatial frequency and spatial orientation of a Gabor pattern from Category A, while each unfilled circle denotes the spatial frequency and spatial orientation of a Gabor pattern from Category B. The solid line in these figures denotes the location of the optimal decision bound. The use of the optimal bound maximizes long-run accuracy.

The Temperament and Character Inventory (Cloninger, Svrakic, & Przybeck, 1993) is a 240-item inventory consisting of 7 independent dimensions. We included two of these (Novelty Seeking and Harm Avoidance), which are variables previously implicated in AN (Wagner et al., 2006).

Figure 1. Computer generated Gabor patch.
The stimuli were computer generated and displayed on a 21-inch monitor with 1360 x 1024 resolution. In each of the three experiments, the stimuli consisted of a single Gabor patch (see Figure 1). Each of the stimuli varied in orientation and spatial frequency. Each Gabor patch was generated using MATLAB routines from Brainard’s (1997) Psychophysics Toolbox. Each random sample \((x_f, x_o)\) was converted to a stimulus by deriving the frequency, \(f = 0.025 \times (x_f/5000)\) cycles per pixel, and orientation, \(o = 0.36x_o\) degrees. The scaling factors were chosen in an attempt to equate the salience of frequency and orientation based on our past experience with these stimuli. Each Gabor patch was 7 cm in diameter, which subtended a visual angle of about 8.8 degrees from a viewing distance of 45 cm.

**Experimental Procedure**

Three hundred and twenty trials were presented and were broken down into four blocks of 80 trials. At the start of the experiment, the participants were told that they were involved in a study that examined their ability to categorize simple stimuli, that a series of stimuli would be presented, and that they would be asked to categorize each as a member of either Category A or Category B. They were also told that at the beginning of the experiment they may feel as though they were guessing, but as the experiment progressed, their accuracy would likely increase. Participants indicated their categorization responses by pressing one key for Category A stimuli and another key for Category B stimuli. For each trial, the stimulus was presented until the participant’s categorization response was made, then immediately following their response, they were given feedback for 1 second that consisted of the word “wrong” if their response was incorrect or “correct” if their response was correct. Once feedback was given, the next trial was initiated 1 second later. Total task duration was between 20 and 25 minutes, depending on individual subject response time.

**Model-Based Analyses**

A major advantage of the task used in this study is that mathematical models have been developed to determine the specific approach used by participants when learning the task (Maddox & Ashby, 2004; Maddox et al., 2003; Zeithamova & Maddox, 2006). These models enable a more detailed investigation of procedural learning in that previous studies have shown that patient groups can differ from controls in implicit learning not because of an impairment in such processes, but instead because they adopted a different strategy than controls when learning (Shohamy, Myers, Onlaor, & Gluck, 2004). Thus, through the use of these models, we are able to identify AN patients who adopted a procedural-based approach to learning compared to control individuals to determine if in fact the patients are impaired in procedural-based learning. Although the main dependent measure in this study was the accuracy of participants’ responses, we also applied a series of models to the data to better understand the category learning processes used by the AN patients and CW participants. These models have been developed to analyze data in this paradigm and have been used extensively to better understand the underlying processes in category learning in both normal and patient populations (Ashby & Maddox, 1992; Ashby, 1999; Filoteo, Maddox, & Davis, 2001; Maddox, 2001; Maddox & Ashby, 1993; Maddox & Ashby, 1996; Maddox & Filoteo, 2001 Maddox, Filoteo, Delis, & Salmon, 1996; Maddox, Filoteo, & Huntington, 1998). Two classes of models were applied to the final block of data separately for each participant (the details of these models can be found in (Maddox & Ashby, 1993; Maddox, 1999). One class of models assumes that the participant used a procedural-based (PB) approach by performing an implicit integration of the spatial frequency and orientation of the stimuli. The optimal PB model assumes that the participant used the rule depicted in Figure 1 as the solid line. This model has one free parameter representing the rule application variability (i.e., perceptual and criterial noise or \(\sigma_f^2\)). The second procedural-based model was the general linear classifier (GLC), which also assumes that the participant’s decision on each trial is based on a linear integration of information from both the spatial frequency and orientation dimensions, although the weighting given to the two dimensions may be unequal. This model has three parameters, including the slope and intercept of the linear bound, and the rule application variability (i.e., \(\sigma_o^2\)).

The second class of models was the hypothesis testing (HT) models, which assumed that the participant used a verbalizable rule when performing the task. Two conjunctive rule models were applied to the data, both of which assume that the participant based her categorization decision on both dimensions. Both models as-

---

**Table 2**

**Category Distribution Parameter Values for the Implicit Category Learning Task**

<table>
<thead>
<tr>
<th>Category</th>
<th>(\mu_f)</th>
<th>(\mu_o)</th>
<th>(\sigma_f^2)</th>
<th>(\sigma_o^2)</th>
<th>(\text{cov}_{f,o})</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>260</td>
<td>125</td>
<td>75</td>
<td>9000</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>339</td>
<td>125</td>
<td>75</td>
<td>9000</td>
<td>0</td>
</tr>
</tbody>
</table>

---

**Figure 2.** Category structure for the experiment.
sumed that the participant set a criterion along the spatial frequency dimension that distinguished low from high frequency stimuli, and set a criterion along the orientation dimension that distinguished shallow from steep angled stimuli. Thus the model assumed that there were four response regions: low frequency/shallow angle, low frequency/steep angle, high frequency/shallow angle and high frequency/steep angle. The criterion along the spatial frequency dimension was a free parameter and the criterion along the orientation dimension was a free parameter. One model assumed that low frequency/steep angle stimuli were in category A and all other stimuli were in category B. The other model assumed that high frequency/shallow angle stimuli were in category B and all other stimuli were in category A. Both models also included a rule application variability parameter (i.e., \(\sigma^2\)), for a total of three parameters in each model. Two additional hypothesis-testing models that were applied were unidimensional models, one that assumes the participant categorized the stimuli based only on the spatial frequency dimension and another that assumes the participant categorized the stimuli based only on the orientation dimension. These two models had two free parameters, a decision criterion on one of the dimensions and \(\sigma^2\). A final model that was applied assumes that the participant responded randomly (RR). Akaike’s (Akaike, 1974) information criterion (AIC) was used to determine the model that provided the best account of the data.

Statistical Analyses

Demographic and behavioral results across groups were analyzed using SPSS 18.0 statistical software package using two-tailed independent \(t\) tests. Task performance was examined by contrasting participants’ accuracy (percent correct), response time (reaction time (RT)), and RT variability (mean standard deviation for each block) across the entire 320 trials in 80 trial blocks using a 2 (group: AN vs. CW) \(\times\) 4 (blocks 1–4) mixed-design ANOVAs.

Results

Clinical Scales

The mean subscale and total scores for AN and CW on the BDI, TCI, EDI-2, and the SPSRQ are provided in Table 3. AN patients scored higher on the BDI than CW patients (\(p < .001\)). For the TCI, AN patients scored higher than the CW on the Harm Avoidance scales (\(p < .05\)). For the EDI-2, the AN group had significantly greater scores than the CW group on all subscales (all \(p < .01\)). For the SPSRQ, the AN group had greater scores on both the SP and SR subscales (both \(p < .05\)). There were no significant differences in scores on any clinical measures between restricting-type and binge/purging-type anorexia.

Task Accuracy Results

Accuracy results for the category learning task are shown in Figure 3. An ANOVA identified a main effect of group, \(F(1, 38) = 4.5, p < .05, \eta^2_p = .11\), with AN participants performing worse than the CW participants, and a main effect of block, \(F(3, 114) = 10.6, p < .001, \eta^2_p = .27\), with both AN and CW participants’ performances improving across the trials. There was a trend for a Group X block interaction, \(F(3, 114) = 2.4, p = .07, \eta^2_p = .08\). Despite the lack of a significant Group X block interaction, follow-up \(t\) tests demonstrated that the two groups did not differ in either blocks 1 or 2 (both \(p > .17\)), but the two groups did differ in blocks 3 and 4 (both \(p < .05\); \(\eta^2_p = .14\) for block 3, and \(\eta^2_p = .16\) for block 4). There was no difference in task performance in the subgroups of AN patients.

Task RT Results

In addition to accuracy, we also examined RT results in the CW and AN groups to help determine if other factors (e.g., attention and/or fatigue) might account for the accuracy differences described above. It was anticipated that if these other factors could account for the accuracy differences, then the AN group would likely have greater and more variable RTs than CW, particularly in the latter two blocks of trials where the accuracy levels differed. For mean RTs, the ANOVA did not identify a main effect of group, \(F(1, 38) = 1.1, p > .05\), but there was a main effect of block, \(F(1, 38) = 11.5, p < .01, \eta^2_p = .24\), with both AN and CW participants’ RTs becoming faster across the blocks. Importantly, there was no Group X block interaction, \(F(3, 114) = 0.8, p > .05\), indicating that RTs did not differ between the groups across blocks. As for the RT standard deviations, there were no main effects of group, \(F(1, 38) = 1.9, p > .05\), or block, \(F(3, 114) = 0.58, p > .05\), and no Group X block interaction, \(F(3, 114) = 0.63, p > .05\).

Task Performance Correlations With Clinical Data

To examine the relationship between performances on the category learning task and the clinical measures, a learning slope was computed by taking the final block accuracy and subtracting the first block accuracy for each participant. This learning slope was then correlated with the BDI and various subscales from the TCI and EDI-2 using Pearson correlations. The correlation coefficients for the TCI and EDI-2 are displayed in Table 4. For the TCI, the learning slope in the AN patients correlated significantly with the

| Table 3
<p>| Behavioral Data as Measured Through the Beck Depression Inventory (BDI), Temperament and Character Inventory (TCI), Eating Disorders Inventory 2 (EDI-2) Raw scores, and the Sensitivity to Punishment and Sensitivity to Reward (SPSRQ) |</p>
<table>
<thead>
<tr>
<th>CW</th>
<th>AN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BDI</strong></td>
<td>1.42</td>
</tr>
<tr>
<td><strong>TCI</strong></td>
<td>20.1</td>
</tr>
<tr>
<td><strong>Novelty Seeking</strong></td>
<td>16.5</td>
</tr>
<tr>
<td><strong>Harm Avoidance</strong></td>
<td>8.5</td>
</tr>
<tr>
<td><strong>EDA-2</strong></td>
<td>20.8</td>
</tr>
<tr>
<td><strong>Drive for Thinness</strong></td>
<td>0.9</td>
</tr>
<tr>
<td><strong>Bulimia</strong></td>
<td>13.5</td>
</tr>
<tr>
<td><strong>Body Dissatisfaction</strong></td>
<td>2.2</td>
</tr>
<tr>
<td><strong>Perfectionism</strong></td>
<td>14.0</td>
</tr>
<tr>
<td><strong>Novelty Seeking</strong></td>
<td>5.3</td>
</tr>
<tr>
<td><strong>Sensitivity to Punishment</strong></td>
<td>13.2</td>
</tr>
<tr>
<td><strong>Sensitivity to Reward</strong></td>
<td>8.2</td>
</tr>
<tr>
<td><strong>SPSRQ</strong></td>
<td>11.3</td>
</tr>
<tr>
<td><strong>Perfectionism</strong></td>
<td>5.4</td>
</tr>
<tr>
<td><strong>ES</strong></td>
<td>13.1</td>
</tr>
</tbody>
</table>
Novelty Seeking subscale, such that patients with lower learning slopes (less learning) scored higher on this scale. For the SPSRQ, there was a significant positive correlation between learning slope and the SP subscale \((p < .05)\) indicating that those AN patients with lower learning slopes were less sensitive to punishment.

Learning slopes were not associated with the BDI, the included subscales of the EDI-2, lowest BMI, duration of illness, or age of disease onset. There were no significant correlations for the CW group between learning slope and any of the clinical measures.

Model Results

To determine whether groups differed in how they solved the categorization task in the last block of trials, we examined the number of participants in each group whose data were best fit by one of the PB model (i.e., the optimal model or the general linear classifier model) versus the number of participants whose data were best fit by one of the HT models described above (i.e., the conjunctive or unidimensional models). If one of the PB models best accounted for a participant’s data, it would indicate that they based their decision on an implicit integration of the two stimulus dimensions, whereas if one of the HT model best accounted for a participant’s data, it would indicate that this individual based their decision on a verbalizable rule. Finally, if the RR model best accounted for a participant’s data it would indicate that they were responding randomly.

The results of these analyses for the final block indicated that, for the CW group, nine data sets were best fit by a PB model, nine were best fit by an HT model, and one was best fit by the RR model. For the AN group, nine data sets were best fit by a PB model, seven were best fit by an HT model, and five was best fit by the RR model. These frequencies did not differ based on a \(\chi^2\) test \((p = .22)\). Thus, although there were differences in accuracy rates between the CW and AN groups, the two groups did not differ in terms of the frequency of participants who used an implicit (PB), verbalizable (HT), or random responding approach.

To further determine whether the model-based subgroups differed, we contrasted accuracy rates in the final block for the CW and AN participants who used either a PB or HT approach. Final block accuracy rates were used because these were the data on which the model analyses were conducted. These accuracy rates are shown in Figure 4. \(T\) tests demonstrated that the CW participants whose data were best fit by a PB model performed more accurately than those AN participants whose data were best fit by a PB model \((p < .05)\). In contrast, accuracy in the final block for the CW and AN participants whose data were best fit by an HT model did not differ significantly \((p = .75)\). Thus, accuracy rates differed between CW and AN participants who used an implicit approach to learning the task, whereas the groups did not differ in those participants who used a verbalizable approach. Within group comparisons between participants who used a PB or HT approach indicated that the AN PB users tended to perform worse than AN HT users (see Figure 4), although this difference was only a trend.

Table 4

<table>
<thead>
<tr>
<th>Clinical Measure</th>
<th>CW</th>
<th>AN</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDI</td>
<td>-.14</td>
<td>.07</td>
</tr>
<tr>
<td>TCI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novelty Seeking</td>
<td>-.34</td>
<td>-.47</td>
</tr>
<tr>
<td>Harm Avoidance</td>
<td>.32</td>
<td>.08</td>
</tr>
<tr>
<td>EDI-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drive for Thinness</td>
<td>-.16</td>
<td>.08</td>
</tr>
<tr>
<td>Bulimia</td>
<td>.12</td>
<td>.11</td>
</tr>
<tr>
<td>Body Dissatisfaction</td>
<td>-.42</td>
<td>-.07</td>
</tr>
<tr>
<td>Perfectionism</td>
<td>.37</td>
<td>.11</td>
</tr>
<tr>
<td>SPSRQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity to Punishment</td>
<td>.04</td>
<td>.44*</td>
</tr>
<tr>
<td>Sensitivity to Reward</td>
<td>-.06</td>
<td>.20</td>
</tr>
</tbody>
</table>

Figure 3. Implicit learning task accuracy results \((p < .05)\).

Figure 4. Accuracy results for the Hypothesis Tested (HT) versus Procedural Based (PB) method of learning \((p < .05)\).
whereas there was no difference between PB and HT users in the CW group ($p = .36$).

Finally, to determine if the AN participants whose final block data were best fit by a PB or HT model differed clinically, we contrasted these subgroups on the clinical variables and measures. The only difference that emerged was that the AN participants whose data was best fit by a PB model had larger scores on the Perfectionism subscale of the EDI-2 as compared to those participants whose data was best fit by an HT model ($13.3 \text{ vs. } 7.6; p < .01$).

**Discussion**

To our knowledge, this is the first study to examine implicit category learning in patients with AN. The study has several key findings. First, individuals with AN performed worse on the implicit learning task than normal weight, age-matched controls, thus providing evidence for a deficit in implicit learning in ill AN. Second, as indicated by our model-based analyses, AN participants who used the task appropriate strategy still performed significantly worse than CW participants, indicating that AN women who learned the appropriate task approach were still impaired in learning, and poor strategy selection, per se, could not account for the learning deficit. Third, certain temperament and personality characteristics were associated with impaired category learning, such as novelty seeking and reinforcement biases (sensitivity to punishment).

Previous work has strongly indicated a role of dopamine in reward-based learning (Schultz, 1998, 2002) and recent research suggests that AN may be associated with alterations in the dopamine system (Fladung et al., 2010; Frieling et al., 2010; Kaye et al., 1984). The finding that AN patients are impaired in implicit category learning, which is believed to rely heavily on reward-based learning, provides support for a possible deficiency in dopamine functioning in AN patients, and is also in line with other studies that have suggested that certain forms of learning in AN are associated with deficits in the reward-based learning system (Lawrence et al., 2003). This possibility is also consistent with previous studies that other patient groups who have involvement of the dopamine system (such as patients with Parkinson’s disease) are also impaired in implicit category learning (Buytenhuijs et al., 1994; Euteneuer et al., 2009; Filoteo, Maddox, Salmon, & Song, 2007; Price, 2005).

Although it is reasonable to speculate that dopamine abnormalities account for AN patients’ deficits in implicit category learning, other possibilities must be considered. Previous studies have consistently identified deficits in AN patients in two other cognitive domains, cognitive set shifting and visual central coherence (Lopez, Tchanturia, Stahl, & Treasure, 2008; Lopez, Tchanturia, Stahl, Booth, Holliday, & Treasure 2007; Roberts, Tchanturia, & Treasure, 2010; Southgate, Tchanturia, & Treasure, 2008; Tchanturia et al., 2004; Tchanturia et al., 2004; Tchanturia, Morris, Surguladze, & Treasure, 2002; Tenco et al., 2010; Zastrow et al., 2009). Cognitive set shifting refers to those processes needed to switch from one cognitive set to another during the course of a task. Visual central coherence, on the other hand, refers to the ability to integrate and process an entire visual scene, as opposed to focusing on the component processes of the visual scene. Deficits in either of these two cognitive processes might also account for AN patients’ deficits in implicit category learning. Specifically, in the initial stages of performing an implicit category learning task, participants are thought to first attempt verbalizable rules in which the rule is often based on a single dimension, and only later do participants shift to a more accurate implicit rule (Ashby et al., 1998). In this study, if AN patients are impaired in cognitive set shifting, they might tend to continue to attempt to implement a verbal rule rather than switching and adopting a more accurate implicit rule, which would lead to a learning deficit. However, a similar proportion of study participants in each group adopted the implicit approach, suggesting that the AN group did not differentially stay with a rule based approach. The second possible explanation of the present findings may be that AN patients failed to see the gestalt of the stimuli because of a deficit in visual central coherence, and as such, they were unable to correctly learn to categorize the stimuli because they could not integrate the two stimulus dimensions. This might be consistent with previous findings that AN individuals had altered strategies and abilities to copy complex figures. Such a deficit might result in the AN patients focusing on only one dimension of the stimulus despite the fact that optimal responding required learning to categorize the stimuli based on both dimensions. Although both are plausible alternative explanations, it should be pointed out that even AN patients who adopted an implicit rule (as demonstrated by the model-based analyses) were impaired in learning the task relative to the controls. Those AN patients who used such an approach were able to shift cognitive set and integrate the stimulus dimensions, but nevertheless demonstrated impaired learning. Thus, although deficits in either cognitive set shifting or visual central coherence might account for the AN patients’ deficits, it appears that a deficit in reward-based learning (possibly secondary to dopamine deficiencies) might best account for our findings.

Other findings from the current study supporting the notion that a deficit in reward-based learning accounts for AN patients’ deficits was the association between poor performance in implicit category learning and personality assessments. In particular, impaired implicit category learning was associated with greater scores on the Novelty Seeking TCI subscale. Greater scores on Novelty Seeking are thought to reflect an individual’s propensity to be more exploratory, impulsive and disorderly. Another finding was that poor category learning in AN patients was associated with less sensitivity to punishment based on the SPSRQ. Taken together, these results indicate that those individuals with AN who dislike rules and are less sensitive to punishment will be more likely to be impaired in implicit category learning.

The notion that AN patients are impaired in reward-based processes is not entirely new and has been the focus of several recent neurobiological and behavioral reviews (Harrison, O’Brien, Lopez, & Treasure, 2010; Kaye, Fudge, & Paulus, 2009; Keating, 2010; Scheurink, Boersma, Nergardh, & Sodersten, 2010). Claes and colleagues examined sensitivity to punishment as measured by the Behavioral Inhibition System subscale (Carver & White, 1994) and found that restricting subtype AN were more sensitive to punishment than purging subtype AN and healthy controls (Claes, Nederkoorn, Vandereycken, Guerrieri, & Vertommen, 2006). A recent study from our group using the SPSRQ found that restricting as well as binge eating/purging type AN had increased sensitivity to punishment and reward compared to age-matched healthy controls (Jappe et al., 2011). The present study also found similar
increased sensitivity to punishment and reward in AN as measured by the SPSRQ, but interestingly category learning performance was only associated with alterations in punishment sensitivity. AN individuals have traditionally been associated with perfectionism and high motivation to perform well (Bachner-Melman et al., 2007; Castro-Fornieles et al., 2007; Kaye, 2008; Maia et al., 2009; Nilsson, Sundborn, Sundborn, & Hagglöf, 2008; Peck & Lightsey, 2008; Wade et al., 2008) and it could be possible that fear of punishment or failure have the benefit of higher effort and better performance. Still, biologic abnormalities in AN might prevent that group from performing similar to controls in implicit learning and other cognitive tasks. Interestingly, only the AN with highest perfectionism scores used the correct, implicit strategy. This could indicate that only those subjects who put forth the most effort actually ended up using the procedural based learning approach, while this did not seem to be a necessary factor for the controls. Further study is needed to examine this finding given that it indicates that there are some personality traits that might in fact positively influence learning in AN.

While the nature of this relationship should be explored further, it is important to point out that this present study is one of the first to link alterations in feedback sensitivity to impaired implicit learning processes in AN.

There are several limitations to the present study that should be addressed. First the sample sizes and cultural homogeneity of our participants limit the power of our statistical analyses and the generalizability of our findings to the general AN population. Second, we provide no direct link between the proposed alteration in the dopamine system in AN patients and their observed impairment on the implicit category learning task. Future research will have to address this hypothesized relationship. Third, at present the clinical implications of our findings is unknown, but we hope that our results provide the impetus for exploring such cognitive and temperament associations and how they impact clinically relevant aspects of the disease (e.g., prevention and treatment). Fourth, we did not have a direct measure of attention and/or fatigue, factors that could have impacted performance and account for the accuracy differences. However, the AN and CW groups did not differ in mean or variability in RT, suggesting that attention or fatigue may not have accounted for the AN learning deficit. Finally, the present study did not characterize the observed category learning deficit relative to other cognitive factors that might have influenced performance, such as general intelligence, so the specificity of this deficit is not known. It will be important for future studies to examine learning deficits in the context of other, potential cognitive impairment in AN individuals.

References


Journal of the International Neuropsychological Society, 7, 710–727. doi:10.1017/S1355617701766076


Received April 19, 2011
Revision received October 5, 2011
Accepted November 21, 2011