

Deficient Goal-Directed Control in a Population Characterized by Extreme Goal Pursuit

Karin Foerde^{1,2}, Nathaniel D. Daw³, Teresa Rufin¹, B. Timothy Walsh^{1,2},
Daphna Shohamy⁴, and Joanna E. Steinglass^{1,2}

Abstract

■ Research in computational psychiatry has sought to understand the basis of compulsive behavior by relating it to basic psychological and neural mechanisms: specifically, goal-directed versus habitual control. These psychological categories have been further identified with formal computational algorithms, model-based and model-free learning, which helps to provide quantitative tools to distinguish them. Computational psychiatry may be particularly useful for examining phenomena in individuals with anorexia nervosa (AN), whose self-starvation appears both excessively goal directed and habitual. However, these laboratory-based studies have not aimed to examine complex behavior, as seen outside the laboratory, in contexts that extend beyond monetary rewards. We therefore assessed (1) whether

behavior in AN was characterized by enhanced or diminished model-based behavior, (2) the domain specificity of any abnormalities by comparing learning in a food-specific (i.e., illness-relevant) context as well as in a monetary context, and (3) whether impairments were secondary to starvation by comparing learning before and after initial treatment. Across all conditions, individuals with AN, relative to healthy controls, showed an impairment in model-based, but not model-free, learning, suggesting a general and persistent contribution of habitual over goal-directed control, across domains and time points. Thus, eating behavior in individuals with AN that appears very goal-directed may be under more habitual than goal-directed control, and this is not remediated by achieving weight restoration. ■

INTRODUCTION

The emerging understanding of the brain's systems for habitual and goal-directed control has offered insight into potential mechanisms of compulsive behaviors (Everitt & Robbins, 2016). Such behaviors occur across a range of psychiatric illnesses, including anxiety, obsessive compulsive disorder (OCD), substance use, and eating disorders. The seemingly compulsive nature of anorexia nervosa (AN) has long been noted, but the clinical and neuropsychological phenomena in AN challenge a simple goal-directed-versus-habitual dichotomy. Whereas the pursuit of thinness appears remarkably goal driven, patterns of eating appear rigidly unchangeable even with treatment (Mayer, Schebendach, Bodell, Shingleton, & Walsh, 2012; Schebendach, Mayer, Devlin, Attia, & Walsh, 2012). Such assumptions about the underlying mechanisms can be tested experimentally aided by advances in cognitive and computational neuroscience (Huys, Maia, & Frank, 2016). In this study, we apply methods for probing habitual and goal-directed behavior in relation to AN.

A dual-system view of automatic or habitual versus controlled or goal-directed behavior is a long-standing organizing principle in psychology and neuroscience (James,

1890). Research shows that these behaviors can be distinguished behaviorally and neurally (Graybiel, 2008; Yin & Knowlton, 2006; Yin, Knowlton, & Balleine, 2004, 2005; Dickinson & Balleine, 2002), and it has been argued that they arise from distinct computational mechanisms for evaluating actions, known as model-based and model-free learning (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Daw, Niv, & Dayan, 2005). In the laboratory, these approaches can be distinguished using instrumental learning tasks that include outcome devaluation procedures and two-step Markov decision tasks, and they have been associated with partially distinct neural substrates (Lee, Shimojo, & O'Doherty, 2014; Daw et al., 2011; Gläscher, Daw, Dayan, & O'Doherty, 2010; Tricomi, Balleine, & O'Doherty, 2009; Valentin, Dickinson, & O'Doherty, 2007; Yin, Knowlton, & Balleine, 2006; Yin et al., 2004; Coutureau & Killcross, 2003; Killcross & Coutureau, 2003). Experimental evidence shows that healthy people generally use a mix of both model-free and model-based approaches rather than one or the other (Weissengruber, Lee, O'Doherty, & Ruff, 2019; Otto, Gershman, Markman, & Daw, 2013; Otto, Raio, Chiang, Phelps, & Daw, 2013; Daw et al., 2011). Compulsive behaviors have been posited to relate to an imbalance in habitual versus goal-directed control (Gillan, Robbins, Sahakian, van den Heuvel, & van Wingen, 2016), either through excessive habitual drive of behavior or through deficient goal-directed control. Thus far, reductions in model-based learning have been

¹New York State Psychiatric Institute, ²Columbia University Irving Medical Center, ³Princeton University, ⁴Kavli Institute for Brain Science

demonstrated in individuals with compulsive behaviors (Wyckmans et al., 2019; Gillan, Kosinski, Whelan, Phelps, & Daw, 2016; Gillan, Apergis-Schoute, et al., 2015; Voon et al., 2015). Such an imbalance in control has not been assessed specifically in AN. Here, we probe these mechanisms to assess their contribution to illness in AN and further address the domain specificity of any deficits and the effect of acute treatment.

First, categorizing the real-world behaviors at the core of psychiatric illnesses is complex, as many include behaviors that can be considered compulsive but are not easily captured by the simple, repetitive, stimulus-evoked habits described in the experimental literature. For example, engagement in a ritualized behavior to reduce anxiety symptoms in OCD may be habitual or automated yet could also be construed as a successful, goal-directed strategy for managing anxiety because of obsessions (Salkovskis, 1985). Relying on self-report or clinical judgment may lead to mislabeling a behavior as goal directed or habitual and misunderstanding the relevant neurocomputational processes. AN provides a particularly intriguing example of this conundrum. Maintenance of significantly low body weight is a defining feature of AN (American Psychiatric Association, 2013). This self-starvation appears to be an unrelenting goal pursuit focused on weight loss (Bruch, 1979), and neuropsychological data support the characterization of patients with AN as engaging excessive control (Lloyd, Yiend, Schmidt, & Tchanturia, 2014). Intertemporal choice studies show that individuals with AN are more patient than healthy comparison participants, consistent with heightened self-control, and that this behavior is associated with abnormal neural activity in the striatum (Steinglass et al., 2012, 2017; Decker, Figner, & Steinglass, 2015). Yet, the restrictive eating that characterizes AN also shares many features of habit (Foerde, Steinglass, Shohamy, & Walsh, 2015; Walsh, 2013): It is learned (not innate) and is inflexibly triggered by certain cues, and individuals with AN are unable to readily change this behavior—even when seeking treatment, suggesting that it may have become outcome independent over time. Although these candidate mechanisms are not exhaustive, they represent prominent views of AN. Here, we assess model-based and model-free behavior in a learning task as one way to test the idea that AN is characterized by abnormalities in goal-directed and/or habitual mechanisms. Although this task provides measures of the strength of both model-based and model-free learning (hypothetically underlying goal-directed and habitual behavior, respectively), previous studies have consistently reported decreased model-based learning (but no changes in the model-free measure) in compulsive disorders and other situations where habits would be expected to dominate (Wyckmans et al., 2019; Gillan, Kosinski, et al., 2016; Gillan, Otto, Phelps, & Daw, 2015; Otto, Raio, et al., 2013). Thus, following this previous work, we primarily focus on the strength of model-based learning in the task as a measure of goal-habit balance.

Second, goal-directed and habitual behaviors are domain-general mechanisms, yet different psychiatric illnesses involve quite specific compulsive behaviors (e.g., specific compulsions in OCD, abuse of specific substances, gambling, or excessive weight loss). Most laboratory studies have used generic outcomes, such as money (Gillan, Kosinski, et al., 2016; Voon et al., 2015), and an important interpretational question left open is whether apparent deficits in goal-directed control in psychiatric patients are because of impairments in a domain-general mechanism or, possibly, to a reorienting of goal-directed behavior toward the object of compulsion at the expense of other domains. The latter view would predict that substance abusers, for instance, might be relatively unmotivated by money and their impairment in goal-directed control could be mitigated or even reversed if their drug of abuse was at stake. Among individuals with AN, the question is whether the influence of model-based control may diverge in the setting of food outcomes compared to monetary outcomes.

Third, model-based deficits in populations with psychopathology indicate a correlation between the two and do not inform whether such deficits are a causal factor in maladaptive compulsions. Indeed, decreased goal-directed control may be a result of compulsive behaviors, as has been argued, for instance, for neurological effects of drugs of abuse (Volkow et al., 2010). Overall, little is known about the longitudinal progression of these deficits. One question is whether deficits in model-based control are remediated by treatment. A recent study suggested that, in OCD, impairments in model-based behavior are not remediated by cognitive behavioral therapy (Wheaton, Gillan, & Simpson, 2019). For AN particularly, the effects of starvation on cognition might be substantial. By studying individuals with AN during acute illness and again after extensive inpatient weight-restoration treatment, the persistence of any learning differences and their relation to changes in psychopathology can be assessed.

In this study, we addressed these three questions with a population of individuals with AN, using a two-step Markov decision task (Figure 1) designed, and widely used in psychiatric populations, to assess the extent of model-based and model-free learning. By examining an illness with complex and seemingly highly controlled (yet maladaptive) behaviors, we aimed to test whether these reflect heightened or deficient model-based behavior (Question 1). We also compared results from food and monetary versions of the task to examine the specificity of findings in the eating disorder domain and tested for the presence of model-based versus model-free behavior in the context of outcomes directly relevant to underlying psychopathology (Question 2). Whereas the monetary task assumes that money is rewarding, in the food task, participants were rewarded with points that they could use to select a food they preferred (e.g., a low-fat food for an individual with AN), thereby allowing the food task to be motivationally relevant to both controls and patients. Finally, we evaluated

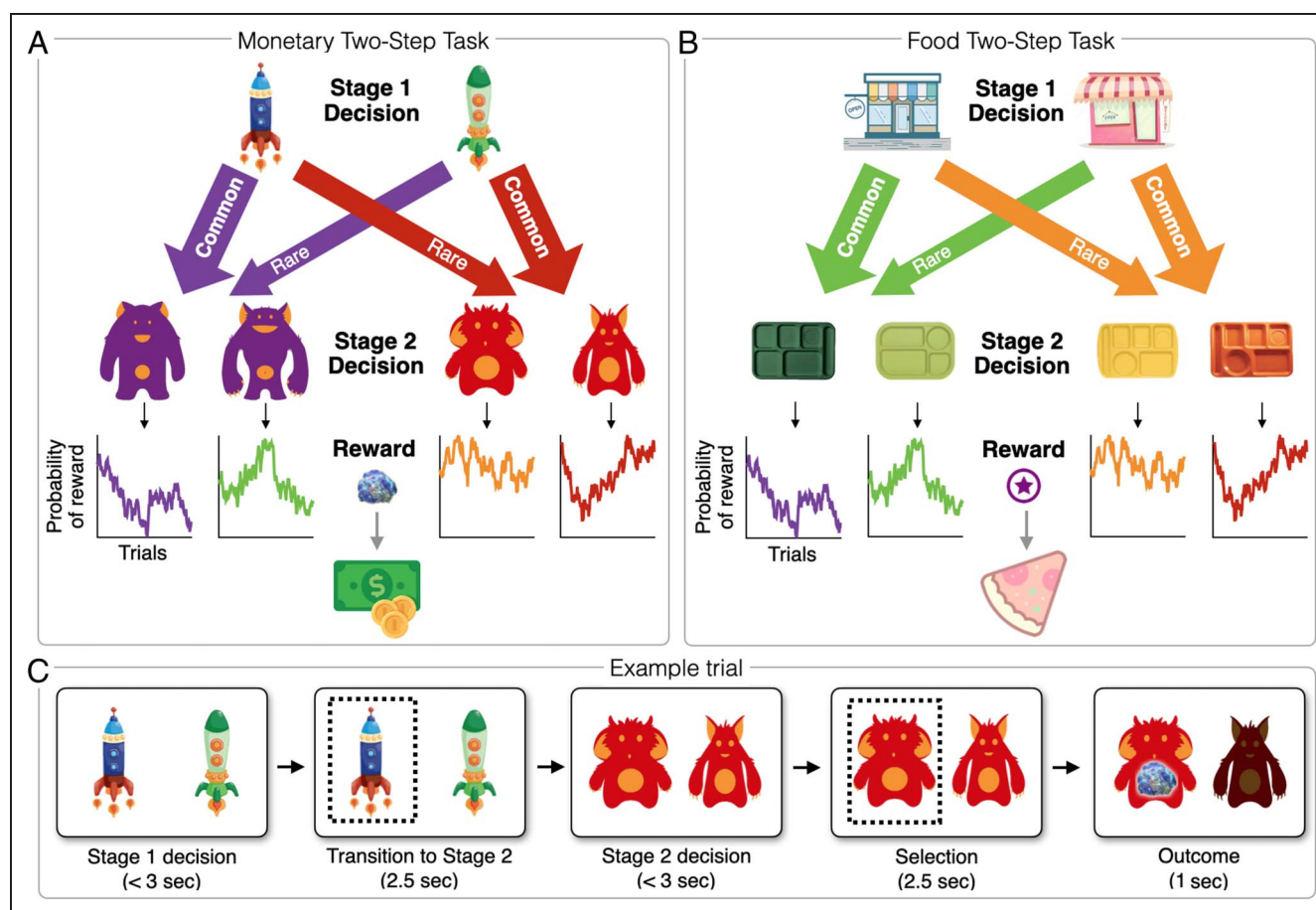


Figure 1. Two-step decision tasks used to assess model-free and model-based learning. (A) Decision task with monetary outcomes. Alien treasure pieces were converted into a monetary bonus paid out at the end of the task to all participants. (B) Task with food outcomes. Food tokens were converted into access to a range of preferred food items, and the selected food was consumed as a snack after the task. (A, B) In both tasks, the Stage 1 choice determined the transition to the next stage according to a fixed probability. One choice was associated with transition to one particular Stage 2 state 70% of the time (Common transition) and the other state 30% of the time (Rare transition). At Stage 2, participants made choices followed by reward or no reward (both monetary and food outcomes were actualized after the task). Each Stage 2 option was associated with a probabilistic reward, which ranged from 0.25 to 0.75 and varied gradually (according to a Gaussian random walk) and independently across trials (see examples in bottom rows of A and B). (C) Example trial steps from the monetary task.

patients before and after inpatient weight-restoration treatment, to begin to study the progression of model-based and model-free processes (Question 3).

METHODS

Participants

Participants were 41 women with AN and 53 healthy comparison women (or healthy controls [HCs]). Individuals were eligible if they were between the ages of 16 and 46 years (Table 1), with an estimated IQ > 80 (measured by Wechsler Abbreviated Scale of Intelligence, Second Edition [WASI-II; Wechsler, 1999]). Eligible patients met *Diagnostic and Statistical Manual of Mental Disorders-5* (DSM-5) (American Psychiatric Association, 2013) criteria for AN—restricting (AN-R, $n = 19$) or binge eating/purging (AN-BP, $n = 22$) subtype—and were receiving inpatient treatment at the New York State Psychiatric Institute

(NYSPI) specialized Eating Disorders Unit. Patients with AN were not eligible if they had a history of a psychotic disorder, were at an imminent risk of suicide, or met criteria for substance use disorder. Anxiety and depressive disorders were not exclusionary, as these commonly co-occur with AN (Hudson, Hiripi, Pope, & Kessler, 2007). Three individuals with AN were taking antidepressant medications (selective serotonin reuptake inhibitors). Treatment at NYSPI is provided at no cost for those interested in and eligible for participation in research, and those with AN were not additionally compensated for their time. HCs were recruited through the community and were compensated \$50 for their time. HCs were group-matched for age and ethnicity and were included if they had no current or past psychiatric illness, including any history of an eating disorder, and had a body mass index (BMI) in the normal range (18–25 kg/m²). The HC group included 19 individuals who endorsed a history of dieting behavior. This study was approved by the NYSPI institutional review board; after

Table 1. Demographic and Clinical Information

	HC (<i>n</i> = 53)	AN (<i>n</i> = 41)	HC vs. AN		AN T1 vs. T2	
	Mean ± SD	Mean ± SD	<i>t</i>	<i>p</i>	<i>t</i>	<i>p</i>
<i>Time 1</i>						
Age, years	25.6 ± 5	27.1 ± 7	1.2	.236		
Estimated IQ	111.8 ± 12.2	105.1 ± 11.3	−2.7	.008		
Dur. ill (years)		9.6 ± 7.1				
BMI	21.3 ± 1.5	16 ± 2	−14.6	<.001		
EDE-Q	0.5 ± 0.6	4.16 ± 1.5	15.8	<.001		
BDI	2.1 ± 2.4	27.6 ± 12.6	14.4	<.001		
STAI	31.3 ± 7.3	61 ± 11.2	15.5	<.001		
YBC-EDS	n/a	22.5 ± 6.4				
YBC-EDS pre.	n/a	11.7 ± 3.2				
YBC-EDS rit.	n/a	10.8 ± 3.7				
<i>Time 2</i>						
	HC (<i>n</i> = 29 ^a)	AN (<i>n</i> = 25 ^a)				
BMI	20.9 ± 1.3	20.3 ± 1	−1.4	.172	−13.0	<.001
EDE-Q	0.4 ± 0.4	2.8 ± 1.2	9.8	<.001	6.7	<.001
BDI	3.0 ± 2.6	16.7 ± 12.1	5.7	<.001	5.2	<.001
STAI	33.6 ± 8.7	55.9 ± 11.4	7.9	<.001	3.4	.003

WASI IQ data are missing from one individual with AN. At Time 1 (T1), BDI data are missing from one individual with AN. At Time 2 (T2), BMI data are missing from one individual with AN. Dur. ill = duration of illness; pre. = preoccupations; rit. = rituals.

^a Thirty-four HCs and 32 individuals with AN participated in the longitudinal study.

complete description of the study to the participants, adult participants provided written informed consent and adolescents provided written assent with parental consent.

Procedure

Psychiatric diagnoses were established using the Structured Clinical Interview for DSM-IV (First, Spitzer, Gibbon, & Williams, 2002) and the Eating Disorders Assessment for DSM-5 (Sysko et al., 2012). Height and weight were obtained on a wall stadiometer and a beam balance scale, respectively. Estimated IQ was assessed with the WASI-II (Wechsler, 1999). Severity of eating disorder psychopathology was measured by the Eating Disorder Examination Questionnaire (EDE-Q; Fairburn, 2008; Fairburn & Beglin, 1994), a 36-item self-report assessment of eating disorder symptoms that has established community norms for adolescents and adults. In addition, participants with AN completed the Yale–Brown–Cornell Eating Disorder Scale (YBC-EDS; Mazure, Halmi, Sunday, Romano, & Einhorn, 1994), an interview measure of eating disorder symptoms with separate subscales related to Preoccupations and Rituals adapted from the Yale–Brown–Cornell Obsessive–Compulsive Scale (Goodman et al., 1989); this is a standard

measure of obsessions and compulsions in eating disorders. Symptoms of anxiety were assessed with the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970); and depression, with the Beck Depression Inventory (BDI; Beck & Steer, 1993). Higher scores indicate greater symptom severity on each measure.

Participants were enrolled in either a longitudinal study that included two assessment time points or a study with a single assessment (Time 1). For all individuals with AN, Time 1 assessments occurred within 1 week of hospital admission; and Time 2 assessments, after weight restoration treatment to at least 90% ideal body weight (Metropolitan Life Insurance, 1959), corresponding to a BMI of approximately 19.5–20.0 kg/m². For HC individuals, Time 1 and Time 2 assessments occurred at an interval group matched to AN ($M_{HC} = 51 \pm 23$ days, $M_{AN} = 59 \pm 22$ days), $t(45) = 1.17$, $p = .25$. Thirty-two individuals with AN participated in the longitudinal study—18 completed both time points, 7 completed only Time 1, and 7 completed only Time 2—and nine individuals participated in the single-assessment study. Food outcome task data at Time 2 of one participant with AN were lost because of computer malfunction. Thirty-four HC individuals participated in the longitudinal study—29 completed both time

points, and 5 completed only Time 1—and 19 individuals participated in the single-assessment study. The dropout rate was 21.8% for the AN group and 14.7% for the HC group.

Two-Step Decision Task

Each participant completed two versions of the two-step decision task (Figure 1; Decker, Otto, Daw, & Hartley, 2016; Sharp, Foerde, Daw, & Shohamy, 2016; Daw et al., 2011)—one that involved playing for money and one that involved playing for a food snack to be consumed after the task (order counterbalanced across participants). In the monetary version of the task, participants collected pieces of “space treasure,” each worth \$0.10, with the total earned paid in cash at the end of the experiment for both HC participants and participants with AN. In the food version of the task, participants collected tokens that were converted into access to food items to be consumed as a snack. Before completing the task, participants were given a “menu” (see Appendix) of 15 food items to rank in order of preference; they were not instructed as to how the ratings would be used in conjunction with task performance to determine later snack options. A greater number of tokens earned on the task translated into the ability to choose among more desirable food items as a snack. For example, earning 110 tokens allowed access to all but the most preferred food item, and earning 100 tokens allowed access to food items ranked 7–15. Upon task completion, participants selected among the food items they had earned access to. In this way, all participants played to gain their most preferred food outcome. This approach holds constant that participants are playing to obtain preferred foods and accounts for the likely inherent group difference in HC and AN food preferences (i.e., most in the AN group prefer low-fat food items). HC participants were given their snack to consume after the task. Patients with AN were given their selected food as an evening snack on the Eating Disorders Unit. The consumption of the evening snack counted toward privileges on the unit, as part of standard behavioral treatment.

A subset of participants (HC: $n = 15$, AN: $n = 17$) completed a version of the food decision task in which the outcome was shown as preferred and nonpreferred food items, rather than as tokens. The specific food items were individually determined per participant: Before the decision task, participants completed a rating task in which nine food items were rated on a scale from 1 to 9 (1 = *highly preferred*). Participants were not instructed that their ratings would relate to their later snack options. The food image participants received on most trials determined which snack they would receive after the task. In this task, participants had 2 sec to make choices in both stages. The pattern of behavior—decreased model-based learning in the AN group—was the same in the two variations of the food decision task. The monetary task was the same for all participants.

The food and monetary tasks were structurally identical but included distinct cover stories and task environments

that matched the outcomes used (see Figure 1A and B). For both tasks, each trial proceeded in two stages (Figure 1C). In the first stage (Stage 1), participants chose between two spaceships (or cafés), revealing a second-stage (Stage 2) choice between two aliens (or food trays). Each second-stage alien (or food tray) had a slowly changing chance of delivering space treasure (or a food token) versus nothing, necessitating continuous learning by trial and error. The four Stage 2 options were determined by independently drifting Gaussian random walks with $SD = 0.025$ and bounded by .25 and .75 probability of reward, such that the reward probability associated with each Stage 2 option changed slowly from trial to trial (Figure 1A and B, bottom). The response window at each stage was 3 sec. Participants completed 200 trials of each task.

A key design feature of the task was the probabilistic association between first- and second-stage choices: Choosing the blue spaceship (or blue café) led to the purple planet (or green kitchen) 70% of the time, that is, a “common” transition, and the red planet (or yellow kitchen) 30% of the time, that is, a “rare” transition. The contingencies were reversed for the green spaceship (or pink café). The transition structure between stages allows dissociation of model-free versus model-based learning strategies: Model-free learning is ignorant of transition structure and favors repeating a first-step choice that ultimately results in reward, even if it does so via a low-probability transition. By contrast, model-based learning is sensitive to the transition contingencies and uses them to infer the first-stage choice most likely to lead to the preferred second-stage environment.

After completing the two-step tasks, participants’ knowledge of the transition structure between Stage 1 and Stage 2 (e.g., “If you picked the blue spaceship, which planet would you most likely land on?”) and their estimates of the transition probabilities were assessed (e.g., “If you picked the blue spaceship, how likely would you be to see the purple planet?”).

Data Analysis

Demographic and clinical variables were compared using independent t tests. Measures of psychopathology at Time 1 and Time 2 were compared within AN using paired t tests.

Assessment of Model-based and Model-free Learning (Computational Model)

To capture the influence of incremental learning across many trials, we fit participants’ choices using a reinforcement learning model in which choices are modeled as arising because of the weighted combination of model-free and model-based reinforcement learning. The model is based on the “hybrid model” originally applied to healthy participants with this task (Daw et al., 2011), but incorporating a set of modifications that have been used

to improve the robustness of parameter estimation and the characterization of population-level parameter estimates in later studies of individual or group differences. Specifically, to eliminate unnecessary free parameters, we (1) use a single learning rate for both levels (i.e., take $\alpha_1 = \alpha_2$; these are typically similar when estimated separately) and (2) set the eligibility trace parameter for model-free learning, $\lambda = 1$ (this is typically near 1 when estimated freely, and constraining it improves the robustness of estimating the model-free effect by making it log-convex conditional only on the learning rate). Next, to improve group-level estimation and the transparency of the model, we (3) use an algebraically equivalent change of variables where β_{MB} equals $w\beta_{stage1}$ in the Daw et al. (2011) parameterization and β_{MF} equals $(1 - w)\beta_{stage1}$ (this makes their estimation log-convex conditional on the learning rate and reexpresses them in the more interpretable form of logistic regression coefficients); (4) further change variables to rescale all three β parameters by α (this reduces the β s' collinearity with the learning rate, which also makes them more comparable and their group-level distributions less dispersed across participants); and (5) decay the Q values of unchosen options by multiplying them by $(1 - \alpha)$ after each trial (this typically improves fit and also makes the model limit to a one-trial back logistic regression analysis, discussed later, as $\alpha \rightarrow 1$). All these modifications except for (2) were used by Otto, Gershman, et al. (2013); Gillan, Kosinski, et al. (2016); and Vikbladh et al. (2019) and also provided the best test–retest reliability relative to two earlier model variants in Brown, Chen, Gillan, and Price's (2020) analysis; we added the additional simplification (2, which was not considered by Brown et al.) in our previous collaborative patient study (Sharp et al., 2016).

On each trial, t , participants make a Stage 1 choice $c_{1,t}$, leading to a transition to a Stage 2 state s_t where another choice, $c_{2,t}$, is made, followed by reward r_t . At Stage 2, it is assumed that participants learn a value function over states and choices, $Q^{stage2}(s, c)$, whose value for the chosen action is updated based on the reward received at each trial according to a delta rule, $Q_{t+1}^{stage2}(s_t, c_{2,t}) = (1 - \alpha)Q_t^{stage2}(s_t, c_{2,t}) + r_t$. Here, α is a free learning rate parameter. (In this, and analogous update equations, a factor of α is omitted from the last term of the update, equivalent to rescaling the rewards and Q s by $\frac{1}{\alpha}$ and the corresponding weighting parameters β by α [Otto, Raio, et al., 2013]). The probability of a particular Stage 2 choice is modeled as governed by these values according to a logistic softmax, with free inverse temperature parameter β_{stage2} : $P(c_{2,t} = c) \propto \exp(\beta_{stage2}Q_t^{stage2}(s_t, c))$, normalized over both options c .

Stage 1 choices are modeled as determined by the weighted combination of both model-free and model-based value predictions about the ultimate, Stage 2 value of each Stage 1 choice. Model-based values Q^{MB} are given by the learned values of the corresponding Stage 2 state, maximized over the two actions: $Q^{MB}(c) = \max_{c_2} Q_t^{stage2}(s, c_2)$, where s is the Stage 2 state predominantly

produced by Stage 1 choice c . Model-free values are learned by TD(1), where $Q_{t+1}^{MF}(c_1, t) = (1 - \alpha)Q_t^{MF} + r_t$. The Stage 1 choice probabilities are given by a logistic softmax, with a contribution from each value estimate weighted by its own free inverse temperature parameter: $P(c_{1,t} = c) \propto \exp(\beta_{MB}Q_t^{MB}(c) + \beta_{MF}Q_t^{MF}(c) + \beta_{stick}I(c = c_{1,t-1}))$. Here, $I(c = c_{1,t-1})$ is a binary indicator of whether a choice repeats the previous trial's choice, so the weight β_{stick} measures the general tendency to persevere or switch regardless of feedback.

At the end of each trial, all value estimates Q for unchosen actions and unvisited states are decayed by multiplying by $(1 - \alpha)$.

The model has five free parameters: four weights β (β_{stage2} , β_{MB} , β_{MF} , and β_{stick}) and a learning rate α . Our main measures of interest are β_{MB} and β_{MF} , measuring the contribution of model-based and model-free learning.

The free parameters of the model were estimated by maximizing the likelihood of each participant's sequence of choices, using a distinct set of parameters for each game (i.e., each combination of participant, session, and task type: up to four games per participant, two task types at two time points). These were estimated jointly with group-level distributions over the entire population using an expectation maximization procedure (Huys et al., 2011) implemented in the Julia language (Bezanson, Karpinski, Shah, & Edelman, 2012). The per-game model-based and model-free weightings β_{MB} and β_{MF} , indicating the strength of each type of learning, were extracted for further analysis.

These estimates were used as dependent variables in a series of regression analyses with group as the main explanatory variable of interest. All analyses controlled for task type, session, IQ, and age as additional independent variables. In follow-up analyses, we included interactions of Group \times Session or Task and measures of eating disorder disease severity (BMI, duration illness, EDE-Q), and specifically of compulsivity (YBC-EDS), as well as anxiety (STAI) and depression (BDI). Group, task type, and session were dummy coded, and all covariates were z scored. The regressions were conducted using mixed-effects logistic regression and estimated using Julia's MixedModels package. All within-participant parameters (e.g., task and session) were taken as random effects per participant, so as to capture the repeated-measure structure of the data (e.g., tasks repeated at two time points) and also the imbalanced data (e.g., not all participants completed both time points).

Examination of Approximate Learning Effects in Raw Choice Data (Logistic Regression)

To visualize approximate model-based and model-free learning effects in more interpretable terms closer to raw data, we followed up our main analysis with a factorial mixed-effects logistic regression analysis (Daw et al., 2011), which considers each trial's Stage 1 choice in terms

only of the events on the previous trial. This corresponds to a limiting case of the computational model in the case where the learning rate $\alpha = 1$; that is, choices are driven only by the most recent feedback, unlike the full model with $\alpha < 1$ where value estimates are built up incrementally over multiple trials and guide choice. Hence, by neglecting the effect of earlier trials' events, we can visualize the patterns of events approximately corresponding to model-based and model-free learning. This analysis considers whether a current Stage 1 choice (coded as stay = 1 and switch = 0, relative to the preceding trial's choice) is influenced by Reward (coded as rewarded = 1 and unrewarded = -1), Transition (coded as common = 1 and rare = -1), and their interaction on the previous trial. The interaction between reward and transition is taken to indicate the contribution of model-based learning, whereas a main effect of reward is taken to indicate a significant contribution of model-free learning (Daw et al., 2011).

All four terms of the model (the intercept, Reward and Transition main effects, and Reward \times Transition interaction) were further interacted with Group, Session, and Task (dummy coded) and Age and IQ (z scored) as in the main analyses. All 12 within-participant coefficients (the four base effects, and their interactions with task and session) were initially taken as random effects by participants. However, this specification encountered numerical errors related to singular covariance estimates, apparently because of near-zero variation between participants in the main effect of transition and its interactions with session and task. (Principal components analysis on the random effects also indicated that the remaining 9 of 12 random effects captured 99.995% of the variance [Bates, Kliegl, Vasishth, & Baayen, 2015].). Accordingly, we omitted these three random effects from the specification to arrive at an estimable model.

Simulation

We verified that the full computational model could capture the key effects in the raw data by extracting the estimated parameters for each participant and game (one 200-trial game for each game in the original data set, including participant, session, and task type, resulting in 281 games in all), simulating model performance on the task (learning and drawing choices according to the model with those parameters, rewarded according to the same procedure as the original experiment). The resulting simulated data set was plotted according to the same factorial stay/switch analyses as for the original data.

Power Analysis

The current study included a sample size comparable to previous studies in psychiatric and neurological patient populations in the laboratory, although the sample was relatively small compared with recent online studies (e.g., Gillan, Kosinski, et al., 2016). A priori power ana-

lysis was not carried out, but we conducted post hoc power analysis to assess whether the current study was adequately powered. On the basis of a recent study that included neurological (patients with Parkinson's disease) and HC groups and used an analysis approach (computational model) similar to ours (Sharp et al., 2016; supplementary material), we computed an effect size of $d = 0.65$ for the key group difference in model-based learning (here viewed, for this purpose, as equivalent to a two-sample t test). Using the software package G*Power (Faul, Erdfelder, Lang, & Buchner, 2007; Version 3.1) with power ($1 - \beta$) set at 0.80 and $\alpha = .05$, two-tailed, we determined that a sample size of 38 per group was required. Thus, the current study was adequately powered. In addition, the inclusion of two assessments (money and food versions) per time point doubled the available data, further increasing power for the overarching group comparison.

Reliability Analysis

Some concerns have been raised about the reliability of the two-step task (Shahar et al., 2019), although more detailed follow-up analyses have been reassuring (Brown et al., 2020). Reliability varies depending on the dependent measure (model-based vs. model-free), computational model, and estimation method, and analytic approaches similar to those used in the current study yield reliability in the fair-to-good range (Brown et al., 2020). Nonetheless, we additionally calculated reliability for model-based and model-free effects from the current data set. Test-retest reliability was assessed by estimating the across-session correlation coefficient from the hierarchical model of Haaf and Rouder (2019), estimated using Markov chain Monte Carlo sampling as implemented in Stan software (Stan Development Team, 2018). We estimated the model over the full data set, comparable to the way the main analyses were conducted, and further (with caution because of the smaller size of the subgroups) investigated reliability broken down by group and task by reestimating the model on subsets of the data. In all cases, we also included the data from participants of the appropriate groups for whom only Session 1 was obtained. (Although these do not directly influence estimated variability on retest, their inclusion does help to achieve better estimates of the Session 1 effects especially in smaller subsets of the data.)

RESULTS

Demographic and Psychopathology Measures

As expected, the HC and AN groups differed on measures of eating disorder severity (EDE-Q), depression (BDI), anxiety (STAI), and BMI at Time 1 but did not differ on age (see Table 1). Although all participants had IQ scores in the normal range, there was a statistically significant group difference; IQ (which has been associated with model-

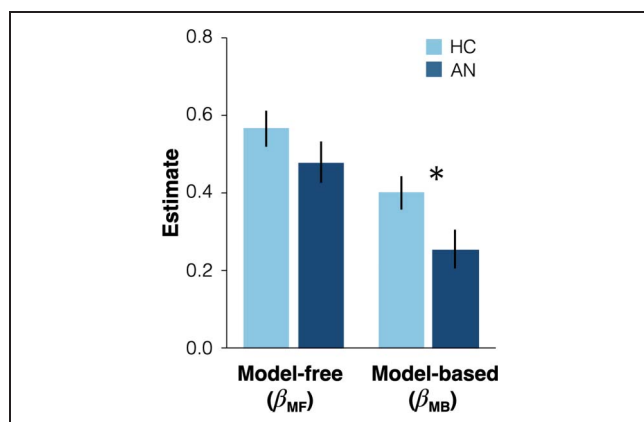


Figure 2. Overall model-free and model-based contributions to learning for HCs and individuals with AN across task type (monetary and food) and session (Time 1 and Time 2). Error bars represent *SEM*.

based learning in previous studies [Gillan, Kosinski, et al., 2016]) was included as a covariate in all analyses. In addition, age was included as a covariate, as the sample spanned adolescents and adults and use of model-based learning has been shown to develop gradually across adolescence (Decker et al., 2016). Mean BMI increased significantly from Time 1 to Time 2 among individuals with AN and did not differ significantly between groups at Time 2, indicating successful weight restoration among individuals with AN. Among individuals with AN, the expected psychological change was seen as measures of eating disorder pathology, depression, and anxiety, significantly improved from Time 1 to Time 2, although these measures remained elevated relative to HC (see Table 1).

Model-based and Model-free Learning in AN vs. HC

To examine task performance, we first tested whether individuals with AN differed from HCs in model-free and model-based learning overall (β_{MB} and β_{MF}), across both monetary and food outcome versions at both assessment points (modeling any effects of Task, Session, IQ, and Age, but considering first the main effect of Group). There

were a significant contribution of model-free learning to behavior in both groups ($ps < 1e-5$) and no significant difference between groups (Est = 0.09, $SE = 0.06$, $z = 1.36$, $p = .174$; Figure 2, Table 2). In contrast, model-based learning was also present in both groups ($ps < 1e-5$) but significantly attenuated in the AN group relative to the HC group (Est = 0.15, $SE = 0.06$, $z = 2.27$, $p = .023$; Figure 2, Table 2). There were no significant Group effects on other parameters (β_{stick} , α , β_{stage2} ; $ps > .14$).

Domain Specificity of Deficits in Model-based Learning and Effects of Treatment

There was a tendency toward greater model-based learning in the monetary task relative to the food task (main effect of task, Table 2; Est = 0.1, $SE = 0.04$, $z = 2.98$, $p = .002$) and a main effect of Session (Time 1 vs. Time 2) reflecting greater model-based learning on retest (Est = 0.11, $SE = 0.04$, $z = 2.55$, $p = .011$; Table 2).

We followed up these results to test whether it differed by group and, conversely, whether the model-based deficit was domain-general or specific to monetary versus food outcomes (modeling the interactions between Group and Task Type, again alongside effects of IQ and Age). Task Type did not interact significantly with group differences in model-based learning (Est = -0.08 , $SE = 0.07$, $z = -1.11$, $p = .27$), indicating that the same pattern of impaired model-based learning in AN was observed regardless of outcome (Figure 3A, Table 3). To further investigate this negative result for the Task Type \times Group interaction, we computed Bayes factors using a Bayesian information criterion approximation to the model evidence, comparing nested models with and without the interaction. The log Bayes factor was 2.2 in favor of the model without the Task Type \times Group interaction. This constitutes “positive” evidence for the null hypothesis under the classification of Kass and Raftery (1995) but misses their threshold (log BF = 3, or 20:1 odds, often viewed as analogous to $p < .05$) for “strong” evidence.

In another follow-up analysis, we also compared the two variants of the food task. The results were qualitatively and

Table 2. Model-based Learning (Related to Figure 2)

	Model-based Learning				Model-free Learning			
	Estimate	SE	z	p	Estimate	SE	z	p
Intercept	0.26	0.05	5.12	$<1e-6$	0.48	0.05	8.79	$<1e-17$
Group: HC	0.15	0.06	2.27	.023	0.09	0.06	1.36	.17
Task type: Money	0.1	0.04	2.98	.002	-0.02	0.04	-0.5	.62
Session: Time 2	0.11	0.04	2.55	.011	-0.02	0.04	-0.35	.72
IQ (z scored)	0.05	0.03	1.47	.142	0.03	0.03	0.87	.38
Age (z scored)	-0.002	0.03	-0.08	.934	0.03	0.03	1.01	.31

Regression analysis including group, session, and task type.

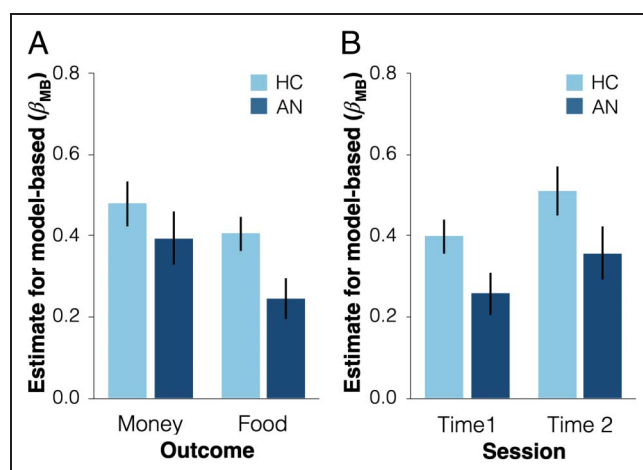


Figure 3. (A) Model-based contributions to learning for the monetary and food tasks collapsed across session (Time 1 and Time 2). (B) Model-based contributions to learning at Time 1 and Time 2 collapsed across task type (monetary and food task). Error bars represent *SEM*.

quantitatively consistent for both variants; there was no significant main effect of the Task version (Est = 0.03, *SE* = 0.08, $z = 0.34$, $p = .74$) nor an interaction with Group (Est = 0.01, *SE* = 0.11, $z = 0.12$, $p = .90$) on model-based learning.

We additionally assessed correlations between model-based estimates for monetary and food outcomes across all individuals and found significant correlations at both Time 1 ($\rho_{85} = 0.33$, $p < .001$) and Time 2 ($\rho_{51} = 0.47$, $p < .001$).

Following up the main effect of Session (Time 1 vs. Time 2) in the initial analysis (Table 2), to test for an interaction between Session and Group on model-based learning (as might be expected if the deficit improved with weight restoration), no significant effect was found (Est = 0.02, *SE* = 0.08, $z = 0.27$, $p = .79$; Figure 3B, Table 3). Interrogating the negative result, the log Bayes factor was 2.8 in favor of the model without the Session \times Group interaction, again constituting “positive” but not “strong” evidence for the null hypothesis (Kass & Raftery, 1995).

Logistic Regression Analysis of Raw Choice Behavior

To visualize how these results are reflected in raw choice behavior, as in previous studies (Daw et al., 2011), we plotted stay versus switch behavior (Figure 4) as a function of the previous trial’s reward and transition type. Here, the model-free effect (with respect to only the previous trial’s events) is approximated by the main effect of Reward, and model-based learning is similarly assessed by the size of the interaction between Reward and Transition (Daw et al., 2011). Similar to the results from the full model described above, a reduction in model-based learning is apparent in the AN group relative to the HC group (Figures 4 and 5A). This is easiest to appreciate, across conditions, by plotting an index of the approximate model-based effect (the size of the interaction using a contrast: the stay probability for common/rewarded plus rare/unrewarded trials, minus rare/rewarded and common/unrewarded; Figure 5C). We also tested this appearance statistically using mixed-effects logistic regression. However, in this coarser analysis, the Group effect on model-based learning was only marginally significant (Est = 0.13, $z = 1.96$, $p = .0501$; Table 4) and did not survive including Age and IQ in the model (Est = 0.08, $z = 1.12$, $p = .26$; Table 4).

To verify that the full computational model was able to recapitulate these patterns of observations, we simulated it playing the task (participant by participant and session by session, using the estimated parameters for each, to produce a full simulated data set) and plotted the same quantities for the simulated data set (Figure 5).

Posttest Assessment of Task Transition Structure

HC and AN groups did not differ significantly in their recall of the transition structure for the monetary or food outcomes task at Time 1 (money: $\chi^2(1, n = 87) = 0.274$, $p = .60$; food: $\chi^2(1, n = 84) = 0.925$, $p = .34$) or Time 2 (money: $\chi^2(1, n = 52) = 0.006$, $p = .94$; food: $\chi^2(1, n = 52) = 0.650$, $p = .42$). HC and AN groups also did not differ significantly

Table 3. Model-based Learning (Related to Figure 3)

	<i>Estimate</i>	<i>SE</i>	<i>Z</i>	<i>p</i>
Intercept	0.25	0.05	4.74	$<1e-5$
Group: HC	0.16	0.07	2.26	.024
Task type: Money	0.15	0.05	2.78	.006
Session: Time 2	0.09	0.06	1.44	.15
IQ (z scored)	0.05	0.03	1.46	.15
Age (z scored)	0.00	0.03	−0.08	.93
Group: HC \times Task Type: Money	−0.08	0.07	−1.11	.27
Group: HC \times Session: Time 2	0.02	0.08	0.27	.79

Regression analysis including interaction between group and task type and between group and session.

Figure 4. Behavioral data—stay probability at Stage 1 as a function of transition type (common/rare) and the outcome (rewarded/unrewarded) on the previous trial—across groups, task types, and sessions. (A) Food outcome task at Time 1. (B) Food outcome task at Time 2. (C) Money outcome task at Time 1. (D) Money outcome task at Time 2. Error bars represent *SEM*.

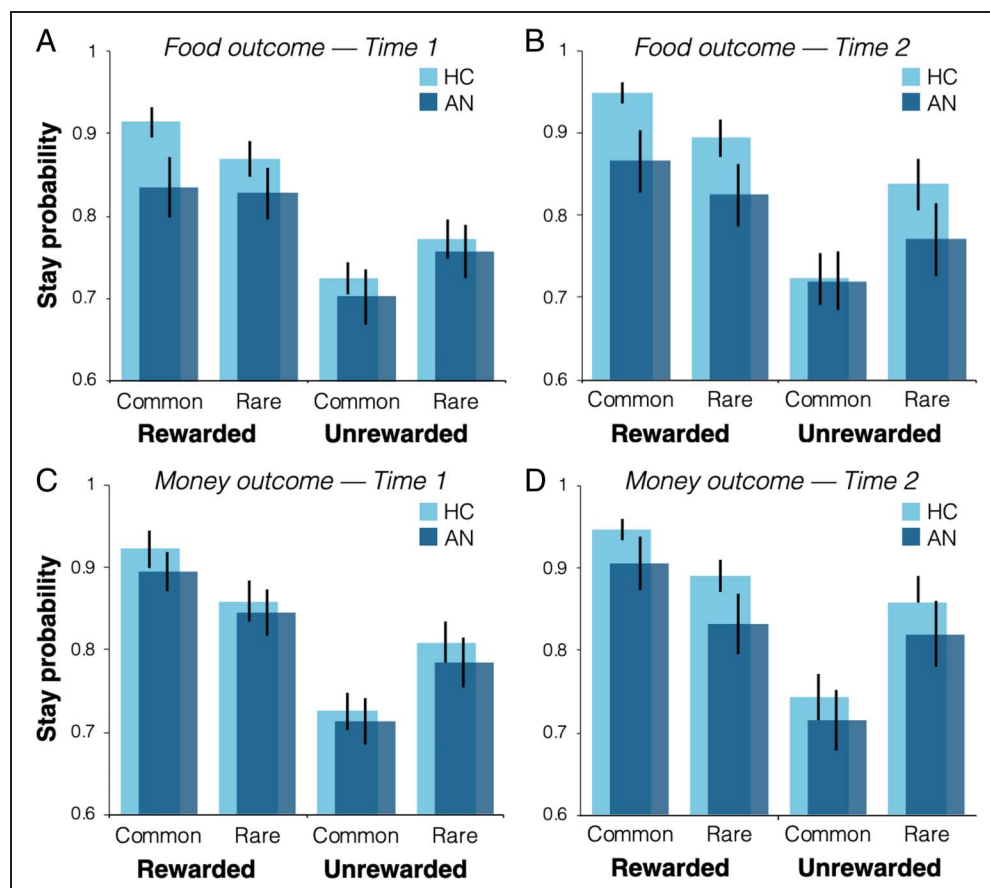


Figure 5. Raw data (A, C) and simulations (B, D) for overall stay probability at Stage 1 (A, B) and model-based effect (contrast measuring size of Reward \times Rare interaction) for task outcome types and sessions (C, D). Error bars represent *SEM*.

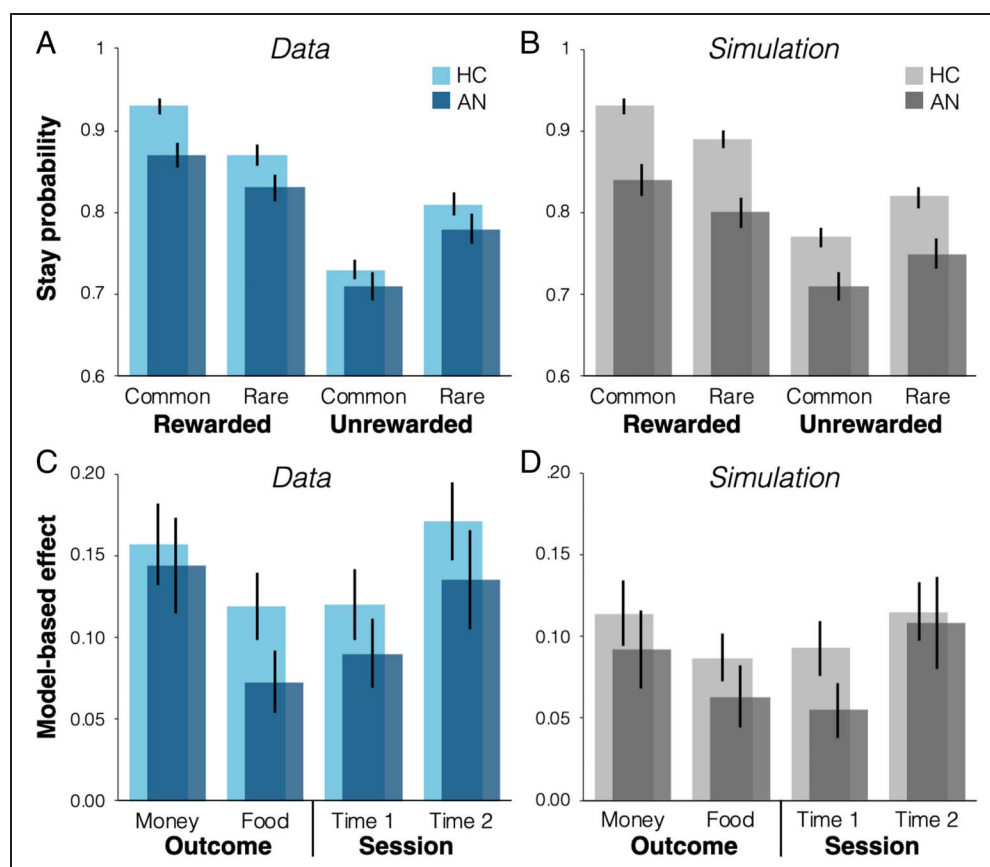


Table 4. Logistic Regression Analyses of Raw Choice Including Group, Session, and Task Type (Related to Figure 4)

	<i>Without Age and IQ</i>				<i>With Age and IQ</i>			
	<i>Est.</i>	<i>SE</i>	<i>z</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>z</i>	<i>p</i>
Intercept	1.28	0.23	5.50	<1e-7	1.31	0.23	5.72	<1e-7
Transition	0.04	0.05	0.91	.36	0.07	0.05	1.50	.13
Reward	0.57	0.11	5.37	<1e-7	0.59	0.10	5.84	<1e-8
Group: HC	0.33	0.21	1.52	.13	0.23	0.22	1.05	.29
Task type: Money	0.13	0.08	1.59	.11	0.12	0.08	1.40	.16
Session	0.40	0.11	3.75	.0002	0.39	0.11	3.63	.0003
Age (<i>z</i> scored)					0.22	0.10	2.15	.032
IQ (<i>z</i> scored)					0.29	0.11	2.71	.0068
Tran × Rew	−0.08	0.08	−1.07	.28	−0.03	0.08	−0.32	.75
Tran × Group: HC	0.04	0.03	1.56	.12	0.03	0.03	0.86	.39
Tran × Task: Money	0.01	0.03	0.43	.67	0.01	0.03	0.41	.68
Tran × Session	−0.02	0.03	−0.82	.41	−0.03	0.03	−1.15	.25
Tran × Age					−0.01	0.02	−0.90	.37
Tran × IQ					0.01	0.01	0.70	.48
Rew × Group: HC	0.13	0.08	1.61	.11	0.10	0.08	1.20	.23
Rew × Task: Money	−0.02	0.06	−0.37	.713	−0.05	0.05	−0.99	.32
Rew × Session	0.02	0.05	0.43	.67	0.03	0.05	0.48	.63
Rew × Age					0.11	0.04	2.65	.008
Rew × IQ					0.09	0.04	2.29	.022
Tran × Rew × Group: HC	0.13	0.07	1.96	.0501	0.08	0.07	1.12	.26
Tran × Rew × Task: Money	0.16	0.04	4.27	<1e-4	0.14	0.04	3.73	.0002
Tran × Rew × Session	0.23	0.05	4.63	<1e-5	0.21	0.05	4.37	<1e-4
Tran × Rew × Age					0.00	0.04	0.13	.90
Tran × Rew × IQ					0.08	0.04	2.35	.019

Rew = reward; Tran = transition.

Table 5. Model-based Learning and Disease Severity

	<i>Estimate</i>	<i>SE</i>	<i>z</i>	<i>p</i>
BMI (<i>z</i> scored)	0.003	0.03	0.09	.928
EDE-Q (<i>z</i> scored)	0.05	0.06	0.86	.39
BDI (<i>z</i> scored)	0.03	0.05	0.68	.50
STAI (<i>z</i> scored)	−0.07	0.06	−1.14	.25
YBC-EDS Total ^a (<i>z</i> scored)	−0.09	0.06	−1.64	.10
Duration illness ^a (<i>z</i> scored)	0.05	0.07	0.80	.42

Regression analyses including measures of eating disorder severity and general psychopathology (BMI, EDE-Q, BDI, STAI, YBC-EDS, and duration of illness). Each measure was included in a separate regression analysis including Group, Session, Task Type, and measure of interest.

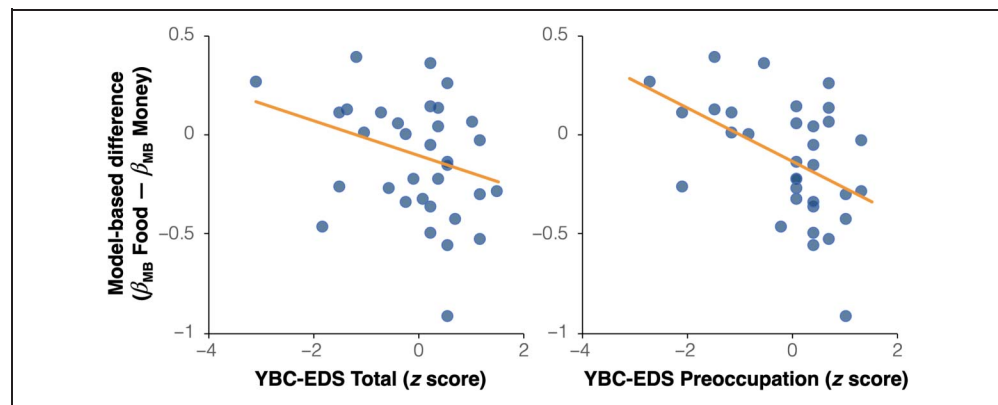
^a YBC-EDS and duration of illness obtained only in the AN group at Time 1.

Table 6. Model-based Learning and YBC-EDS

	<i>Estimate</i>	<i>SE</i>	<i>z</i>	<i>p</i>
Intercept	0.25	0.05	5.02	<1e-6
Task type: Money	0.14	0.048	2.91	.004
YBC Total (<i>z</i> scored)	−0.06	0.047	−1.35	.178
IQ (<i>z</i> scored)	0.01	0.05	0.18	.854
Age (<i>z</i> scored)	−0.02	0.045	−0.49	.623
Task type: Money × YBC	0.09	0.046	1.95	.051
Task type: Money × YBC Rituals	0.04	0.048	0.77	.44
Task type: Money × YBC Preoccupations	0.13	0.042	3.07	.002

Interactions between Task Type (for individuals with AN at Time 1).

Figure 6. Association between model-based learning for food versus money outcomes and the YBC-EDS Total score (left) and Preoccupations subscale (right) in individuals with AN.



in the average discrepancy between their estimates of the transition probabilities and the actual probabilities at Time 1 (money: $t(82) = -0.91, p = .37$; food: $t(82) = -0.59, p = .56$) or Time 2 (money: $t(51) = -1.73, p = .09$; food: $t(51) = -1.34, p = .18$). Responses on the posttest assessment were not collected from three HCs for the monetary task at Time 1 and from one individual with AN on both the monetary and food tasks at Time 2.

Correlations with Clinical Variables

Model-based behavior was not significantly associated with BMI, EDE-Q, BDI, STAI (over and above the group difference), or duration of illness in the patient group (Table 5), although there was a trend-level association with YBC-EDS (Est = $-0.09, SE = 0.05, z = -1.64, p = .10$) at Time 1 (the only time point at which it was

obtained: YBC-EDS data were not collected from HCs, for whom scores would be expected to be zero, or from those with AN after weight restoration). Exploring this last association further, in those with AN at Time 1, higher scores on the YBC-EDS were associated with less model-based learning for food outcomes relative to monetary outcomes (Est = $0.09, SE = 0.05, z = 1.95, p = .051$; Table 6, Figure 6A), a trend mainly driven by (and significant for) the Preoccupations subscale (Est = $0.13, SE = 0.04, z = 3.07, p = .002$; Figure 6B) rather than the Rituals subscale (Est = $0.04, SE = 0.05, z = 0.77, p = .44$).

Although the study was not designed to compare AN-R and AN-BP subtypes as the subgroup sample sizes were small (AN-R/AN-BP, $n = 19/22$), we include the comparison for completeness. We tested the main effect of subgroups on model-based learning overall (β_{MB}), modeling any effects of Task, Session, IQ, and Age. The AN-BP subgroup

Table 7. Reliability Analysis Results

		<i>r, Model-based</i>		<i>r, Model-free</i>	
		<i>Median</i>	<i>(95% Credible Interval)</i>	<i>Median</i>	<i>(95% Credible Interval)</i>
<i>All participants</i>					
Both tasks	$n = 54 + 40$	0.74	(.42, .89)	0.61	(.31, .78)
Food task	$n = 54 + 40$	0.62	(.27, .83)	0.54	(.23, .74)
Money task	$n = 54 + 40$	0.70	(.42, .88)	0.68	(.41, .83)
<i>HC</i>					
Both tasks	$n = 29 + 24$	0.69	(.02, .94)	0.72	(.33, .90)
Food task	$n = 29 + 24$	0.66	(.17, .92)	0.64	(.28, .86)
Money task	$n = 29 + 24$	0.46	($-.03, .79$)	0.63	(.25, .86)
<i>AN</i>					
Both tasks	$n = 25 + 16$	0.74	(.39, .91)	0.34	($-.17, .72$)
Food task	$n = 24 + 16$	0.58	($-.01, .89$)	0.30	($-.24, .69$)
Money task	$n = 25 + 16$	0.90	(.57, .99)	0.71	(.26, .91)

was least model-based, with AN-R falling in between AN-BP and HC (HC vs. AN-BP: $Est = 0.21$, $SE = 0.06$, $z = 3.36$, $p = .0008$; HC vs. AN-R: $Est = -0.09$, $SE = 0.08$, $z = -1.05$, $p = .29$; AN-R vs. AN-BP: $Est = 0.12$, $SE = 0.09$, $z = 1.33$, $p = .18$). Given the small sample size for these comparisons, the result should be treated with caution and examined in a larger study aimed at understanding possible differences between subtypes of AN.

Test-Retest Reliability

Some reports have suggested poor reliability of the two-step decision task under some circumstances (Shahar et al., 2019). However, more detailed follow-up analyses show that these conclusions depend substantially on the dependent measure (model-based vs. model-free), computational model, and estimation method. Overall, analytic approaches similar to those used in the current study yield reliability in the fair-to-good range (.4–.75; (Brown et al., 2020), a range that also held for test-retest reliability estimated from our current data (Table 7).

DISCUSSION

This study addressed three questions about model-based and model-free learning mechanisms in AN. We tested whether AN, which appears at once excessively goal-directed and habitual, is characterized by enhanced or diminished model-based behavior. We further examined model-based versus model-free learning in a food-specific context as well as a monetary context, to test the domain specificity of any differences, and before and after weight restoration, to test whether impairments were present only during acute illness in patients with AN. Individuals with AN showed less model-based learning than HCs, and groups did not differ significantly in model-free learning. This group difference was present when playing for both food and monetary outcomes and persisted after successful weight-restoration treatment.

How Does the Extreme, Yet Inflexible, “Self-control” That Characterizes AN Relate to Model-based Learning?

AN poses a fascinating conceptual challenge for the classic psychological dichotomy between goal-directed, controlled behavior and inflexible, habitual responses. Maladaptive eating behaviors in AN are commonly understood to reflect heightened self-control (King et al., 2019; Wang et al., 2019), which might be expected to relate to enhanced dominance of model-based behavior. Yet, these same behaviors are also rigid and difficult to change—selection of a low-fat, low-calorie diet with limited variety is a stereotyped feature of illness in AN (Mayer et al., 2012; Schebendach et al., 2008; Sysko, Walsh, Schebendach, & Wilson, 2005)—which might suggest the opposite result: relatively weakened model-based control.

Our results suggest that, compared to HCs, those with AN rely significantly less on the model-based approach, similar to individuals with other disorders involving compulsion such as OCD or drug abuse (Gillan, Kosinski, et al., 2016; Voon et al., 2015). The pathology of AN involves complex, multi-step phenomena to avoid food intake. The findings in this study suggest that these computational mechanisms may be relevant even to producing behaviors, like pathological food avoidance, that extend beyond the traditional notion from animal behavioral psychology of habits as simple stimulus–response motor programs. Indeed, a recent study in a large general population sample found that model-based behavior in the two-step task correlated with a set of psychiatric symptoms that included not only simple compulsive actions (such as repetitive checking or morning drinking) that extend easily from a basic notion of habits but also broader, more cognitive symptoms such as intrusive thoughts (Gillan, Kosinski, et al., 2016).

In contrast to the difference in the model-based approach, this study found no significant difference between AN and HC in the model-free approach in the task. That said, as with previous studies using this task (Gillan, Kosinski, et al., 2016), we focused our hypotheses and analyses on the modulation of model-based behavior, as this has proven to be the measure from this task that has most reliably tracked manipulations or individual differences likely to be relevant to the goal–habit balance (Gillan et al., 2020; Wyckmans et al., 2019; Sharp et al., 2016; Gillan, Otto, et al., 2015; Otto, Raio, et al., 2013). Although changes in this balance might be expected also to be reflected in countervailing changes in model-free learning, studies using this task have not generally reported concomitant differences in this measure. This may reflect some combination of at least three factors: first, a true difference in the biological substrate, that is, model-free learning may be less sensitive to neurocognitive dysfunction (Keramati, Dezfouli, & Piray, 2011; Reber, 1989), and second, a difference in the mapping from neuropsychological categories to computational substrate, that is, the link between model-based learning and goal-directed behavior may be tighter than that between model-free learning and habitual behavior (Dezfouli & Balleine, 2012). Third, the task itself may be less sensitive to model-free than model-based learning (e.g., Table 7). Regardless, model-free and model-based learning need not directly trade off between each other.

Does Specific Psychopathology Reflect a Domain-General Failure of Model-based Learning?

Our data indicate that the deficit in model-based learning in AN was not significantly remediated in a food context, in which participants played to receive the food they preferred. Rather, it was consistent across both domain-general (money) and illness-specific (food) outcomes, with no significant task differences in the group effect. This finding cuts against one potential explanation for the current

results on the money task and other similar results showing reduced model-based learning for money in psychiatric populations (Wyckmans et al., 2019; Voon et al., 2015): that generic monetary outcomes are less motivating in psychiatric populations than they are for HC participants. By using an incentive-compatible, individually determined, preferred food outcome (success on the two-step task led to actual snack choices after the task), this study included a condition in which motivation would be expected to be more salient for individuals with AN than for HCs. Yet, even in the food context, when performance directly impacted the food they would need to consume, patients with AN showed less model-based behavior than HCs. Of course, the failure to detect a Group \times Task interaction is a negative result, and as with any such result, the possibility remains that there does exist some interaction smaller than our study was powered to detect. However, the Bayes factor analysis provides some evidence for the lack of effect here, albeit not reaching the level (log Bayes factor > 3) traditionally viewed as comparable to $p < .05$. In addition, this negative effect is mitigated by another relevant positive result: that those with AN were significantly impaired at model-based learning, relative to HCs, even in the food task considered alone ($Est = 0.17$, $SE = 0.067$, $z = 2.6$, $p = .01$). Moreover, model-based estimates for monetary and food outcomes were correlated across individuals, consistent with a domain-general deficit in model-based learning and inconsistent with the idea that a focus on particular outcomes, such as food, precludes focus on other outcomes. The domain generality is interesting given the idiosyncratic nature of compulsions across individuals, and future work is needed to understand how deficits in model-based learning may lead to highly specific maladaptive behaviors.

Exploratory analyses of the relationship between symptom severity and task performance in AN did suggest one result that was selective for the food task: Higher scores on the YBC-EDS, especially the Preoccupations subscale, were associated with less model-based learning for food outcomes relative to monetary outcomes. Such a graded deficit seems plausible, and its specificity to food seems intriguing, but it is of note that the direction of the effect (worse learning with food outcomes) is opposite that predicted by a motivational account. In addition, given the exploratory nature of the analysis and the smaller number of data points underlying it, there is a danger of false positives, and replication will be required to confirm this finding. Assessing a potential lack of domain specificity of model-based deficits in individuals with AN—and other disorders characterized by compulsivity for which specificity has not been examined—is relevant for understanding the role of model-basedness in psychopathology. It is possible that decreased model-based learning constitutes a general vulnerability to illness that manifests in quite specific ways (e.g., food restriction, gambling). If so, it could be interesting to test whether interventions to increase model-based behavior in general

(cf. Patzelt, Kool, Millner, & Gershman, 2019) may affect disorder-specific psychopathology.

Are Model-based Deficits Secondary to Psychopathology or Potentially Primary?

The idea that compulsive psychopathology might relate to impaired model-based learning is appealing in part because it hints at a causal mechanism by which pathological habits might emerge. This study does not address causality, nor even whether the cognitive deficits precede the symptoms. However, we were able to address whether the cognitive differences between individuals with AN and HCs persisted after acute weight restoration. That they did decreases the likelihood that these differences are because of starvation alone. Prior research has shown that, whereas psychological measures improve with weight restoration, eating behavior—namely, the pursuit of low-fat low-calorie diets—does not improve substantially with weight restoration (Mayer et al., 2012; Sysko et al., 2005). Similarly, in a previous study, we found that deficits in feedback-based learning remained after weight restoration (Foerde & Steinglass, 2017). This result also parallels a recent report, using the same task, that cognitive behavioral therapy for OCD does not improve model-based deficits in that population (Wheaton et al., 2019). In the current data set, the absence of an effect of weight restoration is a negative result and subject to similar interpretational caveats as the money versus food contrast. However, again, the Bayes factor analysis provides some evidence for the lack of effect, and the frequentist confidence interval on the effect also suggests that any undetected effect of weight restoration on model-based behavior is likely to be quite small relative to the overall deficit.

Conclusion

Previous work has shown deficits in responding to reward (DeGuzman, Shott, Yang, Riederer, & Frank, 2017; O'Hara, Campbell, & Schmidt, 2015; Frank, Shott, Hagman, & Mittal, 2013) and learning from feedback in individuals with AN (Foerde & Steinglass, 2017; Shott et al., 2012; Lawrence et al., 2003). The present results suggest that a more fine-grained parsing of this deficit is necessary. Basic habitual learning mechanisms may be intact, whereas flexible responses to changing contingencies and the ability to integrate a model of the environment into choices are impaired. Individuals with AN understood the structure of the task as well as did HCs, as both groups were able to report verbally on the rules of transition in the task. Yet, individuals with AN were less able to successfully use this information in their decision-making during the task.

Although decreased model-based learning was present among individuals with AN across monetary and food outcomes, cutting against a general motivational account for the deficit, it is still possible that individuals with AN and HC individuals approached the task with differences in motivation to approach versus avoid outcomes. Future

studies should examine how choices in individuals with AN are related to motivation to obtain low-fat, low-calorie foods versus motivation to avoid high-fat, high-calorie foods.

The divergence between model-based and model-free behavior in individuals with AN could have implications for how illness develops and how it becomes resistant to change. It is particularly important for an illness that often develops in adolescence, a time during which model-free learning reigns and model-based behavior only begins to emerge (Decker et al., 2016). The slower emergence of model-based behavior across development suggests one mechanism that may nudge dieting behavior down the habit (model-free) path. Even further decreased model-based behavior in some individuals may confer added

vulnerability to developing maladaptive behaviors and may be one contributing factor to how eating behavior becomes so entrenched.

More generally, these results also speak to the promise of the goal-directed versus habitual dichotomy as a mechanism of compulsive symptomatology across disorders: The decision-making deficit here was domain general, not food specific, and not apparently secondary to starvation. In summary, we found that model-based, and not model-free, behavior was impaired among individuals with AN relative to HCs. The deficit was not remediated by using real food as the outcomes of choices to enhance motivational salience nor by weight restoration and associated improvement in psychological symptoms.


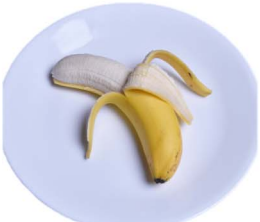





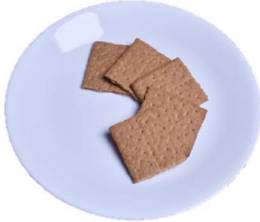






APPENDIX

Sub ID: _____

Date: _____

Session: _____

Use the numbers 1-15 to indicate your order of preference for each of these snack items
(1=Most preferred, 15=Least preferred)

 Apple _____	 Banana _____	 Brownie _____
 Carrots _____	 Cheese cubes _____	 Cookies _____
 Craisins _____	 Croissant _____	 Graham crackers _____
 Granola bar _____	 Reeses pieces _____	 Rice cakes _____
 Strawberries _____	 Sun chips _____	 Popcorn _____

Food task menu used to determine outcomes for the food version of the 2-step task. Participants numbered items in order of preference. Performance on the task determined access to food items with better performance granting access to increasingly preferred items.

Reprint requests should be sent to Karin Foerde, New York State Psychiatric Institute, 1051 Riverside Dr., New York, NY 10032, or via e-mail: kf2265@columbia.edu.

Author Contributions

Karin Foerde: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Visualization, Writing—original draft, Writing—review & editing. Nathaniel D. Daw: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Resources, Writing—original draft, Writing—review & editing. Teresa Rufin: Data curation, Investigation, Project administration. B. Timothy Walsh: Conceptualization, Methodology, Supervision, Writing—original draft, Writing—review & editing. Daphna Shohamy: Conceptualization, Funding acquisition, Writing—review & editing. Joanna E. Steinglass: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing—original draft, Writing—review & editing.

Funding Information

National Institute of Mental Health (<http://dx.doi.org/10.13039/1000000025>). Joanna E. Steinglass: K24 MH113737. Joanna Steinglass and Daphna Shohamy: R01 MH105452.

REFERENCES

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. Arlington, VA: American Psychiatric Association. **DOI:** <https://doi.org/10.1176/appi.books.9780890425596>
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.
- Beck, A. T., & Steer, R. A. (1993). *Beck depression inventory manual*. San Antonio, TX: Psychological Corporation.
- Bezanson, J., Karpinski, S., Shah, V. B., & Edelman, A. (2012). Julia: A fast dynamic language for technical computing. *arXiv preprint arXiv:1209.5145*.
- Brown, V. M., Chen, J., Gillan, C. M., & Price, R. B. (2020). Improving the reliability of computational analyses: Model-based planning and its relationship with compulsivity. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 5, 601–609. **DOI:** <https://doi.org/10.1016/j.bpsc.2019.12.019>, **PMID:** 32249207, **PMCID:** PMC7286766
- Bruch, H. (1979). *Golden cage: The enigma of anorexia nervosa*. Cambridge, MA: Harvard University Press.
- Coutureau, E., & Killcross, S. (2003). Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behavioral Brain Research*, 146, 167–174. **DOI:** <https://doi.org/10.1016/j.bbr.2003.09.025>, **PMID:** 14643469
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, 69, 1204–1215. **DOI:** <https://doi.org/10.1016/j.neuron.2011.02.027>, **PMID:** 21435563, **PMCID:** PMC3077926
- Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8, 1704–1711. **DOI:** <https://doi.org/10.1038/nn1560>, **PMID:** 16286932
- Decker, J. H., Figner, B., & Steinglass, J. E. (2015). On weight and waiting: Delay discounting in anorexia nervosa pretreatment and posttreatment. *Biological Psychiatry*, 78, 606–614. **DOI:** <https://doi.org/10.1016/j.biopsych.2014.12.016>, **PMID:** 25641636, **PMCID:** PMC4478277
- Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. *Psychological Science*, 27, 848–858. **DOI:** <https://doi.org/10.1177/0956797616639301>, **PMID:** 27084852, **PMCID:** PMC4899156
- DeGuzman, M., Shott, M. E., Yang, T. T., Riederer, J., & Frank, G. K. W. (2017). Association of elevated reward prediction error response with weight gain in adolescent anorexia nervosa. *American Journal of Psychiatry*, 174, 557–565. **DOI:** <https://doi.org/10.1176/appi.ajp.2016.16060671>, **PMID:** 28231717, **PMCID:** PMC5607032
- Dezfouli, A., & Balleine, B. W. (2012). Habits, action sequences and reinforcement learning. *European Journal of Neuroscience*, 35, 1036–1051. **DOI:** <https://doi.org/10.1111/j.1460-9568.2012.08050.x>, **PMID:** 22487034, **PMCID:** PMC3325518
- Dickinson, A., & Balleine, B. (2002). The role of learning in the operation of motivational systems. In H. Pashler & R. Gallistel (Eds.), *Steven's handbook of experimental psychology: Learning, motivation, and emotion* (pp. 497–533). New York: Wiley. **DOI:** <https://doi.org/10.1002/0471214426.pas0312>
- Everitt, B. J., & Robbins, T. W. (2016). Drug addiction: Updating actions to habits to compulsions ten years on. *Annual Review of Psychology*, 67, 23–50. **DOI:** <https://doi.org/10.1146/annurev-psych-122414-033457>, **PMID:** 26253543
- Fairburn, C. G. (2008). *Cognitive behavior therapy and eating disorders*. New York: Guilford Press.
- Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16, 363–370. **DOI:** [https://doi.org/10.1002/1098-108X\(199412\)16:4<363::AID-EAT2260160405>3.0.CO;2-%23](https://doi.org/10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-%23), **PMID:** 7866415
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. **DOI:** <https://doi.org/10.3758/bf03193146>, **PMID:** 17695343
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). *Structured clinical interview for DSM-IV axis I disorders, research version, patient edition (SCID-I/P)*. New York: Biometrics Research, New York State Psychiatric Institute.
- Foerde, K., & Steinglass, J. E. (2017). Decreased feedback learning in anorexia nervosa persists after weight restoration. *International Journal of Eating Disorders*, 50, 415–423. **DOI:** <https://doi.org/10.1002/eat.22709>, **PMID:** 28393399, **PMCID:** PMC5869029
- Foerde, K., Steinglass, J. E., Shohamy, D., & Walsh, B. T. (2015). Neural mechanisms supporting maladaptive food choices in anorexia nervosa. *Nature Neuroscience*, 18, 1571–1573. **DOI:** <https://doi.org/10.1038/nn.4136>, **PMID:** 26457555, **PMCID:** PMC4624561
- Frank, G. K., Shott, M. E., Hagman, J. O., & Mittal, V. A. (2013). Alterations in brain structures related to taste reward circuitry in ill and recovered anorexia nervosa and in bulimia nervosa. *American Journal of Psychiatry*, 170, 1152–1160. **DOI:** <https://doi.org/10.1176/appi.ajp.2013.12101294>, **PMID:** 23680873, **PMCID:** PMC3789862
- Gillan, C. M., Apergis-Schoute, A. M., Morein-Zamir, S., Urcelay, G. P., Sule, A., Fineberg, N. A., et al. (2015). Functional neuroimaging of avoidance habits in obsessive-compulsive

- disorder. *American Journal of Psychiatry*, 172, 284–293. DOI: <https://doi.org/10.1176/appi.ajp.2014.14040525>, PMID: 25526600, PMCID: PMC4910868
- Gillan, C. M., Kalanthroff, E., Evans, M., Weingarden, H. M., Jacoby, R. J., Gershkovich, M., et al. (2020). Comparison of the association between goal-directed planning and self-reported compulsivity vs obsessive-compulsive disorder diagnosis. *JAMA Psychiatry*, 77, 77–85. DOI: <https://doi.org/10.1001/jamapsychiatry.2019.2998>, PMID: 31596434, PMCID: PMC6802255
- Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife*, 5, e11305. DOI: <https://doi.org/10.7554/eLife.11305>, PMID: 26928075, PMCID: PMC4786435
- Gillan, C. M., Otto, A. R., Phelps, E. A., & Daw, N. D. (2015). Model-based learning protects against forming habits. *Cognitive, Affective & Behavioral Neuroscience*, 15, 523–536. DOI: <https://doi.org/10.3758/s13415-015-0347-6>, PMID: 25801925, PMCID: PMC4526597
- Gillan, C. M., Robbins, T. W., Sahakian, B. J., van den Heuvel, O. A., & van Wingen, G. (2016). The role of habit in compulsivity. *European Neuropsychopharmacology*, 26, 828–840. DOI: <https://doi.org/10.1016/j.euroneuro.2015.12.033>, PMID: 26774661, PMCID: PMC4894125
- Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, 66, 585–595. DOI: <https://doi.org/10.1016/j.neuron.2010.04.016>, PMID: 20510862, PMCID: PMC2895323
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Delgado, P., Heninger, G. R., et al. (1989). The Yale–Brown obsessive compulsive scale: II. Validity. *Archives of General Psychiatry*, 46, 1012–1016. DOI: <https://doi.org/10.1001/archpsyc.1989.01810110054008>, PMID: 2510699
- Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, 31, 359–387. DOI: <https://doi.org/10.1146/annurev.neuro.29.051605.112851>, PMID: 18558860
- Haaf, J. M., & Rouder, J. N. (2019). Some do and some don't? Accounting for variability of individual difference structures. *Psychonomic Bulletin & Review*, 26, 772–789. DOI: <https://doi.org/10.3758/s13423-018-1522-x>, PMID: 30251148
- Hudson, J. L., Hiripi, E., Pope, H. G., Jr., & Kessler, R. C. (2007). The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biological Psychiatry*, 61, 348–358. DOI: <https://doi.org/10.1016/j.biopsych.2006.03.040>, PMID: 16815322, PMCID: PMC1892232
- Huys, Q. J. M., Cools, R., Gölzer, M., Friedel, E., Heinz, A., Dolan, R. J., et al. (2011). Disentangling the roles of approach, activation and valence in instrumental and Pavlovian responding. *PLoS Computational Biology*, 7, e1002028. DOI: <https://doi.org/10.1371/journal.pcbi.1002028>, PMID: 21556131, PMCID: PMC3080848
- Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, 19, 404–413. DOI: <https://doi.org/10.1038/nn.4238>, PMID: 26906507, PMCID: PMC5443409
- James, W. (1890). *The principles of psychology* (Vol. 1). New York: Henry Holt. DOI: <https://doi.org/10.1037/10538-000>
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90, 773–795. DOI: <https://doi.org/10.1080/01621459.1995.10476572>
- Keramati, M., Dezfouli, A., & Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Computational Biology*, 7, e1002055. DOI: <https://doi.org/10.1371/journal.pcbi.1002055>, PMID: 21637741, PMCID: PMC3102758
- Killcross, S., & Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex*, 13, 400–408. DOI: <https://doi.org/10.1093/cercor/13.4.400>, PMID: 12631569
- King, J. A., Korb, F. M., Vettermann, R., Ritschel, F., Egner, T., & Ehrlich, S. (2019). Cognitive overcontrol as a trait marker in anorexia nervosa? Aberrant task- and response-set switching in remitted patients. *Journal of Abnormal Psychology*, 128, 806–812. DOI: <https://doi.org/10.1037/abn0000476>, PMID: 31657595
- Lawrence, A. D., Dowson, J., Foxall, G. L., Summerfield, R., Robbins, T. W., & Sahakian, B. J. (2003). Impaired visual discrimination learning in anorexia nervosa. *Appetite*, 40, 85–89. DOI: [https://doi.org/10.1016/s0195-6663\(02\)00138-1](https://doi.org/10.1016/s0195-6663(02)00138-1), PMID: 12631509
- Lee, S. W., Shimojo, S., & O'Doherty, J. P. (2014). Neural computations underlying arbitration between model-based and model-free learning. *Neuron*, 81, 687–699. DOI: <https://doi.org/10.1016/j.neuron.2013.11.028>, PMID: 24507199, PMCID: PMC3968946
- Lloyd, S., Yiend, J., Schmidt, U., & Tchanturia, K. (2014). Perfectionism in anorexia nervosa: Novel performance based evidence. *PLoS One*, 9, e111697. DOI: <https://doi.org/10.1371/journal.pone.0111697>, PMID: 25360690, PMCID: PMC4216122
- Mayer, L. E. S., Schebendach, J., Bodell, L. P., Shingleton, R. M., & Walsh, B. T. (2012). Eating behavior in anorexia nervosa: Before and after treatment. *International Journal of Eating Disorders*, 45, 290–293. DOI: <https://doi.org/10.1002/eat.20924>, PMID: 21495053, PMCID: PMC4469276
- Mazure, C. M., Halmi, K. A., Sunday, S. R., Romano, S. J., & Einhorn, A. M. (1994). The Yale–Brown–Cornell Eating Disorder Scale: Development, use, reliability and validity. *Journal of Psychiatric Research*, 28, 425–445. DOI: [https://doi.org/10.1016/0022-3956\(94\)90002-7](https://doi.org/10.1016/0022-3956(94)90002-7), PMID: 7897615
- O'Hara, C. B., Campbell, I. C., & Schmidt, U. (2015). A reward-centred model of anorexia nervosa: A focussed narrative review of the neurological and psychophysiological literature. *Neuroscience & Biobehavioral Reviews*, 52, 131–152. DOI: <https://doi.org/10.1016/j.neubiorev.2015.02.012>, PMID: 25735957
- Otto, A. R., Gershman, S. J., Markman, A. B., & Daw, N. D. (2013). The curse of planning: Dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychological Science*, 24, 751–761. DOI: <https://doi.org/10.1177/0956797612463080>, PMID: 23558545, PMCID: PMC3843765
- Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences, U.S.A.*, 110, 20941–20946. DOI: <https://doi.org/10.1073/pnas.1312011110>, PMID: 24324166, PMCID: PMC3876216
- Patzelt, E. H., Kool, W., Millner, A. J., & Gershman, S. J. (2019). Incentives boost model-based control across a range of severity on several psychiatric constructs. *Biological Psychiatry*, 85, 425–433. DOI: <https://doi.org/10.1016/j.biopsych.2018.06.018>, PMID: 30077331, PMCID: PMC6314918
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118, 219–235. DOI: <https://doi.org/10.1037/0096-3445.118.3.219>
- Salkovskis, P. M. (1985). Obsessional-compulsive problems: A cognitive-behavioural analysis. *Behavior Research and Therapy*, 23, 571–583. DOI: [https://doi.org/10.1016/0005-7967\(85\)90105-6](https://doi.org/10.1016/0005-7967(85)90105-6), PMID: 4051930
- Schebendach, J. E., Mayer, L. E. S., Devlin, M. J., Attia, E., Contento, I. R., Wolf, R. L., et al. (2008). Dietary energy density and diet variety as predictors of outcome in anorexia nervosa. *American Journal of Clinical Nutrition*, 87, 810–816. DOI: <https://doi.org/10.1093/ajcn/87.4.810>, PMID: 18400701

- Schebendach, J. E., Mayer, L. E. S., Devlin, M. J., Attia, E., & Walsh, B. T. (2012). Dietary energy density and diet variety as risk factors for relapse in anorexia nervosa: A replication. *International Journal of Eating Disorders*, *45*, 79–84. DOI: <https://doi.org/10.1002/eat.20922>, PMID: 21448937, PMCID: PMC4469286
- Shahar, N., Hauser, T. U., Moutoussis, M., Moran, R., Keramati, M., NSPN Consortium, et al. (2019). Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. *PLoS Computational Biology*, *15*, e1006803. DOI: <https://doi.org/10.1371/journal.pcbi.1006803>, PMID: 30759077, PMCID: PMC6391008
- Sharp, M. E., Foerde, K., Daw, N. D., & Shohamy, D. (2016). Dopamine selectively remediates “model-based” reward learning: A computational approach. *Brain*, *139*, 355–364. DOI: <https://doi.org/10.1093/brain/awv347>, PMID: 26685155, PMCID: PMC5868097
- Shott, M. E., Filoteo, J. V., Jappe, L. M., Pryor, T., Maddox, W. T., Rollin, M. D. H., et al. (2012). Altered implicit category learning in anorexia nervosa. *Neuropsychology*, *26*, 191–201. DOI: <https://doi.org/10.1037/a0026771>, PMID: 22201300, PMCID: PMC3295920
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *The state-trait anxiety inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Stan Development Team. (2018). Stan modeling language user's guide and reference manual (version 2.18.0). Retrieved from <http://mc-stan.org/>.
- Steinglass, J. E., Figner, B., Berkowitz, S., Simpson, H. B., Weber, E. U., & Walsh, B. T. (2012). Increased capacity to delay reward in anorexia nervosa. *Journal of the International Neuropsychological Society*, *18*, 773–780. DOI: <https://doi.org/10.1017/S1355617712000446>, PMID: 22591835, PMCID: PMC3638253
- Steinglass, J. E., Lempert, K. M., Choo, T.-H., Kimeldorf, M. B., Wall, M., Walsh, B. T., et al. (2017). Temporal discounting across three psychiatric disorders: Anorexia nervosa, obsessive compulsive disorder, and social anxiety disorder. *Depression and Anxiety*, *34*, 463–470. DOI: <https://doi.org/10.1002/da.22586>, PMID: 28009473, PMCID: PMC5869031
- Sysko, R., Roberto, C. A., Barnes, R. D., Grilo, C. M., Attia, E., & Walsh, B. T. (2012). Test–retest reliability of the proposed DSM-5 eating disorder diagnostic criteria. *Psychiatry Research*, *196*, 302–308. DOI: <https://doi.org/10.1016/j.psychres.2011.12.021>, PMID: 22401974, PMCID: PMC3361586
- Sysko, R., Walsh, B. T., Schebendach, J., & Wilson, G. T. (2005). Eating behavior among women with anorexia nervosa. *American Journal of Clinical Nutrition*, *82*, 296–301. DOI: <https://doi.org/10.1093/ajcn/82.2.296>, PMID: 16087971
- Tricomi, E., Balleine, B. W., & O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *European Journal of Neuroscience*, *29*, 2225–2232. DOI: <https://doi.org/10.1111/j.1460-9568.2009.06796.x>, PMID: 19490086, PMCID: PMC2758609
- Valentin, V. V., Dickinson, A., & O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *Journal of Neuroscience*, *27*, 4019–4026. DOI: <https://doi.org/10.1523/JNEUROSCI.0564-07.2007>, PMID: 17428979, PMCID: PMC6672546
- Vikbladh, O. M., Meager, M. R., King, J., Blackmon, K., Devinsky, O., Shohamy, D., et al. (2019). Hippocampal contributions to model-based planning and spatial memory. *Neuron*, *102*, 683–693. DOI: <https://doi.org/10.1016/j.neuron.2019.02.014>, PMID: 30871859, PMCID: PMC6508991
- Volkow, N. D., Wang, G.-J., Fowler, J. S., Tomasi, D., Telang, F., & Baler, R. (2010). Addiction: Decreased reward sensitivity and increased expectation sensitivity conspire to overwhelm the brain's control circuit. *BioEssays*, *32*, 748–755. DOI: <https://doi.org/10.1002/bies.201000042>, PMID: 20730946, PMCID: PMC2948245
- Voon, V., Derbyshire, K., Rück, C., Irvine, M. A., Worbe, Y., Enander, J., et al. (2015). Disorders of compulsivity: A common bias towards learning habits. *Molecular Psychiatry*, *20*, 345–352. DOI: <https://doi.org/10.1038/mp.2014.44>, PMID: 24840709, PMCID: PMC4351889
- Walsh, B. T. (2013). The enigmatic persistence of anorexia nervosa. *American Journal of Psychiatry*, *170*, 477–484. DOI: <https://doi.org/10.1176/appi.ajp.2012.12081074>, PMID: 23429750, PMCID: PMC4095887
- Wang, S. B., Gray, E. K., Coniglio, K. A., Murray, H. B., Stone, M., Becker, K. R., et al. (2019). Cognitive rigidity and heightened attention to detail occur transdiagnostically in adolescents with eating disorders. *Eating Disorders*, 1–13. DOI: <https://doi.org/10.1080/10640266.2019.1656470>, PMID: 31675280, PMCID: PMC7192764
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence manual*. San Antonio, TX: Psychological Corporation. DOI: <https://doi.org/10.1037/t15170-000>
- Weissengruber, S., Lee, S. W., O'Doherty, J. P., & Ruff, C. C. (2019). Neurostimulation reveals context-dependent arbitration between model-based and model-free reinforcement learning. *Cerebral Cortex*, *29*, 4850–4862. DOI: <https://doi.org/10.1093/cercor/bhz019>, PMID: 30888032
- Wheaton, M. G., Gillan, C. M., & Simpson, H. B. (2019). Does cognitive-behavioral therapy affect goal-directed planning in obsessive-compulsive disorder? *Psychiatry Research*, *273*, 94–99. DOI: <https://doi.org/10.1016/j.psychres.2018.12.079>, PMID: 30640057
- Wyckmans, F., Otto, A. R., Sebold, M., Daw, N., Bechara, A., Saeremans, M., et al. (2019). Reduced model-based decision-making in gambling disorder. *Scientific Reports*, *9*, 19625. DOI: <https://doi.org/10.1038/s41598-019-56161-z>, PMID: 31873133, PMCID: PMC6927960
- Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, *7*, 464–476. DOI: <https://doi.org/10.1038/nrn1919>, PMID: 16715055
- Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Journal of Neuroscience*, *19*, 181–189. DOI: <https://doi.org/10.1111/j.1460-9568.2004.03095.x>, PMID: 14750976
- Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2005). Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *European Journal of Neuroscience*, *22*, 505–512. DOI: <https://doi.org/10.1111/j.1460-9568.2005.04219.x>, PMID: 16045503
- Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behavioral Brain Research*, *166*, 189–196. DOI: <https://doi.org/10.1016/j.bbr.2005.07.012>, PMID: 16153716