JNeurosci THE JOURNAL OF NEUROSCIENCE

Research Articles: Behavioral/Cognitive

Model-based planning deficits in compulsivity are linked to faulty neural representations of task structure

https://doi.org/10.1523/JNEUROSCI.0031-21.2021

Cite as: J. Neurosci 2021; 10.1523/JNEUROSCI.0031-21.2021

Received: 6 January 2021
Revised: 29 April 2021
Accepted: 4 May 2021

Model-based planning deficits in compulsivity are linked to faulty neural representations of task structure

Abbrev title: Task structure representations in compulsivity

Tricia X.F. Seow,[1,2], Edith Benoit[1], Caoimhe Dempsey[1], Maeve Jennings[1], Aoibheann Maxwell[1], Redmond O'Connell[1,2] and Claire M. Gillan[1,2,3]

[1]Department of Psychology, Trinity College Dublin, Dublin, Ireland

[2]Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland

[3]Global Brain Health Institute, Trinity College Dublin, Dublin, Ireland

Correspondence: Tricia Seow, *School of Psychology, Áras an Phiarsaigh, Trinity College Dublin, Dublin 2, Ireland, +353-01-896-8506, seowx@tcd.ie*

Number of pages: 50
Number of figures: 8
Abstract word count: 237
Introduction word count: 645
Discussion word count: 1497

Conflict of interests: The authors declare no competing financial interests.

Acknowledgements: This work was supported by a Postgraduate Ussher fellowship from Trinity College Dublin to TXFS and a 'Institutional Strategic Support Fund' grant (204814/Z/16/A) to Trinity College Dublin funded by the SFI-HRB-Wellcome Trust partnership. TXFS is now a post-doctoral fellow at the Max Planck UCL Centre for Computational Psychiatry and Ageing Research.

Abstract

Compulsive individuals have deficits in model-based planning, but the mechanisms that drive this have not been established. We examined two candidates—that compulsivity is linked to (i) an impaired model of the task environment and/or (ii) an inability to engage cognitive control when making choices. To test this, 192 participants performed a two-step reinforcement learning task with concurrent EEG recordings and we related the neural and behavioral data to their scores on a self-reported transdiagnostic dimension of compulsivity. To examine subjects' internal model of the task, we used established behavioral and neural responses to unexpected events (reaction time (RT) slowing, P300 and parietal-occipital alpha-band power) measured when an unexpected transition occurred. To assess cognitive control, we probed theta power at the time of initial choice. As expected, model-based planning was linked to greater behavioral (RT) and neural (alpha power, but not P300) sensitivity to rare transitions. Critically, the sensitivity of both RT and alpha to task structure was weaker in those high in compulsivity. This RT-compulsivity effect was tested and replicated in an independent pre-existing dataset (N = 1413). We also found that mid-frontal theta power at the time of choice was reduced in high compulsive individuals though its relation to model-based planning was less pronounced. These data suggest that model-based planning deficits in compulsive individuals may arise, at least in part, from having an impaired representation of the environment, specifically how actions lead to future states.

Significance Statement

Compulsivity is linked to poorer performance on tasks that require model-based planning, but it is unclear what precise mechanisms underlie this deficit. Do compulsive individuals fail to engage cognitive control at the time of choice? Or do they have difficulty in building and maintaining an accurate representation of their environment, the foundation needed to behave in a goal-directed manner? With reaction time and EEG measures in 192 individuals who performed a two-step decision making task, we found that compulsive individuals are less sensitive to surprising action-state transitions, where they slow down less and show less alpha band suppression following a rare transition. These findings implicate failures in maintaining an accurate model of the world in model-based planning deficits in compulsivity.

Introduction

Compulsive behavior manifests as out-of-control and repetitive actions, often leading to functionally impairing outcomes (Robbins et al., 2012). This symptomology is characteristic of psychiatric disorders like obsessive-compulsive disorder (OCD) and addiction, and is thought to arise from an imbalance between two modes of action control (Gillan and Robbins, 2014): (i) goal-directed 'model-based' planning relying on knowledge of how actions lead to specific outcomes and (ii) rigid habits depending on reflexive stimulus-response associations which form slowly over time (Dickinson, 1985; Balleine and O'Doherty, 2010). The compulsivity literature has largely focused on testing if a dysfunctional imbalance in the competitive interactions between these decision systems cause habitual behaviors to dominate (Lee et al., 2014; Gruner et al., 2016), but rather than being solely an arbitration failure, recent evidence suggest that compulsivity may be primarily associated with goal-directed control impairments. For example, OCD patients have performance deficits in the two-step reinforcement task (Voon et al., 2015) where 'model-based' planning, a reinforcement-learning model of goal-directed action, is operationalized as the extent to which individuals make decisions using knowledge of how their actions relate to subsequent events (Daw et al., 2005, 2011). Recent work has shown that this dysfunction has a developmental course (Vaghi et al., 2020) and is best captured by a compulsivity dimension in both general population samples (Gillan et al., 2016).

However, it remains unclear what underlies model-based planning problems in compulsivity—a multifaceted cognitive capacity, model-based planning depends upon several functions including: (i) the construction/maintenance of an internal model (i.e., a representation of the environment, like the knowledge of relevant action-outcome and state-state relationships), which is a pre-requisite for (ii) implementation of this model in behavior through prospective planning. Model-based failures could theoretically stem from mechanistic issues underlying either component (and others not the focus of the present study). Though direct tests to resolve this have been lacking, patients show goal-directed deficits even when they have requisite explicit knowledge of simple action-outcome contingencies (Gillan et al., 2014), suggesting that OCD patients may have issues solely with implementation. But, paradigms that feature more numerous and/or taxing contingency structures revealed problems in learning action-outcome associations in OCD and addiction (Gillan et al., 2011; Ersche et al., 2016), which correlated with goal-directed control failures in OCD (Gillan et al., 2011). Overall, the evidence remain equivocal because these devaluation-style tasks conflate goal-directed control deficits with increases in stimulus-response habit learning (Watson and de Wit, 2018) and were not designed to assess participants' ability to represent the task environment.

Recent data has suggested that goal-directed failures in compulsivity might arise from the latter. For example, compulsivity is linked to poorer learning of the consequences of their actions (Sharp et al., 2020) and at the meta-level, high compulsive individuals have abnormalities in how they view their own actions, exhibiting an over-confidence, which is relatively impervious to corrective evidence (Rouault et al., 2018; Seow and Gillan, 2020). Though studied in a different context, these findings suggest the possibility that individuals high in compulsivity have fundamental issues in acquiring and maintaining an accurate internal model of the world. To date, no study has examined neural representations of task structure in compulsive individuals as they perform a model-based planning task. The present study aimed to fill this gap—testing if compulsivity is characterized by a disruption in constructing/maintaining an accurate representation of the task environment, or the use/implementation of this model in their choices. To do this, we analyzed reaction time (RT) and electroencephalography (EEG) data to define signatures of state transition knowledge (RT, P300 and posterior alpha) and of a well-established cognitive control marker (mid-frontal theta) as 192 subjects performed a two-step reinforcement learning task (Daw et al., 2005, 2011). With single-trial regression analyses, we sought to characterize several candidate neural correlates of the representation and implementation of the mental model and test if they associated with individual differences in model-based planning and compulsivity.

Materials and Methods

Power estimation. We determined a minimum sample size from a prior study that investigated the association of goal-directed control (on a different task) with OCI-R scores from non-clinical participants who were also tested in-person (r = -0.26, p < 0.05) (Snorrason et al., 2016). The effect size indicated that N = 150 participants were required to achieve 90% power at 0.05 significance. Our final sample was larger than this to achieve the required power for another study that the same subjects participated in (Seow et al., 2020).

Participant exclusion criteria. During recruitment, all participants were ensured to be ≥18 to 65 years, had no personal/familial history of epilepsy and no personal history of neurological illness/head trauma nor unexplained fainting. Participants' data were excluded from analysis if they failed any of the following on a rolling basis: Participants whose/who (i) EEG data were incomplete (N = 5) (i.e., recording was prematurely terminated before the completion of the task) or corrupted (N = 2), (ii) EEG data which contained excessive noise (i.e., >70% EEG epochs from the individual failing epoch exclusion criteria, see **EEG recording & pre-processing**) (N = 4), (iii) responded with the same key in stage one >90% (n > 135 trials) of the time (N = 10), (iv) probability of staying after common-rewarded trials was significantly worse than chance, measured as <5% probability of fitting a binomial distribution with 50% (chance) probability and the total number of common-rewarded trials experienced by each subject (N = 11), (v) missed more than 20% of trials (n > 30 trials) (N = 3), and (vi) incorrectly responded to a "catch" question within the questionnaires: "If you are paying attention to these questions, please select 'A little' as your answer” (N = 7). Combining all exclusion criteria, 42 participants (17.95%) were excluded with N = 192 participants left for analysis (115 females (59.90%), between 18-65 ages (mean = 31.55, SD = 11.75 years). Excluded participants did not significantly differ in any of the three psychiatric dimension scores (see **Self-report psychiatric questionnaires, transdiagnostic dimensions & IQ**; all *ps* > 0.06) from participants whose data were analyzed.

Procedure. Before presenting to the lab for in-person EEG testing, participants completed a brief at-home assessment via the Internet. They provided informed electronic consent, and submitted basic demographic data (age and gender), listed any medication they were taking for a mental health issue and completed a set of 9 self-report psychiatric questionnaires (see **Self-report psychiatric questionnaires, transdiagnostic dimensions & IQ**). During the in-person EEG session, participants completed two tasks: the modified Eriksen flanker task (Eriksen and Eriksen, 1974) and the two-step reinforcement learning task (Daw et al., 2005, 2011). Data from the former task are published elsewhere (Seow et al., 2020), but note that we also reported the basic association with compulsivity and model-based planning in that paper, which served to contextualize a null result. Once participants had completed both tasks, they completed a short IQ evaluation before debriefing. A subset of the participants (N = 110, 47%) completed a short psychiatric interview (Mini International Neuropsychiatric interview English Version 7.0.0; M.I.N.I.) (Sheehan et al., 1998) before the experimental tasks to establish their diagnostic status.

Disorder prevalence (M.I.N.I.). After exclusion, 80 participants (41.67%) completed the M.I.N.I., which was introduced part-way through the study to add additional clinical context above our self-report measures. Of these participants, 35 (43.75%) met the criteria for one or more disorder. Broken down by recruitment arm, all 7 subjects (100%) recruited from the clinical setting met criteria, while 28 (38.36%) from university channels met criteria. This rate is close to published reports on the prevalence of mental health disorders in college student samples (Auerbach et al., 2018; Evans et al., 2018). Of the total sample, 33 (17.19%) were currently medicated for a mental health issue. Broken down by recruitment arm, all individuals recruited from the clinic were medicated, while 26 (14.05%) of those recruited through normal channels were medicated.

Two-step reinforcement learning task. The sequence of events was presented in the same manner as a prior study that conducted the two-step task in the EEG (Eppinger et al., 2017) with the exception that we used the standard 70/30% transition probabilities (whereas Eppinger et al. (2017) instead contrasted blocks of 60/40% vs 80/20%) and had a slightly shorted time to make a choice (1500ms here versus 2000ms in their paper) (***Figure 1***). On each trial, participants were first presented with a fixation cross for 500ms, and then shown a choice between two spaceships. They had 1500ms to respond; after which, an outline over the chosen option would indicate their choice (feedback) for 500ms. A fixation cross was shown for 500ms before transition, where the transitioned planet was shown (a blank color block) for 1000ms. Two aliens of that particular planet would then appear, with 1500ms for choice, with feedback of the chosen option subsequently shown for 500ms. Each of the aliens led to a probabilistic reward with a picture of ‘space treasure', or no reward with 'space dust', that was presented for 1000ms. Responses were indicated using the left ('Q') and right ('P') keys. Color of blocks behind rockets and those representing planets were randomized across all participants. Participants performed two blocks of 75 trials, i.e., 150 trials in total.

The task captures both model-based and model-free behavior. A participant who performs the task purely in a model-free way will make their first-stage choices solely on whether they were rewarded on the last trial (choosing the same option if rewarded previously), regardless of the transition type that occurred. In contrast, a model-based strategy will take into account both the history of reward and the transition structure of the task when making the first-stage choice. For instance, if a first-stage choice led to a rewarded second-stage option via a rare transition, a model-based learner would be more likely to choose the alternative first-stage choice on the next trial as a common transition would then lead to the previously rewarded second-stage option. However, a model-free learner would *not* make this adjustment in choice based on transition type, and instead repeat the same first-stage choice again.

Prior to the experimental task, participants completed a tutorial that explained the key concepts of the paradigm; the probabilistic association between the aliens and rewards (10 trials) and the probabilistic transition structure of rockets to planets (10 trials). After this practice phase, they had to answer a 3-item basic comprehension test regarding the key rules of the task. If participants failed to answer all questions correctly, the experimenter would reiterate the key concepts of the paradigm to the participant, allowing clarification.

Self-report psychiatric questionnaires, transdiagnostic dimensions & IQ. In order to quantify compulsivity in our sample, we applied a previously defined transdiagnostic definition (Gillan et al., 2016) that is based on a weighted combination of items drawn from 9 self-report questionnaires (which were fully randomized). The questionnaires utilized were the Alcohol Use Disorder Identification Test (AUDIT) to asses *alcohol addiction* (Saunders et al., 1993), the Apathy Evaluation Scale (AES) for *apathy* (Marin et al., 1991), the Self-Rating Depression Scale (SDS) for *depression* (Zung, 1965), the Eating Attitudes Test (EAT-26) for *eating disorders* (Garner et al., 1982), the Barratt Impulsivity Scale (BIS-11) for *impulsivity* (Patton et al., 1995), the Obsessive-Compulsive Inventory - Revised (OCI-R) for *OCD* (Foa et al., 2002), the Short Scales for Measuring Schizotypy (SSMS) for *schizotypy* (Mason et al., 2005), the Liebowitz Social Anxiety Scale (LSAS) for *social anxiety* (Liebowitz, 1987), the trait portion of the State-Trait Anxiety Inventory (STAI) for *trait anxiety* (Spielberger et al., 1983). The short IQ evaluation was the International Cognitive Ability Resource (I-CAR) (Condon and Revelle, 2014). Questionnaires were fully randomized in their presentation. Correlations between questionnaire total scores ranged greatly (r = -0.08 to 0.79).

We used weights derived from a previous study (Gillan et al., 2016) to transform our scores as our sample size had too low a subject-to-variable ratio (N = 192) for *de novo* factor analysis, as compared to the original study (N = 1413). Prior studies have demonstrated the stability of the factor structure in new data (Rouault et al., 2018; Seow and Gillan, 2020). Consistent with prior work, the resulting dimension scores were moderately intercorrelated (r = 0.33 to 0.42).

Behavioral data pre-processing. Individual missed trials and trials with very fast (<150ms) reaction times at the first-stage (indicating inattention or poor responding) were excluded from analyses. A total of 1082 trials (3.76%) were removed across participants (per participant mean = 5.64 (3.76%) trials).

Quantifying model-based planning. The extent to which participants exhibited model-based (goal-directed) behavior was estimated from the stay/switch behavior of the first-stage choice (see **Two-step reinforcement learning task**) using mixed-effects models written in R, version 3.6.0 via RStudio version 1.2.1335 (http://cran.us.r-project.org; RRID:SCR_001905) with the *glmer()* function from the *lme4* package (RRID:SCR_015654), with Bound Optimization by Quadratic Approximation (bobyqa) with 1e5 functional evaluations. The basic model tested if participants' choice behaviour to *Stay* (repeat a choice they made on the last trial) (stay: 1, switch: 0) was influenced by the previous trial's *Reward* (rewarded: 1, unrewarded: -1), *Transition* (common (70%): 1, rare (30%): -1) and their interaction (***Figure 1***). Within-subject factors (the intercept, main effects of reward, transition, and their interaction) were taken as random effects (i.e., allowed to vary across participants). In R syntax, the model was: Stay ~ Reward * Transition + (Reward * Transition + 1 | Subject).

As a model-based strategy depends on the history of reward and the transition structure, the extent to which model-based planning contributed to choice was indicated by the presence of a significant interaction effect between Reward and Transition (MB). Split half-reliability, where the data were split into two subsets (even versus odd trials) and correlated and adjusted with Spearman-Brown prediction formula, was estimated for model-based planning. To test if the compulsive dimension was associated with model-based deficits, we included the total scores of all three dimensions (*AD*: anxious-depression, *CIT*: compulsive behavior and intrusive thought, *SW*: social withdrawal) as z-scored fixed effect predictors into the basic model described above. The extent to which compulsivity is related to deficits in model-based planning was indicated by the presence of a significant negative Reward*Transition**CIT* interaction.

Sensitivity to task structure: Reaction time (RT). Recent work has shown that one effective way to index an individual's sensitivity to the structure of the task is via reaction times (RT) (Shahar et al., 2019). In a similar fashion, we conducted a mixed effect linear regression of transition type (*Transition*; common: -1, rare: 1) on second-stage reaction time (S2-RT). In the syntax of R with the *lmer()* function and *lmerTest* package for statistical tests (RRID: SCR_015656) (as with for all subsequent mixed effect models), the model was: S2-RT ~ Transition + (Transition + 1 | Subject). We asked if compulsivity was associated with a reduction in RT sensitivity to the transition structure (*RT-Trans*) with an interaction of the total scores of the three dimensions (*AD, CIT, SW*) as z-scored fixed effect predictors into the original model above; indicated by the presence of a significant negative Transition*CIT interaction.

We report the standardized beta coefficients and standard errors (applicable for all subsequent regression analyses).

EEG recording & pre-processing. EEG was recorded continuously using an ActiveTwo system (BioSemi, The Netherlands) from 128 scalp electrodes and digitized at 512 Hz. The data were processed offline using EEGLab (Delorme and Makeig, 2004) (RRID:SCR_007292) version 14.1.2 in MATLAB R2018a (The MathWorks, Natick, MA) (RRID:SCR_001622). Data were imported using A1 as a reference electrode, then down-sampled to 250 Hz and band-pass filtered between 0.05 and 45 Hz. Bad channels were rejected with a criterion of 80% minimum channel correlation. All removed channels were interpolated, and the data were re-referenced to the average. To remove ocular and other non-EEG artefacts, ICA was run on continuous data with runica, *PCA* option on, and its components were rejected automatically with Multiple Artifact Rejection Algorithm (MARA) (Winkler et al., 2011), an EEGLab toolbox plug-in, at a conservative criterion of >90% artefact probability. For all EEG analyses, other non-specific artefacts were removed after epoching using a criterion of any relevant electrode examined showing a voltage value exceeding ±100μV. If participants had a rate of >70% of total epochs failing this criterion, their data were excluded from all analyses (N = 4 as reported in **Participant exclusion criteria**). The remaining participants had mean = 147.46 (SD = 2.98) epochs left.

Single-trial analyses with EEG signals. All analyses described below (including time-frequency single-trial analyses) were conducted with mixed effects models. For every single-trial analysis, we excluded single-trial EEG estimates which were ±5 SD away from the mean of the group. A maximum of <0.79% (n = 215) of the total trials across all participants were excluded for any measure. The regression model-based estimate (MB; defined in **Quantifying model-based planning**) was used as the individual between-subjects model-based estimate in all EEG analyses.

Sensitivity to task structure: P300 and transition type. The P300 has well-established sensitivity to stimulus probability (Polich and Margala, 1997) and prior research in healthy humans hypothesized the P300 as a sensitivity marker of state transition knowledge on the two-step task, although the direction of the reported effects have varied (Eppinger et al., 2017; Sambrook et al., 2018; Shahnazian et al., 2019). Likewise, here we sought to investigate if the P300 would be sensitive to individual subjects' sensitivity to transition structure and if the effect were linked to model-based planning/compulsivity.

We first measured the P300 component at four parietal electrodes over the topography of the stimulus-locked peak (D16 (CP1), A3 (CPz), B2 (CP2), A4); ***Figure 3A***). Data were epoched from -500ms to 1700ms relative to the onset of the second-stage stimulus (aliens presented) and baselined corrected from -200ms to 0ms. Stimulus-locked single-trial P300 amplitudes were estimated as the mean of ±100ms around the individual's averaged latency of their positive peak within a search window 250ms to 1000ms after stimulus onset. To eliminate amplitude biases owing to latency variances due to RT, we subsequently aligned the epochs (measured at A4, A5, A19 (Pz), A32, the response-locked peak; ***Figure 3B***) to the time of choice response execution. The response-locked P300 amplitude was quantified as the mean amplitude -100ms to 0ms before response. We also measured the build-up rate of the response-locked signal as the slope of a straight line fitted to each single-trial waveform using the interval -400ms to -200ms. To investigate if the P300 was sensitive to rare versus common transitions and whether this depended on model-based control/compulsivity, we regressed both stimulus- and response-locked P300 measures against transition type *Transition*: rare: 1, common: 0) interacting with z-scored model-based estimates (*MB*) or compulsivity (*CIT*, controlled for the other psychiatric dimensions *AD* and *SW*), taking *Transition* and the intercept as random effects. In R syntax, the models were EEG ~ Transition*MB + (Transition + 1 | Subj) and EEG ~ Transition*(CIT + AD + SW) + (Transition + 1| Subj) respectively.

Time-frequency analysis. EEG data were epoched for both first and second-stages of the task for time-frequency analyses (alpha (9-13Hz) and theta (4-8Hz) power) detailed in the subsequent sections: -1700ms to 2200ms stimulus-locked at the first-stage (rockets) as well as -2000ms to 3500ms stimulus-locked at second-stage (aliens). Time-frequency calculations were computed using custom-written MATLAB (The MathWorks, Natick, MA) routines. The EEG time series in each epoch was convolved with a set of complex Morlet wavelets, defined as a Gaussian-windowed complex sine wave: $e^{(-i2*time*f)}$ $e^{(-time^{\wedge}2/2\sigma^{\wedge}2\ )}$ where *i* is the complex operator, *time* is time, *f* is frequency, which increased from 2 to 40 Hz in 40 logarithmically spaced steps. *σ* defines the cycle (or width) of each frequency band and was set to cycle/2πf, where cycle increased from 4 to 12 in 40 logarithmically spaced steps in accordance with each increase in frequency step. The variable number of cycles leverages the temporal precision at lower frequencies and increases frequency precision at higher frequencies. From the resulting complex signals of every epoch, we extracted estimates of power. Power is defined as the modulus of the resulting complex signal: $Z(\mathit{time})$ (power time series: $\rho(\mathit{time}) = \mathrm{real}[z(\mathit{time})]^2 + \mathrm{imag}[z(\mathit{time})]^2$).

Stimulus-locked first-stage epoch was baselined corrected to the average frequency power for each frequency band examined (i.e., alpha or theta) from -400ms to -100ms (corresponding to first-stage fixation) while for stimulus-locked second-stage epoch used -1400ms to -1100ms (corresponding to second-stage fixation, *before* presentation of the coloured squares (i.e., planets)) as the baseline. The latter baseline window was chosen as the colour of the planets were predictive of the aliens; as such, choice-relevant neural activity may potentially merge in the interval between the onset of the planets and aliens. For single-trial estimates of frequency power, as baselining with division induces spurious power fluctuations due to trial-to-trial fluctuations, power at each individual trial was baselined corrected with the linear subtraction method (Cohen, 2014) with its corresponding baseline activity: (power(*time*) – power(*baseline*)), at each frequency, at each channel. For visualisation purposes in the figures presented, power was normalized by conversion to a decibel (dB) scale: (10*log10[power(*time*)/power(*baseline*)]).

Sensitivity to task structure: Alpha power and transition type. We were also interested in the idea that more sustained post-planning processes may be important for explaining model-based deficits in compulsive individuals. As such, we focused on posterior alpha-band (9-13Hz), which in addition to reflecting surprising outcomes (Fouragnan et al., 2017), is considered a general marker of mental activity and attention (Laufs et al., 2003; Klimesch, 2012) and is suppressed in conditions where increased mental effort is needed (Stipacek et al., 2003; Pesonen et al., 2007). Much like the P300, we hypothesized that in model-based planners, alpha power would be greater following rare transitions. Potentially reflecting more than just an acute surprise, we predicted that alpha would show a more sustained pattern of increased suppression on rare versus common trials; speculatively, the sort that might be required to correctly update the (alternative) top stage choices following reward receipt on those rare trials. As a putatively core constituent of model-based planning, we hypothesized that the degree of this alpha sensitivity to transition type would be associated with individual differences in model-based choice. Moreover, if individuals high in compulsivity have an impoverished model of the task, we predicted they would show reduced alpha sensitivity to the transition types.

Alpha power was measured at five parietal-occipital electrodes (A18, A19 (Pz), A20, A21, A31; surrounding A20 electrode; ***Figure 8A***) in an epoch centred on the onset of the second-stage stimuli (aliens) and baseline corrected with activity before the onset of the transition (planets) (see **Time-frequency analysis**). Single-trial stimulus-locked alpha power estimates were measured as the mean power ±250ms around the average latency of the negative peak, specific for each individual, found within a search window 0ms to 1000ms after stimulus (alien) onset. We additionally obtained alpha power estimates quantified across four 1000ms rolling time bins by the mean amplitude within each time window. We labelled the time bins as they began from transition (planet presentation) to the stimuli (aliens presentation) from 0ms to 1000ms, followed by three windows spanning choice to reward from 1000ms to 2000ms, 2000ms to 3000ms, and 3000ms to 4000ms. The same approach of mixed effect models with **P300 and transition type** was used to examine the influence of model-based estimates/compulsivity on alpha power representation of rare versus common transitions, except for where *Transition* was coded differently (rare: -1, common: 1) for ease of interpreting the direction of interaction effects.

Cognitive control: theta power during choice. Mid-frontal theta (4-8Hz) power is a well-established EEG signature of exerting 'cognitive control' over lower level impulses (Sauseng et al., 2010; Cavanagh and Frank, 2014), including Pavlovian biases (Cavanagh et al., 2013). We therefore considered theta power as a candidate signature associated with implementing model-based decisions and overriding more habitual model-free choices. If deficits in model-based planning in high compulsive individuals arise due to a failure of implementation, theta power during choice would be negatively linked to compulsivity.

For theta power (4-8Hz), power estimates were measured at four frontal midline electrodes (C21 (Fz), C22, C23 (FCz), A1 (Cz); see ***Figure 8B***) at the first-stage (see **Time-frequency analysis**). The mean power ±250ms around the individual's average latency of the positive peak found within a search window 0ms to 500ms after stimulus onset was taken for every epoch. Similar to preceding analyses, we tested if single-trial theta power at the time of first-stage choice was associated with individual differences in model-based choice (*MB*), RT sensitivity to the transition structure (*RT-Trans*) or to compulsivity (*CIT*, controlled for *AD* and *SW*). We did this by taking each of them as z-scored predictor variables in their own linear regression model of trial by trial theta power using the following notation in R, which allows for a random intercept for each subject: S1-Theta ~ predictor variable + (1 | Subject). We also carried out a *post-hoc* analysis to test if theta modulates participants' trial-by-trial RT (*S2-RT)* sensitivity to transition (*Transition*; common: 1, rare: -1) by testing a model of S2-RT ~ Transition * S1-Theta + (Transition * S1-Theta + 1 | Subject).

Specificity with psychiatric questionnaire scores versus transdiagnostic dimensions. Additionally, we examined the advantages of utilizing a transdiagnostic definition of compulsivity as opposed to investigating single psychiatric questionnaires. We repeated the above time-frequency analyses (alpha and theta) with the individual total questionnaire scores (*QuestionnaireScore*, z-scored) replacing the three psychiatric dimensions (*CIT, AD, SW*) in their respective regression models detailed above. Separate mixed effects regression models were performed for each individual questionnaire as correlation across questionnaire scores ranged greatly from r = -0.09 to 0.79 as opposed to the transdiagnostic analysis where all three dimensions (that correlated moderately: r = 0.33 to 0.42) were included in the same model.

Supplemental analyses. Finally, to ensure the specificity of any observed effects to the task events outlined above, we also tested for an association between model-based planning and compulsivity with our candidate EEG signatures in reverse. That is, we tested if between model-based planning and compulsivity were linked to (i) alpha power at the first-stage or (ii) theta power sensitivity to transition type at the second-stage. See ***Figure 8A*** and ***Figure 8B*** for the respective analyses.

For all analyses, we report the standardized beta coefficients and standard errors.

Code and data availability. The code and data to reproduce the main figures are available at https://osf.io/mx9kf/.

Results

Compulsivity and model-based planning. Logistic regression analysis of choice behavior on the two-step task revealed clear evidence for model-based planning in this sample via a significant interaction between Reward and Transition (β = 0.20, *standard error (SE)* = 0.03, p < 0.001; ***Figure 1***). Individual subject coefficients for this interaction term were extracted and used as an individual difference measure for EEG analysis (split half-reliability was r = 0.71). Consistent with prior work, there was also evidence for model-free learning, where subjects were more likely to repeat choices if they were followed by reward (main effect of Reward: β = 0.55, *SE* = 0.05, p < 0.001), and an overall tendency to repeat choices from one trial to the next (Intercept: β = 1.46, *SE* = 0.07, p < 0.001). Importantly, we replicated prior work in finding that individual differences in compulsivity and intrusive thought (hereafter: 'compulsivity') were associated with reduced model-based planning (β = -0.07, *SE* = 0.04, p = 0.05) (***Figure 2A***), while anxious-depression (β = 0.05, *SE* = 0.04, p = 0.14) and social withdrawal were not (β = -0.01, *SE* = 0.04, p = 0.73).

Reaction time (RT) sensitivity to task structure. Someone who is aware of the task structure should expect to be presented with the second-stage state that is most commonly associated with their first-stage choice. As such, when a rare transition occurs, they will react to this violation in expectancy, requiring more time to respond and 're-plan' (Decker et al., 2016). We therefore hypothesized that participants would have a slower RT after a rare versus common transition and that this difference would be greater in participants who exhibit the most model-based behavior. We found support for both hypotheses; participants had a slower mean RT for rare versus common trials after transition (β = 0.17, *SE* = 0.01, p < 0.001) (***Figure 2B***) and this effect was larger in those with higher levels of model-based control (β = 0.28, *SE* = 0.07, *p* < 0.001). Crucially, we found that this effect was reduced in high compulsive individuals (β = -0.03, *SE* = 0.01, *p* = 0.01) (***Figure 2C***). Prior studies using this task did not test for an association between compulsivity and this RT cost, but the data is readily available. To test the robustness of this finding, we therefore re-analyzed a prior dataset (N = 1413) collected entirely online (Gillan et al., 2016) using a similar variant of the two-step task and the same measure of compulsivity. We replicated this effect (β = -0.02, *SE* = 0.004, *p* < 0.001) (***Figure 2C***). This is, to our knowledge, the first evidence that compulsivity is associated with muted behavioral reactions to violations in transition expectancy, suggestive of disruption in the quality of the mental model of the task itself.

P300 sensitivity to task structure. The P300 or P3b has well-established sensitivity to stimulus probability, exhibiting larger peak amplitudes for less probable stimuli (Polich and Margala, 1997). Prior research in healthy humans thus hypothesized that the P300 may be a marker of sensitivity to state transitions on the two-step task, though these studies have yielded inconsistent results, with some finding greater P300 amplitudes for rare versus common transitions (Sambrook et al., 2018; Shahnazian et al., 2019) and one finding the opposite (Eppinger et al., 2017). Here, we examined the second-stage stimulus-locked P300 and found a significant main effect of transition type (β = 0.03, *SE* = 0.01, *p* = 0.02), consistent with Sambrook et al. (2018) and Shahnazian et al. (2019) whereby greater P300 amplitude was observed after rare versus common transitions (***Figure 3A***). However, this differential rare versus common signal was not larger in individuals high in model-based planning (β = 0.01, *SE* = 0.01, p = 0.35), nor did it show any association to compulsivity (β = 0.02, *SE* = 0.02, p = 0.24).

Recently, it has been suggested that P300 is more accurately characterized as a response-locked signal (O'Connell et al., 2012; Twomey et al., 2015). This raises the possibility that the stimulus-locked signal measurements favored in previous studies of the two-step task may have yielded cross-condition effects that were partly or entirely determined by RT differences. In light of these considerations, we complemented the stimulus-locked analyses with an examination of response-locked signal measurements. When we repeated the analysis using response-locked P300 amplitude, we found that the transition effect was no longer significant and its direction was in fact reversed (β = -0.02, *SE* = 0.01, p = 0.23) (***Figure 3B***). Again, there was no association with model-based planning (β = -0.01, *SE* = 0.01, p = 0.49) or compulsivity (β = 0.01, *SE* = 0.02, p = 0.67). We also examined the build-up rate of the response-locked P300 as a measure of how quickly evidence for the decision was accumulated (Kelly and O'Connell, 2013). The build-up rate was steeper for common versus rare trials (β = -0.04, *SE* = 0.01, p = 0.002) but this measure was again not linked to model-based planning (β = -0.01, *SE* = 0.01, p = 0.46) nor compulsivity (β = 0.01, *SE* = 0.02, p = 0.25). Thus, we concluded that the P300 may not provide the most reliable or sensitive measure of neural sensitivity to task structure.

Alpha power sensitivity to task structure. ERPs principally reflect activity changes that are short-lived and strictly time-locked to particular events (Makeig and Onton, 2012). We thus investigated if time-frequency measures such as alpha power (9-13Hz), which has been previously linked to OCD (Perera et al., 2019), would be able to provide a superior and more sustained neural representation of the task's transition structure. Specifically, we examined if parietal-occipital alpha power locked to the second-stage stimulus was able to distinguish between rare and common transitions across a series of time bins in our task. This allowed us to ascertain not just if participants showed sensitivity to task structure following a transition, but for how long they sustained that representation (e.g., as they made subsequent choices and received a reward). We reasoned that short-lived responses might reflect surprise stemming from arriving at a rare versus common state, but more sustained patterns could reflect post-planning processes required to update model-based top stage choice values.

In line with our hypothesis, alpha power overall differentiated between the two transition types ($\beta = 0.02$, $SE = 0.01$, $p < 0.001$), such that parietal-occipital alpha was more suppressed after rare versus common transitions (***Figure 4A***). We found that in a manner sustained over three rolling time bins beginning from the state transition (planet) (0ms to 1000ms: $\beta = 0.02$, $SE = 0.01$, $p = 0.03$) to the end of choice feedback (1000ms to 2000ms: $\beta = 0.02$, $SE = 0.01$, $p = 0.03$; 2000ms to 3000ms: $\beta = 0.01$, $SE = 0.02$, $p < 0.05$), individuals high in model-based control showed the largest alpha power differentiation (***Figure 4B***). Importantly, this same signature was negatively related to compulsivity, with a significant association observed at the time after state transition (0ms to 1000ms: $\beta = -0.03$, $SE = 0.01$, $p = 0.007$) (***Figure 4C***). Overall second-stage alpha power was also associated with compulsivity ($\beta = -0.09$, $SE = 0.03$, $p < 0.001$), however, this effect was not related to model-based control (β = 0.03, *SE* = 0.02, *p* = 0.25) nor RT differences in transition types (β = -0.03, *SE* = 0.02, *p* = 0.20)—highlighting that it is the sensitivity of alpha to task structure, not alpha overall, that best tracks model-based performance at this task.

Control analyses demonstrate that this transition sensitivity effect is present even if alpha estimates were locked to response times (***Figure 5***) and is not found with second-stage theta-power (***Figure 8B***) (which we examine later in the context of cognitive control at first-stage). In terms of alpha specificity to compulsivity, there were no associations to the other two transdiagnostic dimensions; anxious-depression (β = 0.007, *Ss* = 0.01, *p* = 0.47) or social withdrawal (β = -0.001, *SE* = 0.01, *p* = 0.91). When we examined the association between alpha-band sensitivity to transition structure and all nine of the original psychiatric questionnaire total scores, we found diminished sensitivity in those with elevated OCD (β = -0.02, *SE* = 0.01, *p* = 0.006) and eating disorder symptoms (β = -0.02, *SE* = 0.01, *p* = 0.05) (***Figure 6***).

Theta power at the time of choice. Finally, moving beyond participants' sensitivity to the transition structure of the task, we tested if during the crucial time of first-stage choice, when model-based planning manifests in behavior, we could detect differences in a neural signature previously linked to cognitive control, mid-frontal theta (4-8Hz). As theta has previously been shown to reflect computations crucial to goal-directed action (Sauseng et al., 2010; Cavanagh et al., 2013; Cavanagh and Frank, 2014), we hypothesized that model-based planning would be positively linked to theta power while compulsivity would be negatively associated with the neural oscillation.

We tested this using a mixed effects regression analysis with trial-by-trial estimates of theta power as the dependent variable and individual differences in model-based choice (coefficients of the effect of reward*transition from the logistic regression of stay/switch behavior) as the predictor variable. Theta power during choice was not significantly associated with model-based planning (β = 0.02, *SE* = 0.01, *p* = 0.11); though, the trend was in the expected direction. When we used RT sensitivity to transition structure, instead of model-based choice, as an alternative manifest variable of the brain's capacity for model-based planning, we found a significant positive relationship with theta (β = 0.04, *SE* = 0.01, *p* = 0.002), indicating that those participants who had higher theta power during their first-stage choice also had larger differences in their RT between rare and common transitions at the second-stage. Finally, using the same analysis approach, this time with individual differences in compulsivity as the predictor variable, we found an overall effect of lower theta at the time of choice in individuals high in compulsivity (β = -0.03, *SE* = 0.01, *p* = 0.04) (***Figure 7A***). Similar to alpha power modulations, reduced theta power at the first-stage was linked to more than one questionnaire score—schizotypy (β = -0.03, *SE* = 0.01, *p* = 0.01), depression (β = -0.03, *SE* = 0.01, *p* = 0.02) and OCD (β = -0.03, *SE* = 0.01, *p* = 0.03) and were associated with the compulsive dimension (β = -0.03, *SE* = 0.01, *p* = 0.03) (***Figure 7B***).

One explanation for the somewhat closer association between theta and RT sensitivity (compared to model-based choice) is that theta at the time of choice reflects participants' mental simulation of future states. We tested this *post-hoc* using a within-subject analysis by examining whether on trials where theta was highest, subjects showed even greater RT sensitivity to transition type. We did not find evidence in support of this within-subject, such that the interaction between theta and transition type was not significant ($\beta = 0.004$, $SE = 0.01$, $p = 0.57$). Finally, by way of control analysis, we tested if alpha power at first-stage (***Figure 8A***) was associated with compulsivity ($\beta = -0.14$, $SE = 0.05$, $p = 0.002$), model-based planning ($\beta = 0.03$, $SE = 0.04$, $p = 0.45$) or RT differences in transition types ($\beta = -0.004$, $SE = 0.04$, $p = 0.92$), but none were significant.

Discussion

Model-based planning deficits linked to compulsivity have been theorized to arise from issues with the balance/arbitration between competing model-based and model-free influences during choice (Gillan and Robbins, 2014; Lee et al., 2014; Gruner et al., 2016; Lloyd and Dayan, 2019), but these presumed planning failures might, at least partially, arise from an impoverished internal model of task structure. Here, we found that high compulsive individuals lacked neural and behavioral sensitivity to state transition probabilities, evidenced in their RT and parietal-occipital alpha power suppression in response to unexpected transitions. Speaking to the potential for more general cognitive control problems to also contribute to model-based deficits, we additionally took mid-frontal theta as its candidate neural signature and observed that high compulsive individuals had reduced theta when they made their first-stage choices. These findings have important implications for refining theories of compulsivity, which may be associated with more fundamental problems in constructing and maintaining a model of the causal structure of the environment necessary for goal-directed “model-based” control than just cognitive control failures.

In line with prior research, participants exhibited longer RTs following rare transitions which was also previously shown to relate to model-based planning (Deserno et al., 2015; Decker et al., 2016; Shahar et al., 2019). Crucially, the opposite was true of compulsivity, with the most compulsive individuals showing the smallest difference in RT between these trial types. This finding was robust—the effect replicates in a former dataset (N=1413) tested online (Gillan et al., 2016). This may reflect a number of processes including uncertainty arising from the presentation of unexpected options (Deserno et al., 2015), lower discriminability of the options presented following rare transitions (Shahar et al., 2019) or, as per our original hypothesis, a reduced awareness of the task structure (Decker et al., 2016) including action-state transitions necessary to build an accurate causal model of the world.

Moving beyond behavior, analysis of alpha power revealed a similar picture. Much like RT, alpha suppression at the second-stage was sensitive to transition probabilities, with rare than common transitions associated with greater alpha suppression, possibly reflecting the greater mental effort required after rare transitions to call to mind action values associated with the unexpected options presented. In line with this, previous studies using *n*-back paradigms have shown greater parieto-occipital alpha suppression when working memory load increases (Stipacek et al., 2003; Pesonen et al., 2007). Importantly, this mental activity was sustained beyond second-stage choice right up until reward receipt, which might reflect that one must not only re-plan, but also that task structure information is used together with trial outcome to update first-stage choices. Consistent with this interpretation, individual difference analysis demonstrated that this difference in alpha suppression had important behavioral correlates. Model-based planners showed the largest differences in alpha power between transition type, while higher levels of compulsivity were associated with less of a distinction in alpha power between transition type. Building upon the RT findings, we present neural evidence that compulsivity may be characterized by failures in representing the kind of causal action-state relations necessary to behave in a model-based manner. The notion that sustained alpha differentiation across common/rare trials reflects a post-planning process is speculative and future research should aim to distinguish this from the effects of surprise.

Our data do not exclude the possibility that compulsive individuals also face issues with *implementing* model-based planning even when they have the requisite state knowledge. Indeed, we also found that mid-frontal theta, which is thought to support adaptive cognitive control in a variety of contexts (Cavanagh et al., 2012) was reduced in compulsive individuals during first-stage choice. In addition to being negatively related to compulsivity, theta power was also elevated in those whose RT was most sensitive to task structure, and trended towards being elevated in model-based planners, supporting the view that theta activity at the time of choice at least in part reflects mental operations relevant to executing a model-based plan. However, disentangling the specific theta-driven processes is beyond the scope of our current experimental design. Theta power at choice time could reflect a host of executive processes such as selecting between competing options (including suppressing distracting stimuli) (Nigbur et al., 2011), inhibiting unhelpful associations (Cavanagh et al., 2013) and the mental simulation/search of future states (Doll et al., 2015).

Previous EEG studies of the two-step task (Eppinger et al., 2017; Sambrook et al., 2018; Shahnazian et al., 2019) showed that the P300 was associated with state transitions. However, the inconsistent effect direction raises doubt as to how these differences should be interpreted. Recent literature conceptualizes the P300 as an evidence accumulation process that builds towards a peak at choice time (Twomey et al., 2015) and as such variances in RT will influence the latency of the stimulus-locked P300 amplitude peak (Kelly and O’Connell, 2015). Our results comparing stimulus-locked and response-locked analysis approaches suggest that it is the build-up rate of the P300 that is sensitive to transitions and that previously reported stimulus-locked amplitude modulations are attributable to RT differences. We also found that none of the analyzed P300 metrics were predictive of individual differences in model-based planning.

In this study, we utilized a transdiagnostic compulsive dimension which was previously shown to provide the best mapping to model-based deficits in an online general population sample (Gillan et al., 2016). We replicated this finding here and extend it to EEG correlates of behavior, where our alpha and theta modulations were relatively non-specific to the DSM-defined questionnaires compared to our *a priori* dimensional factor 'compulsivity'. This research pipeline illustrates how mental health dimensions may be defined in large online samples and then used in smaller studies that can avail of the harder tools of neuroscience, like EEG (Gillan and Seow, 2020). While continued over-arching criticism will question its applicability to diagnosed patients, recent work suggests the core mechanisms we capture in general population samples is broadly equivalent, at least in compulsivity. Similar to large-scale general population findings, model-based deficits in diagnosed patients are linked to individual differences in self-reported compulsivity, and model-based deficits mapped onto this compulsivity dimension more strongly than the diagnosis of these patients (e.g., whether they had an OCD diagnosis) (Gillan et al., 2019). As such, there is growing evidence that the specific associations between cognition and compulsivity observed in the general population are likely clinically relevant.

Overall, our findings suggest that model-based difficulties in compulsivity may be linked to an impoverished mental model of environmental contingency—an interpretation bolstered by recent findings implicating diminished transition learning in compulsivity in a task devoid of value representations (Sharp et al., 2020). Future work should carry on in this vein, perhaps asking: are failures in memory encoding or retrieval are responsible for the deficits observed in compulsivity following transitions? Are these effects specific to learning about actions and their consequences, or more distributed failures to learn about causality? Moreover, there are several facets of model-based planning beyond the learning/maintaining knowledge of the transition structure that may also be implicated, like the inhibition of opposing model-free signals at choice time, forward simulation of future states at choice time, attention at reward receipt and using that information for updating the correct first-stage option. Understanding these factors will provide a clearer picture of the neural mechanisms that lead to compulsive disorders and hopefully, provide scope for intervening more effectively. The clear advantage of the use of EEG here is its temporal resolution, which was crucial in allowing us to capture the sustained differentiation of alpha power to transitions. With this, of course, comes with a lack of spatial precision. Future work combining fMRI and EEG might prove fruitful, particularly for dissecting potentially multiple processes at the time of first-stage choice. Finally, there is growing recognition that the dichotomization of two decision systems is over-simplified; model-based/model-free processes are partially synergistic, overlapping in certain situations and/or hierarchically organized (Cushman and Morris, 2015; Balleine and Dezfouli, 2019; da Silva and Hare, 2020). Future research must go beyond dichotomized frameworks to advance our mechanistic understanding of how deficits in building a model of the world translate to behavior irregularities such as compulsive habits.

Our findings may have implications for understanding how compulsive behaviors and obsessive beliefs develop in concert, in a more integrated fashion than previously considered. Clinical cognitive models of OCD have long presumed that compulsions are performed to reduce anxiety induced by obsessive beliefs (Salkovskis and McGuire, 2003; Matthews and Wells, 2008), in contrast to a more recent hypothesis suggesting that obsessions are post-hoc rationalizations to explain the performance of compulsive behavior (Gillan and Sahakian, 2015). These data may suggest that the hard distinction between obsessions and compulsions might be less clear than these models propose. Failures in accurately representing the relationship between actions and their consequences may be a common source of both compulsive habitual behaviors in OCD and also faulty metacognitive beliefs that form the basis of obsessions. One might imagine that with a less stable world model representation, the more likely a patient may develop faulty beliefs and rely on habitual representations.

References

Allen M, Poggiali D, Whitaker K, Marshall TR, Kievit RA (2019) Raincloud plots: a multi-platform tool for robust data visualization. Wellcome open Res 4:63.

Auerbach RP, Mortier P, Bruffaerts R, Alonso J, Benjet C, Cuijpers P, Demyttenaere K, Ebert DD, Green JG, Hasking P, Murray E, Nock MK, Pinder-Amaker S, Sampson NA, Stein DJ, Vilagut G, Zaslavsky AM, Kessler RC (2018) WHO world mental health surveys international college student project: Prevalence and distribution of mental disorders. J Abnorm Psychol 127:623.

Balleine BW, Dezfouli A (2019) Hierarchical action control: Adaptive collaboration between actions and habits. Front Psychol 10:2735.

Balleine BW, O'Doherty JP (2010) Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology 35:48–69.

Cavanagh JF, Eisenberg I, Guitart-Masip M, Huys Q, Frank MJ (2013) Frontal theta overrides Pavlovian learning biases. J Neurosci 33:8541–8548.

Cavanagh JF, Frank MJ (2014) Frontal theta as a mechanism for cognitive control. Trends Cogn Sci 18:414–421.

Cavanagh JF, Zambrano-Vazquez L, Allen JJB (2012) Theta lingua franca: A common mid-frontal substrate for action monitoring processes. Psychophysiology 49:220–238.

Cohen MX (2014) Analyzing neural time series data: theory and practice. MIT press.

Condon DM, Revelle W (2014) The international cognitive ability resource: Development and initial validation of a public-domain measure. Intelligence

43:52–64.

Cushman F, Morris A (2015) Habitual control of goal selection in humans. Proc Natl Acad Sci 112:13817–13822.

da Silva CF, Hare TA (2020) Humans primarily use model-based inference in the two-stage task. Nat Hum Behav:1–14.

Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. Neuron 69:1204–1215.

Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci 8:1704–1711.

Decker JH, Otto AR, Daw ND, Hartley CA (2016) From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. Psychol Sci 27:848–858.

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21.

Deserno L, Huys QJMM, Boehme R, Buchert R, Heinze H-JJ, Grace AA, Dolan RJ, Heinz A, Schlagenhauf F (2015) Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. Proc Natl Acad Sci 112:1595–1600.

Dickinson A (1985) Actions and habits: the development of behavioural autonomy. Philos Trans R Soc London B, Biol Sci 308:67–78.

Doll BB, Duncan KD, Simon DA, Shohamy D, Daw ND (2015) Model-based choices involve prospective neural activity. Nat Neurosci 18:767–772.

Eppinger B, Walter M, Li S-CC (2017) Electrophysiological correlates reflect the integration of model-based and model-free decision information. Cogn Affect Behav Neurosci 17:1–16 Available at: http://link.springer.com/10.3758/s13415-016-0487-3.

Eriksen BA, Eriksen CW (1974) Effects of noise letters upon the identification of a target letter in a nonsearch task. Percept Psychophys 16:143–149.

Ersche KD, Gillan CM, Simon Jones P, Williams GB, Ward LHE, Luijten M, De Wit S, Sahakian BJ, Bullmore ET, Robbins TW (2016) Carrots and sticks fail to change behavior in cocaine addiction. Science (80-) 352:1468–1471.

Evans TM, Bira L, Gastelum JB, Weiss LT, Vanderford NL (2018) Evidence for a mental health crisis in graduate education. Nat Biotechnol 36:282.

Foa EB, Huppert JD, Leiberg S, Langner R, Kichic R, Hajcak G, Salkovskis PM (2002) The obsessive-compulsive inventory: Development and validation of a short version. Psychol Assess 14:485–496.

Fouragnan E, Queirazza F, Retzler C, Mullinger KJ, Philiastides MG (2017) Spatiotemporal neural characterization of prediction error valence and surprise during reward learning in humans. Sci Rep 7:1–18.

Garner DM, Bohr Y, Garfinkel PE (1982) The eating attitudes test: Psychometric features and clinical correlates. Psychol Med 12:871–878.

Gillan CM, Kalanthroff E, Evans M, Weingarden HM, Jacoby RJ, Gershkovich M, Snorrason I, Campeas R, Cervoni C, Crimarco NC, Sokol Y, Garnaat SL, McLaughlin NCR, Phelps EA, Pinto A, Boisseau CL, Wilhelm S, Daw ND, Simpson HB (2019) Comparison of the association between goal-directed planning and self-reported compulsivity vs obsessive-compulsive disorder diagnosis. JAMA Psychiatry 77:77–85.

Gillan CM, Kosinski M, Whelan R, Phelps EA, Daw ND (2016) Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. Elife 5:e11305.

Gillan CM, Morein-Zamir S, Urcelay GP, Sule A, Voon V, Apergis-Schoute AM, Fineberg NA, Sahakian BJ, Robbins TW (2014) Enhanced avoidance habits in obsessive-compulsive disorder. Biol Psychiatry 75:631–638.

Gillan CM, Papmeyer M, Morein-Zamir S, Sahakian BJ, Fineberg NA, Robbins TW, De Wit S (2011) Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. Am J Psychiatry 168:718–726.

Gillan CM, Robbins TW (2014) Goal-directed learning and obsessive–compulsive disorder. Philos Trans R Soc B Biol Sci 369:20130475.

Gillan CM, Sahakian BJ (2015) Which is the driver, the obsessions or the compulsions, in OCD? Neuropsychopharmacology 40:247–248.

Gillan CM, Seow TXF (2020) Carving out new transdiagnostic dimensions for research in mental health. Biol Psychiatry Cogn Neurosci Neuroimaging 5:932–934.

Gruner P, Anticevic A, Lee D, Pittenger C (2016) Arbitration between action strategies in obsessive-compulsive disorder. Neurosci 22:188–198.

Kelly SP, O’Connell RG (2013) Internal and external influences on the rate of sensory evidence accumulation in the human brain. J Neurosci 33:19434–19441.

Kelly SP, O'Connell RG (2015) The neural processes underlying perceptual decision making in humans: Recent progress and future directions. J Physiol Paris 109:27–37.

Klimesch W (2012) Alpha-band oscillations, attention, and controlled access to stored information. Trends Cogn Sci 16:606–617.

Laufs H, Kleinschmidt A, Beyerle A, Eger E, Salek-Haddadi A, Preibisch C, Krakow K (2003) EEG-correlated fMRI of human alpha activity. Neuroimage 19:1463–1476.

Lee SW, Shimojo S, O'Doherty JP (2014) Neural computations underlying arbitration between model-based and model-free learning. Neuron 81:687–699.

Liebowitz MR (1987) Social phobia. Mod Probl Pharmapsychiatry 22:141–173.

Lloyd K, Dayan P (2019) Pavlovian-instrumental interactions in active avoidance: The bark of neutral trials. Brain Res 1713:52–61.

Makeig S, Onton J (2012) ERP Features and EEG Dynamics: An ICA Perspective. In: The Oxford Handbook of Event-Related Potential Components.

Marin RS, Biedrzycki RC, Firinciogullari S (1991) Reliability and validity of the apathy evaluation scale. Psychiatry Res 38:143–162.

Mason O, Linney Y, Claridge G (2005) Short scales for measuring schizotypy. Schizophr Res 78:293–296.

Matthews G, Wells A (2008) Rumination, Depression, and Metacognition: The S-REF Model. In: Depressive Rumination: Nature, Theory and Treatment.

Nigbur R, Ivanova G, Stürmer B (2011) Theta power as a marker for cognitive interference. Clin Neurophysiol 122:2185–2194.

O'Connell RG, Dockree PM, Kelly SP (2012) A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. Nat Neurosci 15:1729–1735.

Patton JH, Stanford MS, Barratt ES (1995) Factor structure of the Barratt impulsiveness scale. J Clin Psychol 51:768–774.

Perera MPN, Bailey NW, Herring SE, Fitzgerald PB (2019) Electrophysiology of obsessive compulsive disorder: a systematic review of the electroencephalographic literature. J Anxiety Disord 62:1–14 Available at: https://doi.org/10.1016/j.janxdis.2018.11.001.

Pesonen M, Hämäläinen H, Krause CM (2007) Brain oscillatory 4–30 Hz responses during a visual n-back memory task with varying memory load. Brain Res 1138:171–177.

Polich J, Margala C (1997) P300 and probability: comparison of oddball and single-stimulus paradigms. Int J Psychophysiol 25:169–176.

Robbins TW, Gillan CM, Smith DG, de Wit S, Ersche KD (2012) Neurocognitive endophenotypes of impulsivity and compulsivity: Towards dimensional psychiatry. Trends Cogn Sci 16:81–91.

Rouault M, Seow T, Gillan CM, Fleming SM (2018) Psychiatric symptom dimensions are associated with dissociable shifts in metacognition but not task performance. Biol Psychiatry 84:443–451.

Salkovskis PM, McGuire J (2003) Cognitive-behavioural theory of OCD. Obs

Compuls Disord Theory, Res Treat:59–78.

Sambrook TD, Hardwick B, Wills AJ, Goslin J (2018) Model-free and model-based reward prediction errors in EEG. Neuroimage 178:162–171.

Saunders JB, Aasland OG, Babor TF, De La Fuente JR, Grant M (1993) Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. Addiction 88:791–804.

Sauseng P, Griesmayr B, Freunberger R, Klimesch W (2010) Control mechanisms in working memory: a possible function of EEG theta oscillations. Neurosci Biobehav Rev 34:1015–1022.

Seow TXF, Benoit E, Dempsey C, Jennings M, Maxwell A, McDonough M, Gillan CM (2020) A dimensional investigation of error-related negativity (ERN) and self-reported psychiatric symptoms. Int J Psychophysiol 158:340–348.

Seow TXF, Gillan CM (2020) Transdiagnostic Phenotyping Reveals a Host of Metacognitive Deficits Implicated in Compulsivity. Sci Rep 10:2883 Available at: https://doi.org/10.1038/s41598-020-59646-4.

Shahar N, Hauser TU, Moutoussis M, Moran R, Keramati M, Dolan RJ (2019) Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. PLOS Comput Biol 15:e1006803.

Shahnazian D, Ribas-Fernandes JJF, Holroyd CB (2019) Electrophysiological correlates of state transition prediction errors. bioRxiv:544551.

Sharp PB, Dolan RJ, Eldar E (2020) Disrupted state transition learning as a computational marker of compulsivity and anxious arousal.

Sheehan D V., Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC (1998) The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 59:22–33.

Snorrason I, Lee HJ, de Wit S, Woods DW (2016) Are nonclinical obsessive-compulsive symptoms associated with bias toward habits? Psychiatry Res 241:221–223.

Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA (1983) Manual for the state-trait anxiety inventory. Palo Alto, CA: Consulting Psychologists Press.

Stipacek A, Grabner RH, Neuper C, Fink A, Neubauer AC (2003) Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load. Neurosci Lett 353:193–196.

Twomey DM, Murphy PR, Kelly SP, O'Connell RG (2015) The classic P300 encodes a build-to-threshold decision variable. Eur J Neurosci 42:1636–1643.

Vaghi MM, Moutoussis M, Váša F, Kievit RA, Hauser TU, Vértes PE, Shahar N, Romero-Garcia R, Kitzbichler MG, Bullmore ET (2020) Compulsivity is linked to reduced adolescent development of goal-directed control and frontostriatal functional connectivity. Proc Natl Acad Sci 117:25911–25922.

Voon V, Derbyshire K, Rück C, Irvine MA, Worbe Y, Enander J, Schreiber LRN, Gillan C, Fineberg NA, Sahakian BJ, Robbins TW, Harrison NA, Wood J, Daw ND, Dayan P, Grant JE, Bullmore ET (2015) Disorders of compulsivity: A common bias towards learning habits. Mol Psychiatry 20:345.

Watson P, de Wit S (2018) Current limits of experimental research into habits and future directions. Curr Opin Behav Sci 20:33–39.

Winkler I, Haufe S, Tangermann M (2011) Automatic classification of artifactual ICA-components for artifact removal in EEG signals. Behav Brain Funct 7:30.

Zung WW (1965) A self rating depression scale. Arch Gen Psychiatry 12:63–70.

Figure Legends

Figure 1. Two-step reinforcement learning task. *Paradigm consists of two stages where participants take a rocket that has a common (70%) or rare (30%) transition to one of two second-stage planets (states). Aliens on these planets each have a unique probability of reward ('space treasure' (reward) or 'space dust' (non-reward)) that drifts slowly throughout the entire experiment. Participants have to take into consideration the task transition structure and their history of rewards to make choices that maximize reward. The sequence of events as presented for EEG is the same as that of Eppinger et al. (2017), except they included a manipulation of transition probabilities in their study (comparing 60/40% to 80/20%) and used a longer choice window (2000ms). On the top right inset, model-based behavior is reflected as the probability of repeating the first-stage choice (stay) as a function of the occurrence of a transition from the previous trial (common: 70%, rare: 30%) and whether a reward was received (reward, non-reward). In a purely model-free learner, stay probabilities after reward should be higher than when no reward was presented regardless of transition type. In a purely model-based learner, stay probabilities after common-reward and rare-non reward should be higher than common-non reward and rare-reward. In our empirical data here, the stay probabilities obtained across conditions is a mix of both model-based and model-free behavior. Error bars reflect standard errors of mean.*

Figure 2. Model-based behavior and reaction times in compulsivity.

*(**A**) Model-based control estimated by a logistic regression of choice behavior with one-trial back reward and transition. Regressions were conducted in a model with all three dimensions: 'anxious-depression' (AD), 'compulsivity and intrusive thought' (CIT) and 'social withdrawal' (SW). Model-based control is reduced high compulsive individuals.*

*(**B**) Participants have on average a longer mean response time (RT) at second-stage choice after a rare transition than a common one (paired t-test: t_{191}=16.16, 95% Confidence Interval (CI) [79.85 102.05], p<.001). Circles in raincloud plot* (Allen et al., 2019) *depict mean RT of rare or common trials for each individual, with black marker indicating grand average mean and standard deviation (SD).*

*(**C**) RT difference between transition type (RT-Trans) is diminished in high compulsive individuals. We replicated the same effect in a prior dataset of N = 1413* (Gillan et al., 2016)*.*

*For (**A**) & (**C**), error bars denote standard error. The Y-axes indicate the percentage change in model-based planning/RT-Trans as a function of 1 SD of psychiatric dimension scores. *p≤.05, ***p<.001.*

Figure 3. Second-stage P300 and transition type.

*(**A**) Grand average waveforms of rare and common trials stimulus-locked to second-stage stimuli (aliens). Waveform is baselined -200ms to 0ms. The mean amplitude for stimulus-locked P300 was obtained over 4 centro-parietal electrodes (D16 (CP1), A3 (CPz), B2 (CP2), A4) as indicated by the white dots in the topography plot. This transition effect was no longer significant when second-stage P300 signal was response-locked (**Figure 3B**).*

***(B)** Topography plot represents the P300 component -100ms to 0ms before second-stage response. White dots indicate parietal electrode sites (A4, A5, A19 (Pz), A32) where the positive component was measured. Grand average second-stage P300 is plotted response-locked comparing the waveforms following rare versus common transitions. Single-trial analyses indicate that the P300 amplitude, measured as the mean amplitude -100ms to 0ms (shaded grey), does not distinguish transition type ($\beta = -0.02$, $SE = 0.01$, $p = 0.23$).*

Figure 4. Stimulus-locked alpha power at transition. *Alpha power was measured across 4 time bins of 1000ms each separated by vertical dashed lines, starting from the transition (0ms) until after reward (4000ms), at parietal-occipital electrode sites (see Figure 4-2).*

*(**A**) Grand average second-stage alpha power waveforms between rare and common transitions. Continuous analyses revealed that alpha difference (rare – common) is significant in time bin 2-3 (all β>.03, SE<.01, p<.001).*

*(**B**) Alpha power difference between transitions (common minus rare) is depicted above by comparing top/bottom* 50^{th} *percentile (N=96 per group) of participants grouped by model-based estimates (MB). Continuous analyses revealed that alpha difference (rare – common) is enhanced for more model-based participants in time bins 1-3 (all β>.01, SE<.02, p<.05).*

*(**C**) Alpha power difference between transitions (common minus rare) comparing top/bottom* 50^{th} *percentile (N=96 per group) of participants grouped by or compulsivity (CIT). Continuous analyses revealed that alpha difference (rare – common) is diminished for more compulsive participants in time bin 1 (β=-.03, SE=.01, p=.007).*

*Stars in time bins indicate significance from continuous analyses. *p<.05, **p<.01, **p<.001.*

*These second-stage transition effects were specific to alpha power, and not present with theta power (**Figure 8B**).*

Figure 5. Grand average waveforms of rare versus common transitions for second-stage response-locked alpha power. *RT differences between rare and common transitions was only significantly associated with stimulus-locked alpha power differentiation of states in the time bin before reward presentation (2000ms to 3000ms: $\beta = 0.01$, SE = 0.01, $p = 0.04$; all other time bins: ps > 0.30;* ***Figure 4****). To complement our main result based on stimulus-locked alpha, we repeated the transition analysis with single-trial response-locked alpha estimates (measured as the mean of ±100ms centered around each participant's averaged latency of the negative peak), which also yielded a significant association overall effect ($\beta = 0.03$, SE = 0.01, $p < 0.001$;* ***Figure 5****). Similar to stimulus-locked alpha, rare transitions showed greater depression of alpha during choice selection for rare versus common transitions, suggesting that alpha transition effect is not explained by RT.*

Figure 6. Second-stage alpha power sensitivity to transition at time bin 1 (0ms to 1000ms) with psychiatric symptoms and dimensions (AD: 'anxious-depression', CIT: 'compulsive behavior and intrusive thought', SW: 'social withdrawal'). *Alpha power differentiating rare versus common transitions was associated with both OCD and eating disorder symptoms. The transdiagnostic analysis showed the effect was captured by a compulsive dimension (CIT). The Y-axis show the percentage change in alpha power sensitivity to transition type (%) as a function of 1 SD increase of psychiatric questionnaire/dimension scores. Error bars denote standard errors. *p ≤ .05, **p<.01.*

Figure 7. First-stage theta power with psychiatric symptoms and dimensions (AD: 'anxious-depression', CIT: 'compulsive behavior and intrusive thought', SW: 'social withdrawal'). *Theta power was measured at mid-frontal electrode sites (see Figure 6-1).*

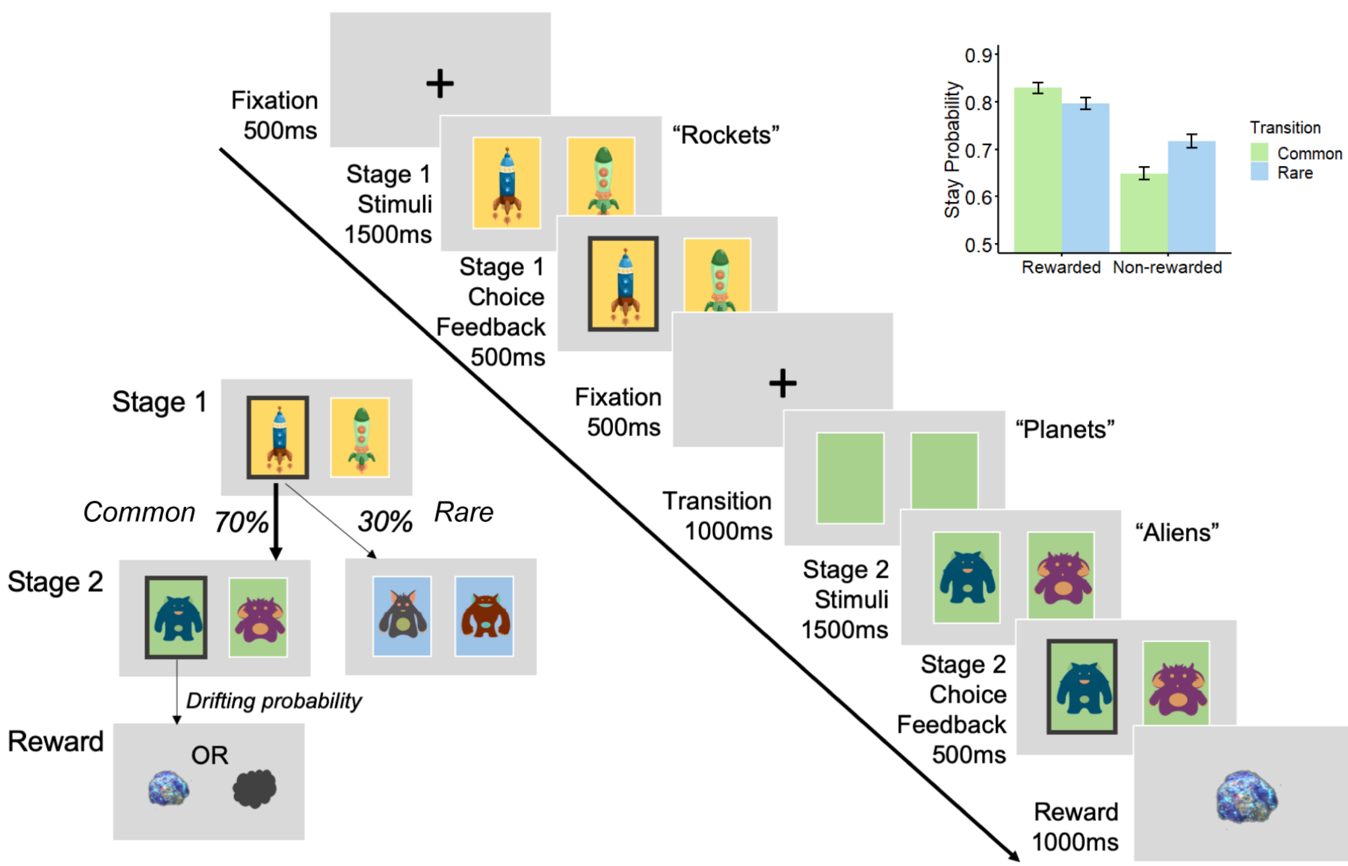
*(**A**) Grand average waveforms of first-stage theta power comparing the top/bottom 50th percentile (N=96 per group) individuals based on their compulsivity (CIT) estimates. Single trial analyses (with all participants) indicate high compulsive individuals exhibit a decrease in theta power (β=-.03, SE=.01, p=.03). In contrast, first-stage alpha power was not associated with compulsivity (**Figure 8A**).*

*(**B**) Reduced theta power at first-stage was linked to several questionnaire scores but the effect was ultimately specific to compulsivity. The Y-axis shows the change in theta power (μV^2) as a function of 1 SD increase of psychiatric questionnaire/dimension scores. Error bars denote standard errors. *p<.05.*

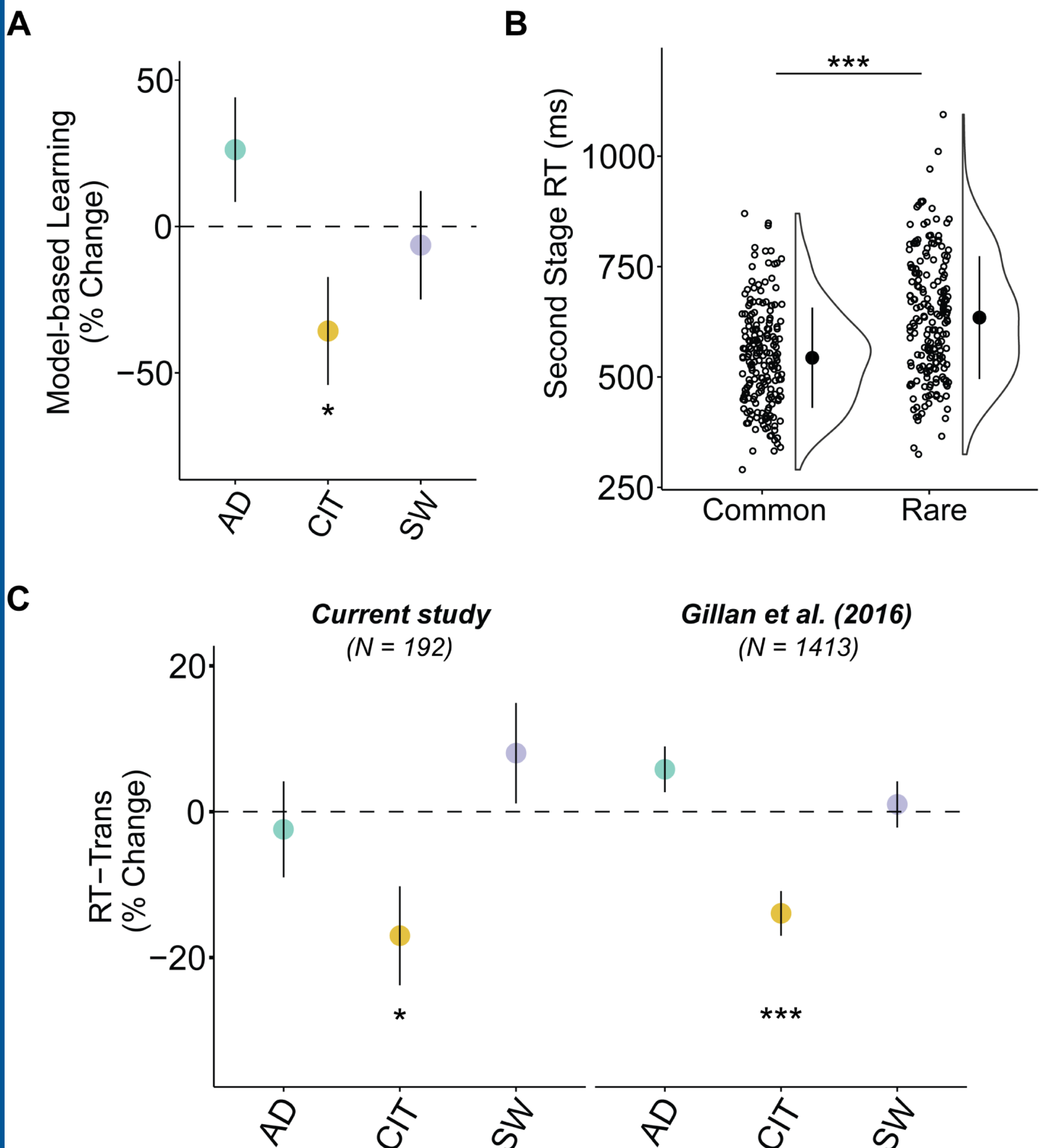
Figure 8. Supplemental analyses.

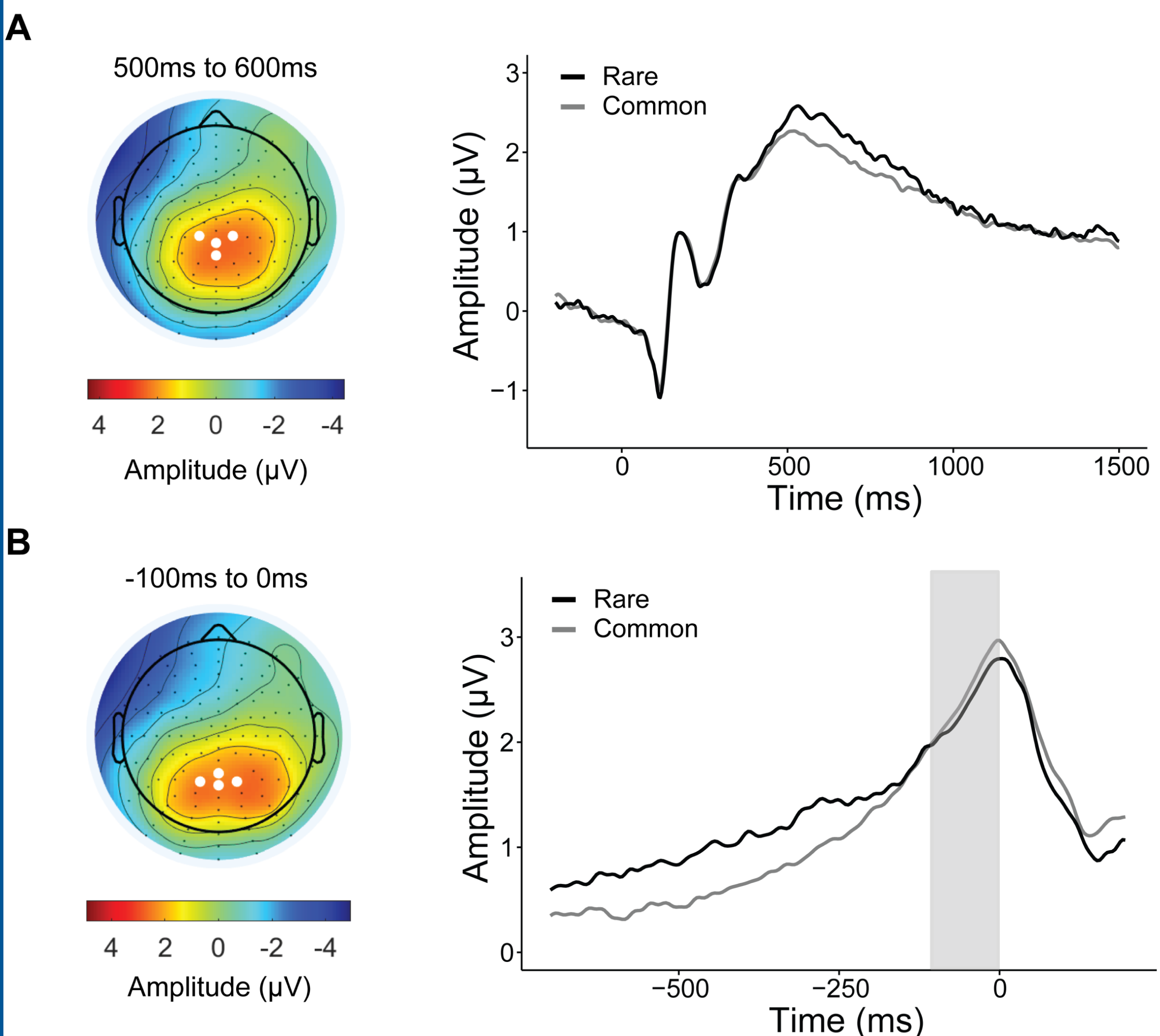
*(**A**) First-stage stimulus-locked alpha power. Topography and line plot (locked to first-stage rockets) show alpha depression during the making a choice at the first-stage. White dots on the topography plot indicate parietal-occipital electrode sites (A18, A19 (Pz), A20, A21, A31) where alpha was measured for both first and second-stages.*

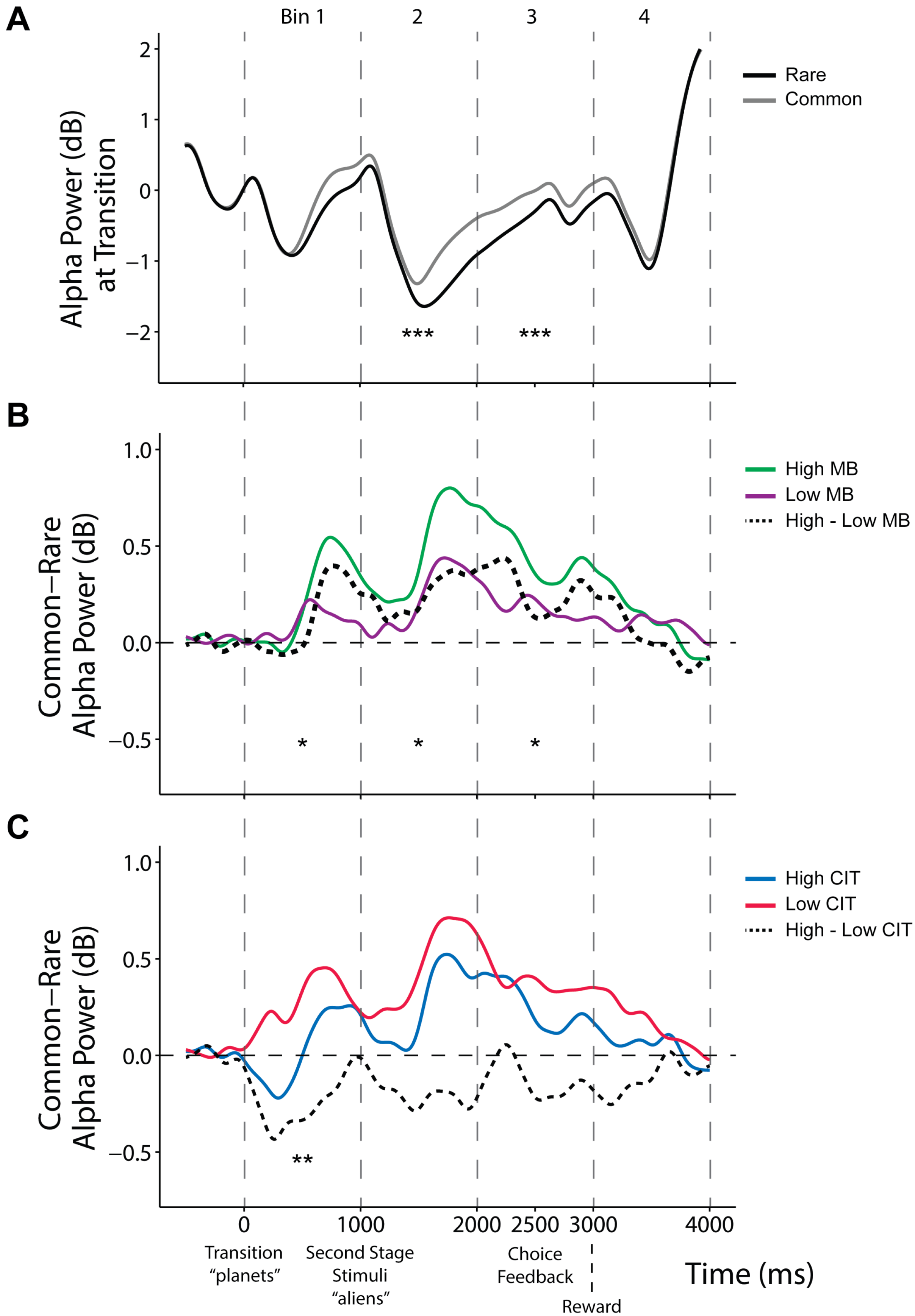
*(**B**) Second-stage stimulus-locked theta power. Topography plot shows theta power increase after stimulus-onset at the mid-frontal scalp. White dots indicate electrode sites (C21 (Fz), C22, C23 (FCz), A1 (Cz)) where theta power was measured for both first and second-stages. Theta power at the first-stage was not associated to compulsivity ($\beta = -0.004$, $SE = 0.02$, $p = 0.84$) nor model-based planning ($\beta = 0.01$, $SE = 0.02$, $p = 0.51$). Theta power was also not linked to transition type ($\beta = -0.01$, $SE = 0.01$, $p = 0.20$) and had no transition interaction effects with compulsivity ($\beta = 0.01$, $SE = 0.01$, $p = 0.14$) nor model-based planning ($\beta = -0.004$, $SE = 0.01$, $p = 0.65$).*

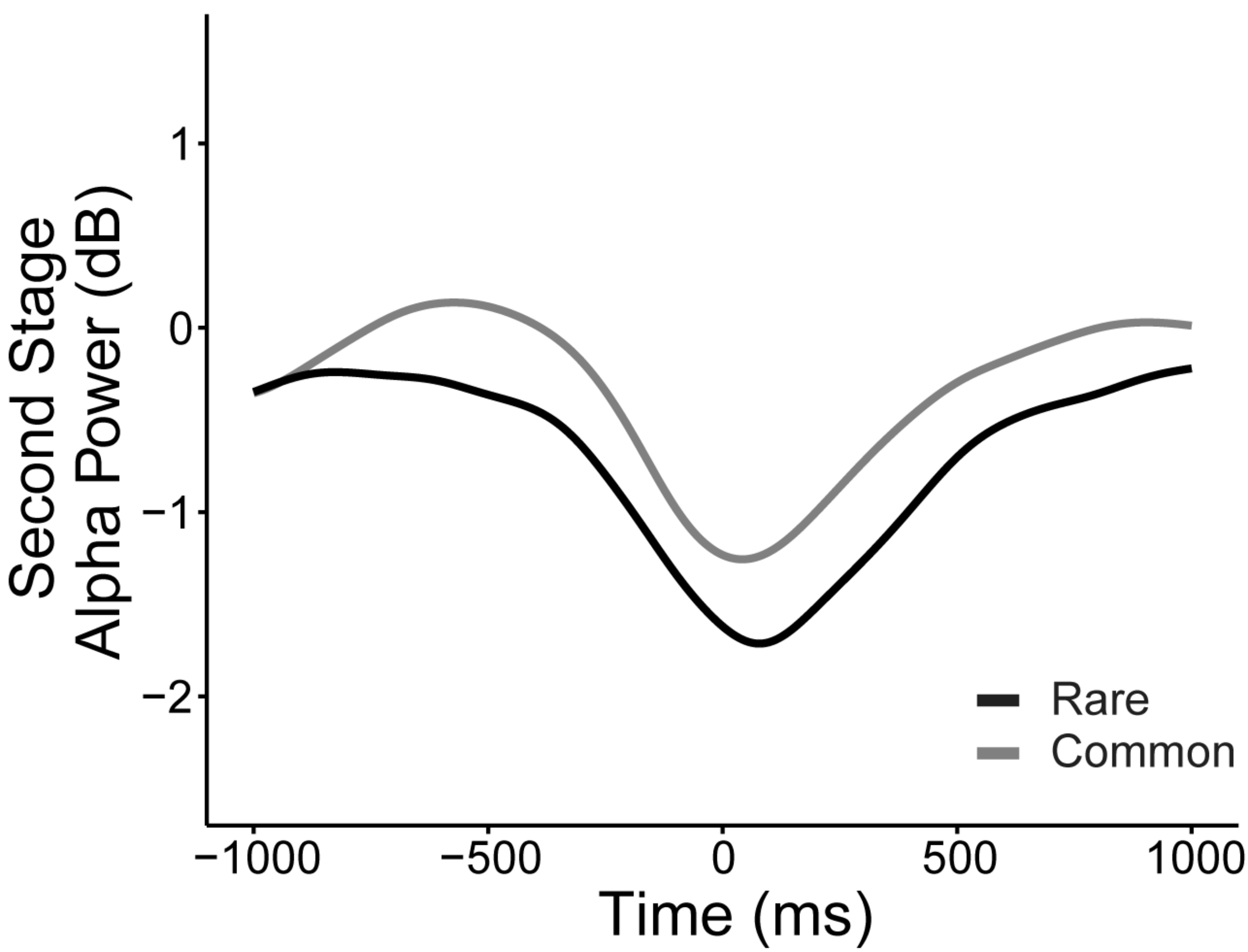


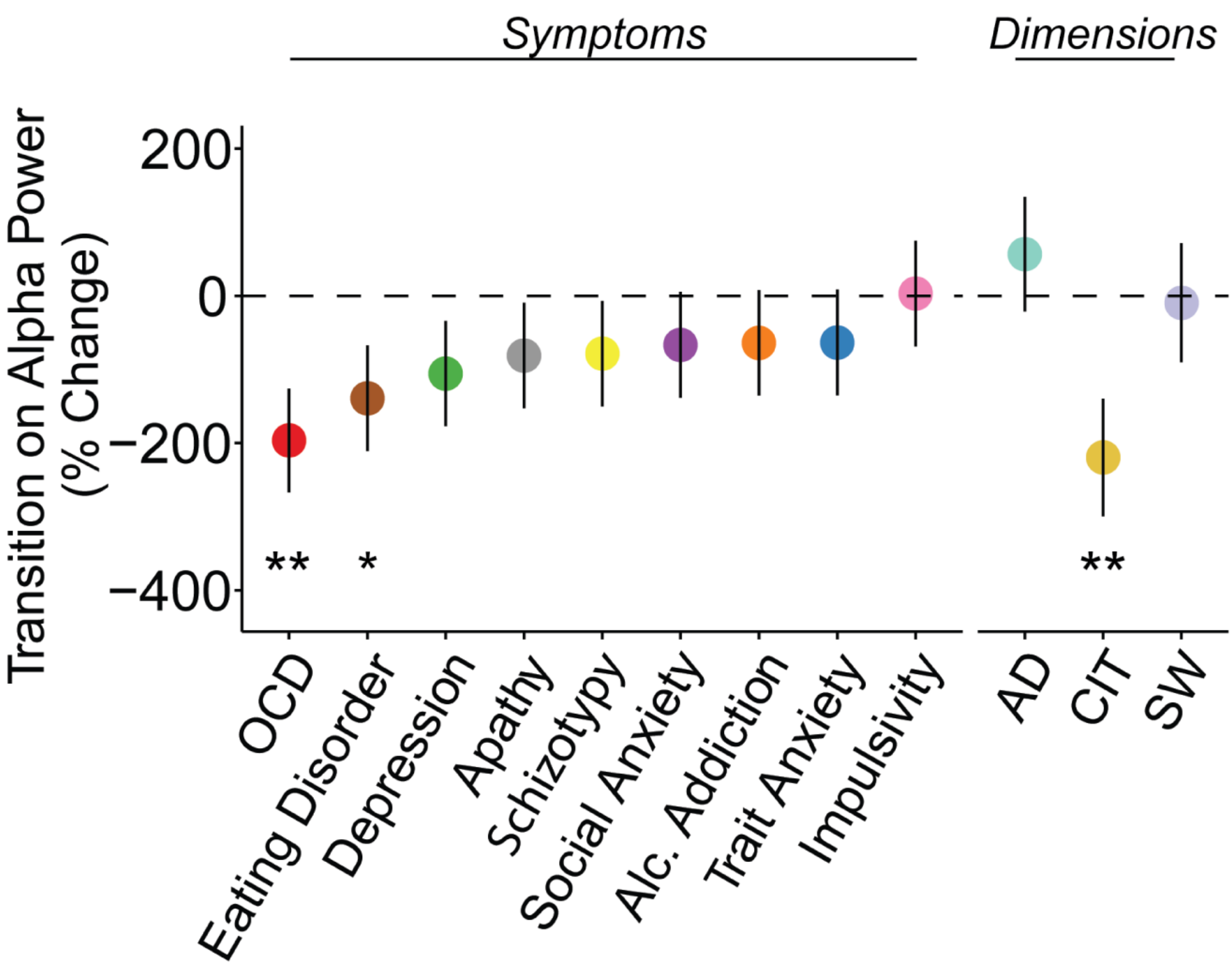
Fixation
500ms
Stage 1
Stimuli
1500ms
"Rockets"
Stage 1
Choice
Feedback
500ms
Fixation
500ms
Transition
1000ms
"Planets"
Stage 2
Stimuli
1500ms
"Aliens"
Stage 2
Choice
Feedback
500ms
Reward
1000ms
Stage 1
Common
70%
30%
Rare
Stage 2
Drifting probability
Reward
OR
Stay Probability
0.9
0.8
0.7
0.6
0.5
Rewarded
Non-rewarded
Transition
Common
Rare

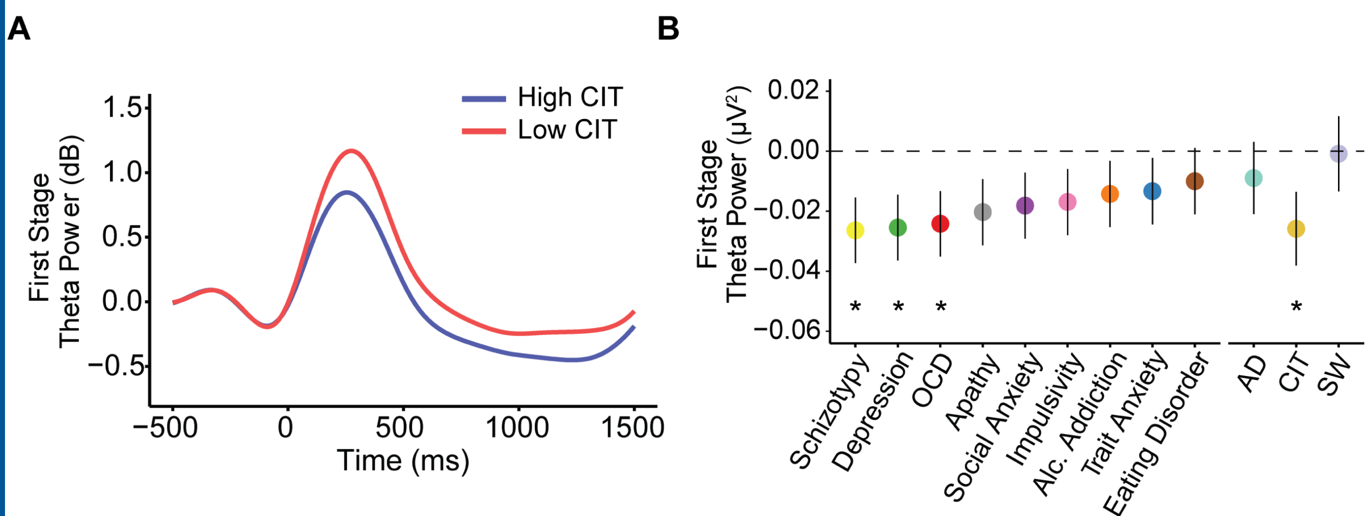

A
Model-based Learning
(% Change)
50
0
−50
*
AD
CIT
SW
B
***
Second Stage RT (ms)
1000
750
500
250
Common
Rare
C
Current study
(N = 192)
Gillan et al. (2016)
(N = 1413)
RT−Trans
(% Change)
20
0
−20
*
***
AD
CIT
SW
AD
CIT
SW


A
500ms to 600ms
4 2 0 -2 -4
Amplitude (μV)
Rare
Common
Amplitude (μV)
Time (ms)
B
-100ms to 0ms
4 2 0 -2 -4
Amplitude (μV)
Rare
Common
Amplitude (μV)
Time (ms)

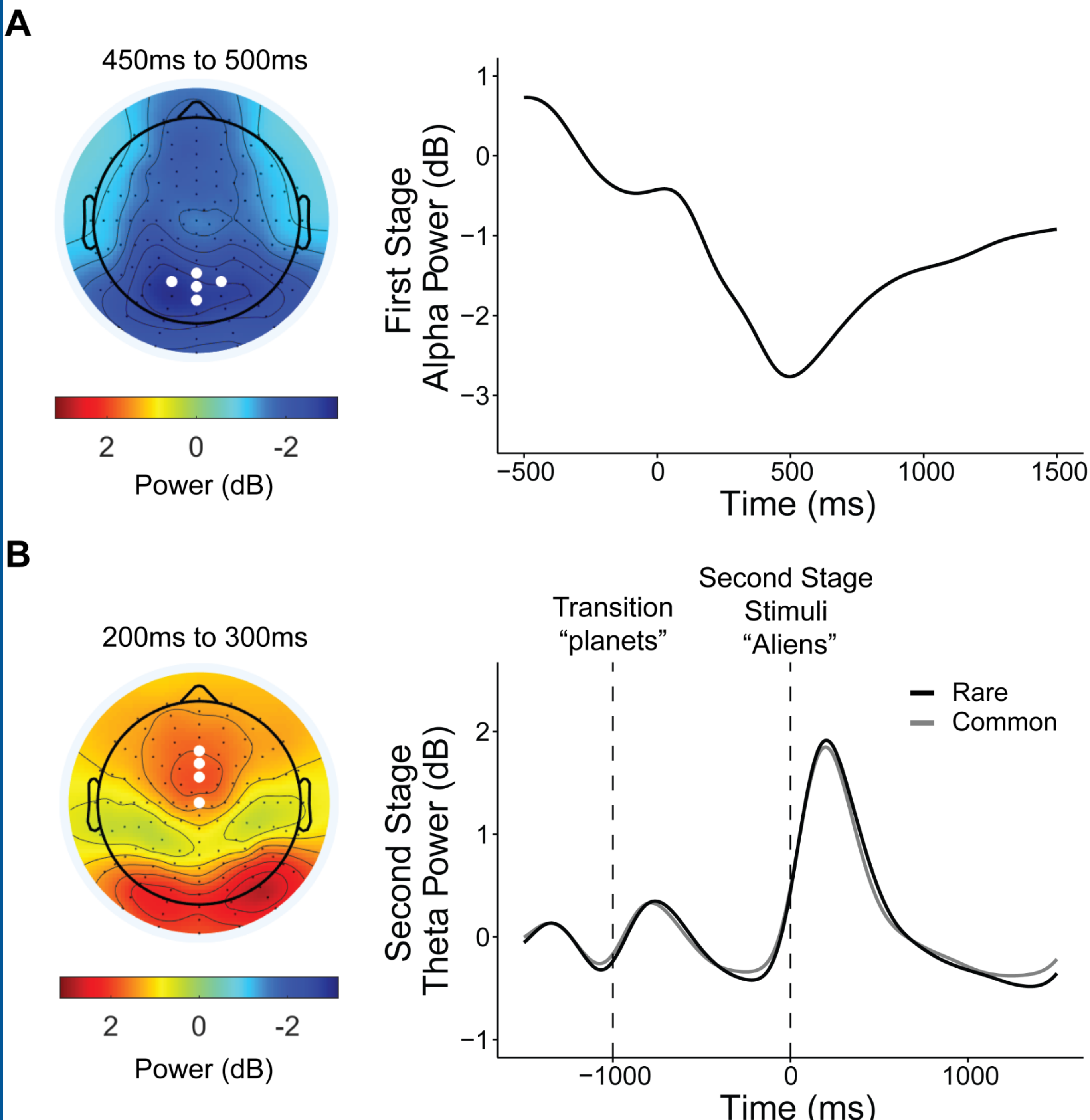

A
Bin 1
2
3
4
Alpha Power (dB) at Transition
Rare
Common
***
***
B
Common−Rare Alpha Power (dB)
High MB
Low MB
High - Low MB
*
*
*
C
Common−Rare Alpha Power (dB)
High CIT
Low CIT
High - Low CIT
**
0
1000
2000
2500
3000
4000
Transition "planets"
Second Stage Stimuli "aliens"
Choice Feedback
Reward
Time (ms)


Second Stage
Alpha Power (dB)
1
0
−1
−2
−1000
−500
0
500
1000
Time (ms)
Rare
Common


Symptoms
Dimensions
Transition on Alpha Power
(% Change)
200
0
−200
−400
**
*
**
OCD
Eating Disorder
Depression
Apathy
Schizotypy
Social Anxiety
Alc. Addiction
Trait Anxiety
Impulsivity
AD
CIT
SW



A
B
High CIT
Low CIT
First Stage
Theta Power (dB)
Time (ms)
First Stage
Theta Power (μV²)
Schizotypy
Depression
OCD
Apathy
Social Anxiety
Impulsivity
Alc. Addiction
Trait Anxiety
Eating Disorder
AD
CIT
SW


A
450ms to 500ms
2
0
-2
Power (dB)
First Stage
Alpha Power (dB)
1
0
−1
−2
−3
−500
0
500
1000
1500
Time (ms)
B
200ms to 300ms
2
0
-2
Power (dB)
Transition
"planets"
Second Stage
Stimuli
"Aliens"
Rare
Common
Second Stage
Theta Power (dB)
2
1
0
−1
−1000
0
1000
Time (ms)