Evidence Accumulation in Rumination: the Role of Uncertainty in Information-Seeking Behavior

Dallas Brodersen ‘24

Submitted to
The Department of Psychology, Princeton University
in partial fulfillment of the requirements for the degree of Bachelor of Arts.

Faculty Advisor: Dr. Yael Niv
April 22, 2024

Authorization Page

This paper represents my own work in accordance with University Regulations.

I authorize Princeton University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Princeton University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Dallas Brodersen

Dallas Brodersen

Abstract

Rumination is a common symptom of depression involving repetitive, self-focused thoughts of negative valence. Past studies have hypothesized that the repeated sampling of events during rumination may function to achieve certainty about the consequences of potential future actions. Here, we test the predictions of a computational model of this theory with an online information-seeking task. Using a modified version of the 'beads task' (Phillips & Edwards, 1966), we investigated whether increased information seeking (i.e., increased draws-to-decision) is associated with the severity of rumination symptoms, as determined by responses to a battery of self-report questionnaires. In our task, participants (n = 73) had to choose between two actions (each of which led to either reward or loss) or sample more information (which may increase knowledge of which choice would lead to reward). We hypothesized that increasing perceptual overlap between the distributions (i.e., making it unclear if the stimulus was from the distribution that would require action 1 or the distribution that would require action 2) would increase draws-to-decision to resolve the ambiguity of the current distribution. We further predicted more draws-to-decision in individuals who are high-trait ruminators. Our model-based analysis demonstrated a relationship between depression score and decision threshold. We found that participants with higher depression scores exhibit less of a difference between medium and low uncertainty condition decision thresholds. This result suggests that more depressed participants required a higher degree of certainty before making a choice, possibly due to perceiving the samples as more ambiguous.

Keywords: rumination, perceptual decision-making, belief updating, uncertainty

1. INTRODUCTION

The high rate of comorbidity among mental disorders has motivated the study of transdiagnostic processes, or features present across disorders that are associated with their etiology and maintenance (McLaughlin and Nolen-Hoeksema, 2011). One example of a transdiagnostic process strongly associated with major depression and generalized anxiety disorder is rumination. This thinking style is characterized by involuntary and compulsive repetitive negative thoughts focused on past events, which persist in the absence of immediate environmental demands necessitating the thoughts (Martin and Tesser, 1996; McLaughlin and Nolen-Hoeksema, 2011).

To effectively target ruminative thinking in transdiagnostic treatment approaches for emotional disorders, we must first disambiguate the factors contributing to the behavior (McLaughlin & Nolen-Hoeksema, 2011). Given it is a thinking process that is inherently unpleasant, traditional studies on rumination typically view the behavior as suboptimal (Ward et al., 2003; Berg et al., 2021; McLaughlin and Nolen-Hoeksema, 2010). An alternative approach that assumes rumination is optimal, however, may help uncover nuanced conditions that drive the behavior (Bedder et al., 2023).

In line with this alternative approach, rumination has been framed as a repeated sampling of information from memory to understand the cause of some negative affective experience (Bedder et al., 2023). Bedder et al. (2023) conceptualize this process as an attempt at hidden-state inference, that is, trying to infer the (hidden) state of the world that a negative experience belongs to by sampling thematically similar memory episodes. For example, one might infer that their boss yelling at them is either part of the state ‘everyone is in a bad mood on Mondays’ or ‘I am underperforming at

my job.' To infer the hidden state, one might recall other instances where co-workers acted unfavorably towards them at the beginning of the week.

To infer a hidden state with sufficient certainty is adaptive because it can enable an individual to make an informed choice that maximizes reward (Bedder et al., 2023; Schulz et al., 2020). That is, by inferring the hidden state with sufficient certainty, an individual can select the action that will lead to a positive outcome (e.g., opt to work from home if the state is 'everyone is in a bad mood on Mondays,' or stay late if the state is 'I am underperforming'). To be effective, however, this strategy requires an eventual transition into a mode of active problem-solving, which is rarely the result of a ruminative episode (Nolen-Hoeksema, 1991; Nolen-Hoeksema et al., 2008; Takano & Tanno 2009; Watkins & Roberts 2020).

In a two-part study, Ward et al. (2003) offered insight into why high-trait ruminators might passively fixate on their thoughts without transitioning into an action-oriented state. Identifying uncertainty as a substantial barrier to instrumental behavior, Ward et al. (2003) hypothesized that when asked to generate plans in response to two types of community and personal problems, ruminators would remain uncertain about which action to take, even after extensive deliberation. To test this prediction, ruminators and non-ruminators were grouped based on their scores on the Ruminative Responses Scale (RRS), a ten-item self-report rumination questionnaire (Nolen-Hoeksema & Morrow, 1991). Study 1 compared each group's evaluation of a self-generated plan to revise their university housing system. Participants were then tasked with presenting their proposals for approximately 5 minutes with essentially no time to 'rehearse.' In Study 2, participants were asked to formulate a plan to redesign

the current undergraduate curriculum for approximately 5 min while being videotaped. Both studies demonstrated that, compared to non-ruminators, ruminators expressed lower levels of satisfaction and confidence in their plans. That is, ruminators were significantly less likely to commit to their self-generated solutions and more likely to report requiring more time to research them.

Berg et al. (2021) created a model of rumination development and maintenance consistent with the findings of the Ward et al. (2003) study (See Appendix A). The primary hypothesis proposed by Berg et al. (2021) involves the concept of policy candidate oversampling, an attempt at mental problem-solving that involves the premature rejection of potential actions (policy candidates). Their model describes a two-step policy selection process and demonstrates how oversampling of policy candidates inhibits effective mental problem-solving and contributes to an endless cycle of sampling and discarding candidate policies. For instance, someone aiming for the perfect verbal response might become stuck in a cycle of sampling and discarding potential options for what to say (i.e., candidate policies), causing them to refrain from speaking altogether (Berg et al., 2021).

This prediction of how certain alterations in the inference and action-selection policies result in a protracted thinking process has been formalized in computational psychiatry. In particular, Bedder et al. (2023) model rumination as sampling in a partially observable Markov decision process (POMDP), a mathematical framework used to model decision-making under uncertainty. Rumination was represented in this context as prolonged sampling (i.e., the agent's need to sample more information before selecting an action). Simulations of this model revealed that extended sampling was the

optimal policy under high information ambiguity or high magnitude of prospective loss. This outcome aligns with the Berg et al. (2021) model in which protracted thinking results from insufficient certainty.

The relevance of high information ambiguity to information-seeking (i.e., prolonged sampling) aligns with sequential sampling models of decision-making. That is, the less clear a decision is, the more samples appear to be drawn to alleviate uncertainty, particularly when the decision holds significant importance (Shadlen & Shohamy, 2016; Baker et al., 2019). In other words, decision-making requires the accumulation of sensory evidence.

This notion is exemplified in classic paradigms like the beads task (Phillips & Edwards, 1966), a probabilistic reasoning task widely used for investigating decision-making under uncertainty. In this task, participants draw and replace beads from one of two hidden urns containing colored beads in equivalent but opposite proportions (e.g., one urn contains 85 blue and 15 red beads and another 85 red and 15 blue beads; participants do not know which urn they are sampling from) until they feel confident enough to guess the identity of the hidden urn given the observed sequence of beads (Baker et al., 2019). The primary measure, ‘draws-to-decision’ (i.e., the number of beads drawn by a participant before making a guess), offers insight into abnormalities in decision-making (Baker et al., 2019).

Researchers have used variations of the beads task to elucidate differences in decision-making among individuals with psychiatric illnesses (Romero-Ferreiro et al., 2022; Jacoby et al, 2016; Lawrance et al., 2022; Baker et al., 2019). For instance, Baker et al. (2019) investigated delusions in schizophrenia, demonstrating that increased

draws-to-decision correlated specifically with delusion severity. The study identified deficient belief updating, characterized by a stronger reliance on prior beliefs formed early in the inferential process, as the underlying mechanism driving increased information-seeking in more delusional patients. This finding explains the canonical form of delusions as rigid beliefs that persist despite the presentation of new, and possibly contradictory, evidence (Baker et al., 2019).

Other studies have used perceptual ambiguity paradigms to study evidence accumulation and decision-making processes within clinical populations. For instance, Banca et al. (2015) utilized a dot-motion task to investigate the interaction between uncertainty and obsessive-compulsive disorder (OCD). Here, nine different motion coherence levels representing high, medium, and low uncertainty, were defined by varying the proportion of dots that were moving in the same direction, and participants had to report the direction of motion. Results showed that compared with healthy controls, OCD patients required more evidence under high uncertainty conditions, as evidenced by greater response time and higher decision boundaries, suggesting that uncertainty exerts greater interference on their perceptual judgments (Banca et al., 2015).

Utilizing stimuli of varying degrees of ambiguity is a common method for isolating the role of uncertainty in decision-making. Sun et al. (2017) investigated value-based decision-making under ambiguity with a face judgment task with three ambiguity levels (anchor, intermediate, and high). In Study 1, participants interpreted morphed faces expressing fear-happy emotions, and Study 2, involved an animal judgment task with cat-dog morphs (Sun et al., 2017). Maksimenko et al. (2020) used Necker cubes of two

degrees of ambiguity (two distributions) as ambiguous visual stimuli, with half of them considered left-oriented and another half right-oriented.

This present study aimed to validate the prediction of the Bedder et al. (2023) model that draws-to-decision is modulated by information ambiguity (i.e., draws-to-decision increases as a function of information ambiguity). The primary objective was to investigate whether individuals with a greater tendency to ruminate (as determined by responses to a battery of self-report questionnaires) exhibit increased sensitivity to uncertainty. The Baker et al. (2019) beads task provides a useful basis for some experimental design choices. Our study modified their approach by integrating elements from the Banca et al. (2015), Sun et al. (2017), and Maksimenko et al. (2020) task designs to test sensitivity at various degrees of ambiguity.

We created a novel online study, the ‘mushroom categorization task.’ Participants sampled visual stimuli (mushrooms) coming from one of two states, with various degrees of perceptual uncertainty, to decide whether to take one of two actions (one of which led to reward in state 1, and the other for the other state; wrong actions led to punishment). At each time point, the participant could choose an action or sample another mushroom. Building on Bedder et al.’s (2023) model, we predicted that increasing the information ambiguity would protract the sampling process (i.e., increase draws-to-decision) for all participants. We further predicted that participants who score higher on self-report rumination questionnaires would show higher sensitivity to uncertainty in terms of increased draws-to-decision. These predictions were formed based on Bedder et al.’s (2023) computational framework for rumination, which suggests an interaction among trait rumination, uncertainty, and draws-to-decision.

Thus, this study represents a first attempt to test the predictions of Bedder et al's model of rumination.

2. METHODS

2.1. Participants

73 participants (ages 19-67; mean age = 37.66, 30 female, 43 male) were recruited through Prolific, an online research platform. Participants were compensated 15 dollars for completing the task, which they were informed would take under 60 minutes. Participants received a bonus between $0 and $5 based on their task performance. They received $1 if their performance was at chance level (chance level was calculated as the number of points for equal correct and incorrect distribution evenly between blocks). They received between $1 and $5 for points above chance level and between $0 and $1 for points below chance.

2.2. Clinical Measures

Participant depression, anxiety, and rumination symptom severity were quantified by self-report responses to a battery of three questionnaires: the Beck Depression Inventory-Short Form (BDI-SF; Beck & Beck, 1972), the Beck Anxiety Inventory (BAI; Beck & Steer, 1990), and the Ruminative Responses Scale (RRS), a subscale of the Response Styles Questionnaire (RSQ; Nolen-Hoeksema & Morrow, 1991). These questionnaires were selected due to their use in a McLaughlin & Nolen-Hoeksema (2010) study on whether rumination accounts for the co-occurrence of depression and anxiety to a significant degree.

The 13-item BDI-SF, an abbreviated version of the standard 21-item BDI, was used to assess depressive symptoms. The BDI-SF is a self-report measure with possible scores ranging from 0 to 39 (calculated by assigning scores of 0, 1, 2, or 3 to a statement that best describes the way they have been feeling the past week). Due to

IRB restrictions, item 7, which pertains to suicide, was excluded from our study, resulting in a possible score range of 0 to 36. Score ranges of 5–7 indicate mild depression, 8–15 indicate moderate depression and 16 or higher indicate severe depression. Scores on the BDI-SF ranged from 0-25.

The 21-item self-report BAI was administered to assess anxiety symptoms. Possible scores on the BAI range from 0 to 63 (calculated by assigning scores of 0, 1, 2, or 3 to the response categories ‘not at all,’ ‘mildly but it didn’t bother me much,’ ‘moderately – it wasn’t pleasant at times,’ and ‘severely – bothered me a lot,’ respectively). A score of 0-21 indicates low anxiety, a score of 22-35 indicates moderate anxiety and a score greater than 36 indicates potentially concerning levels of anxiety. Scores on the BAI ranged from 0-42.

The 22-item RRS was used to assess the tendency to ruminate in response to a depressed mood. Items describe responses to depressed mood that are either focused on the self or possible causes and consequences of the mood. Possible scores on the RRS range from 22 to 88, with higher scores indicating higher degrees of ruminative symptoms (calculated by assigning scores of 1, 2, 3, or 4 to the response categories ‘almost never,’ ‘sometimes,’ ‘often,’ and ‘almost always,’ respectively). Scores on the RRS ranged from 22-71.

2.3. Uncertainty Visual Stimuli

We used a mushroom icon as an ambiguous visual stimulus. Mushroom stimuli were generated using a custom CSS function that draws the stem, base, and head of the fungi according to specified dimensions and colors. We varied the width of the stem by sampling from one of two distributions. The first distribution (‘tall fungi’) had a mean of

0.4 which made the mushroom stems tall and narrow (Fig. 1a). The second distribution ('short fungi') had a mean of 1.6 which made the mushroom stems short and wide (Fig. 1b).

The standard deviation of the distributions was manipulated to create three different uncertainty levels. In other words, we varied the overlap between the tall and short fungi distributions with more overlap corresponding with greater uncertainty. The low uncertainty distributions had an SD of 0.2, medium uncertainty distributions had an SD of 0.8, and high uncertainty distributions had an SD of 1.04.

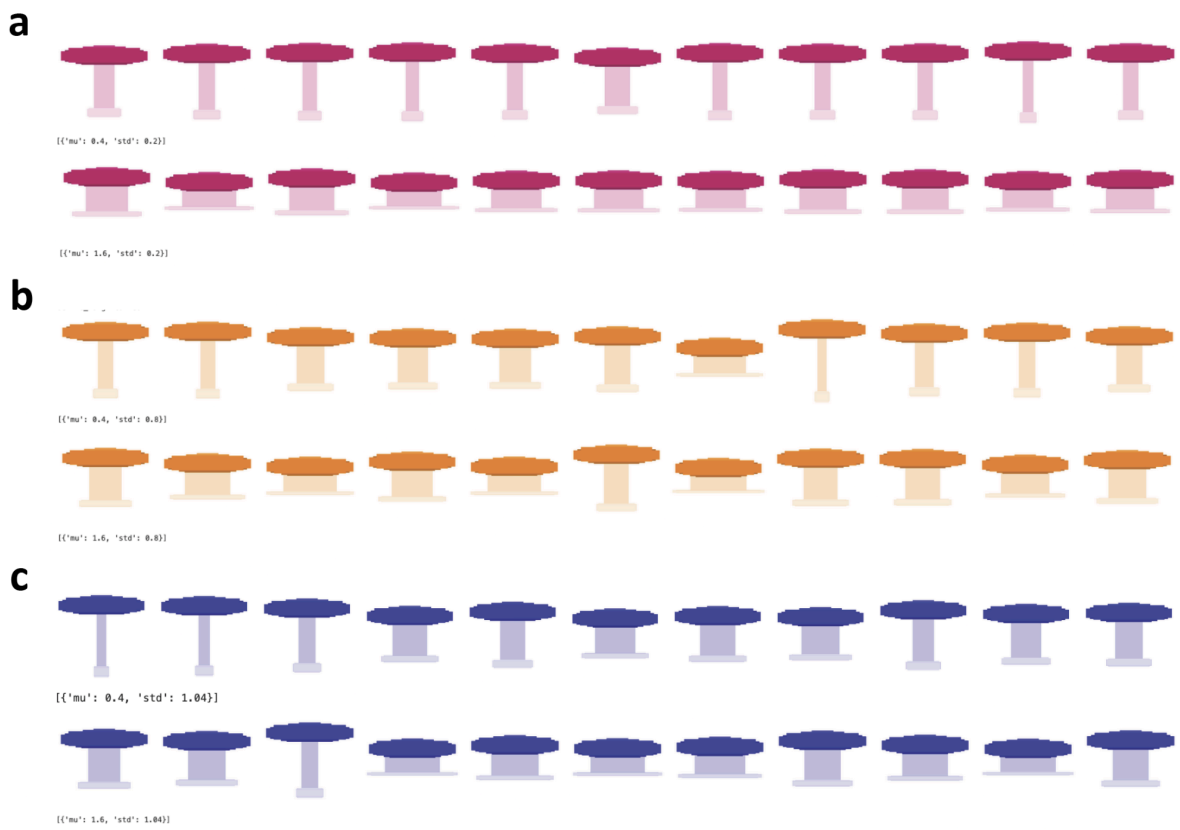

Fig. 1 | Fungus distributions from low, medium, and high uncertainty conditions. The upper distribution is tall (M = 0.4) and the lower distribution is short (M = 1.6) for each uncertainty condition. **a**, Mushrooms from the low uncertainty condition (SD = 0.2). **b**, Mushrooms from the medium uncertainty condition (SD = 0.8). **C**, Mushrooms from the high uncertainty condition (SD = 01.04).

2.4. Procedures

Participants reviewed the instructions and completed comprehension checks and practice trials for the task (3 practice blocks of 2 trials each). Participants then moved on to the task (see section 2.5). After completion of the task, participants completed a demographics survey, a battery of questionnaires to assess symptoms of depression, anxiety, and rumination (see section 2.2), and a debriefing questionnaire to provide feedback on their experience of the task and strategies used. In both pilot sessions, the demographics survey and symptom assessments were omitted.

2.5. Mushroom Categorization Task

Participants completed 8 blocks of each uncertainty level in a randomized order (24 blocks in total). Thus, each uncertainty level had 40 trials. Half of the blocks corresponded with a high reward and the other half corresponded with a neutral reward. On each trial tall or short was randomly selected to be correct. Results for the different reward levels were included for a separate project and are not analyzed here.

Blocks were referred to as forests. Fungi color varied between blocks but was consistent within each block (i.e. the high gain, low certainty block would always be the same color). The color allocation for each gain and certainty combination was randomized for each participant. Due to an error in the code, the forest color was set to black for the final 18 blocks, resulting in only black mushrooms being generated. However, stimulus color was not relevant to any task elements and should not have affected the results. At the onset of each block, participants were informed of the new stimulus color, along with the uncertainty level and reward and punishment magnitudes

for that block (see Appendix B3). Certainty and reward levels were shown on screen, except when a mushroom was being shown, or the fungus was being labeled.

The three uncertainty levels were represented as mushroom maturity stages: elder for low uncertainty, mature for medium uncertainty, and immature for high uncertainty. The maturity level for each block was indicated by a ‘maturity bar’ (see Appendix B1).

Potential reward and punishment magnitudes were communicated to participants using happy or sick emojis, the options being either 3,1 (high gain) or 2,2 (neutral reward) (see Appendix B2). Participants were rewarded for correctly labeling the fungus species on each trial. Each happy emoji earned 50 points, and each sick emoji lost -50 points. Total points were later converted to a financial bonus (see section 2.1).

Participants labeled each fungus (tall or short) by looking at mushrooms from that fungus. On each trial, participants were presented with a choice between ‘sample’ or ‘decide’ (Fig. 2). If they chose to sample, participants next saw a mushroom (duration = 1000 ms) with a stem width that had been sampled from a distribution from the uncertainty level of that block. Participants were able to sample a maximum of 30 times (they were unaware of this maximum unless they reached it). If participants chose to label, they were presented with a choice between tall or short (duration = 10,000 ms) (Fig. 3). Once they made a choice, the correct fungus was shown to participants by showing the rewards if correct, or punishments if incorrect. Correct responses were dictated by which of the two states had generated the observations. After responding, participants proceeded to the next trial.

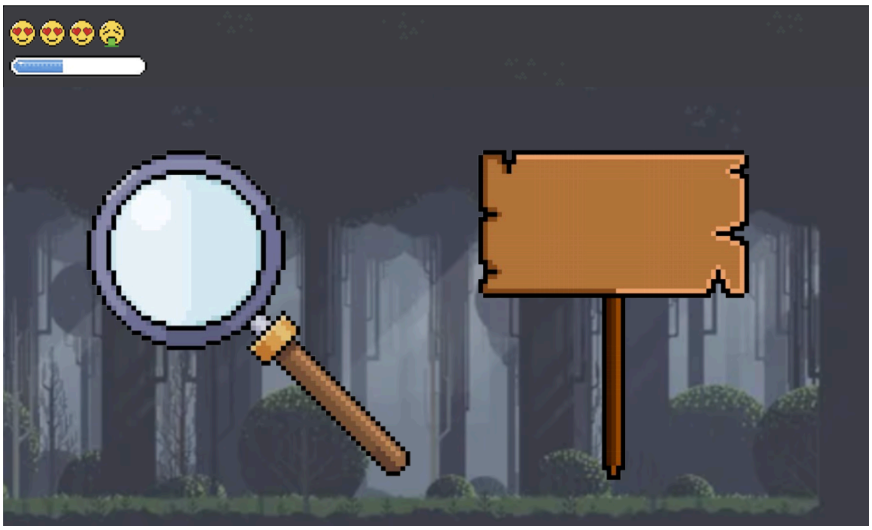

Fig. 2 | Participants were presented with a choice between sampling or labeling. Full-screen view. Participants could choose the magnifying glass by using a key press, allowing them to view a mushroom with a stem width that had been sampled from a distribution from the uncertainty level assigned to that block. Participants could sample as many (with a maximum of 30) or as few mushrooms as they desired. Participants could choose the sign by using a key press, allowing them to label the fungus and move on to the next trial.

Fig. 3 | Participants were presented with a choice between 'tall' and 'short.' Full-screen view of the fungus labeling screen. Participants could choose either 'tall' or 'short' using a key press. The tea kettle and bowl of soup icons visible on each sign were a part of the cover story and helped participants understand why labeling was important. These icons were consistent with each fungus (i.e., tall was always tea).

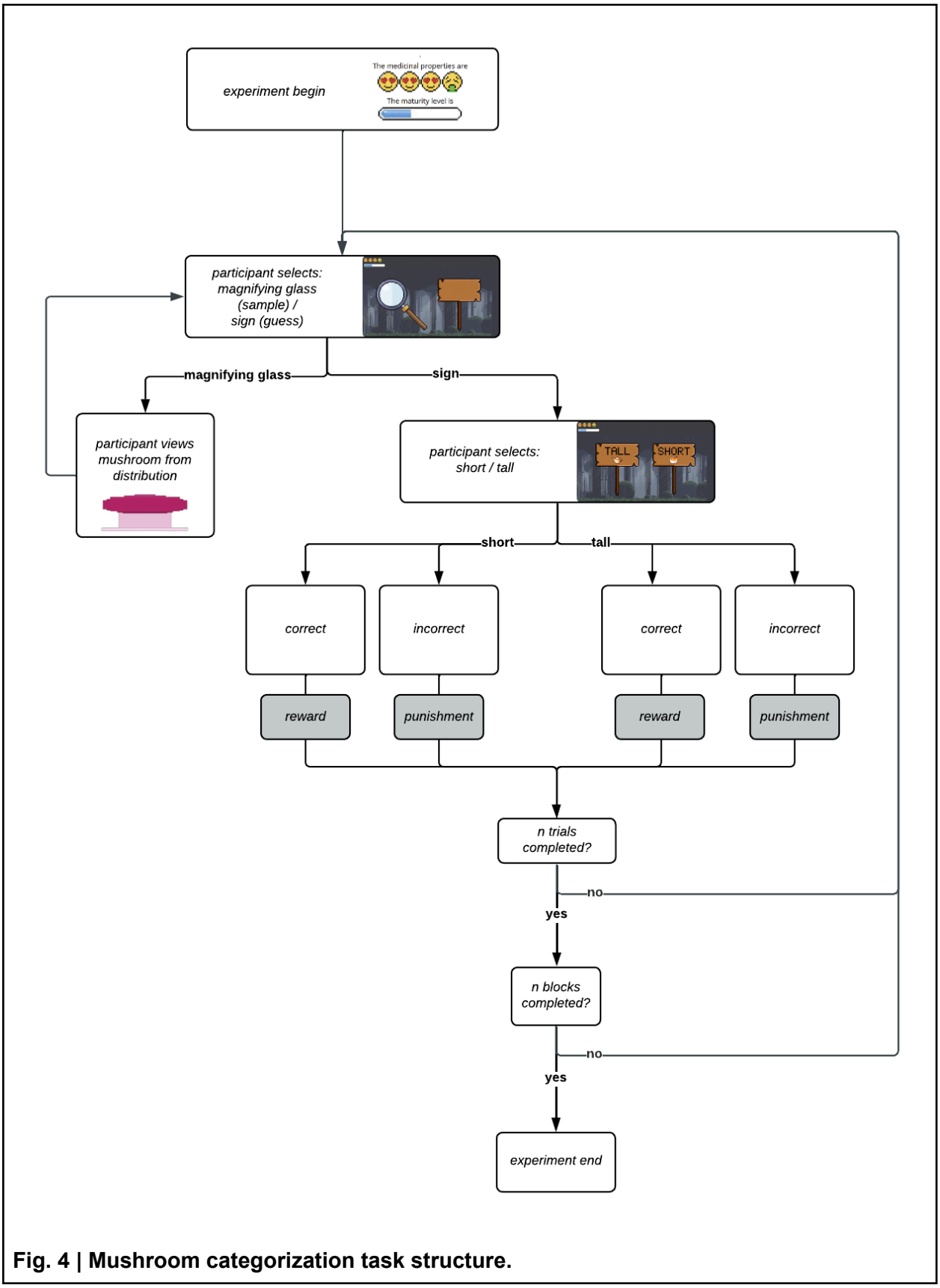

Fig. 4 | Mushroom categorization task structure.

2.6. Pilot

We ran two pilots to validate the clarity of the task instructions and whether participants were sensitive to changes in the uncertainty levels. 19 adult participants (n = 5 for Pilot 1; n = 14 for Pilot 2) were recruited through Prolific.

The task included three different uncertainty levels (low SD = 0.2, medium SD = 0.8, high SD = 1.04; see section 2.3). Participants completed 2 blocks of each uncertainty level in a randomized order. Each block contained 5 trials where tall was correct on 50% of the trials. Thus, each uncertainty level had 10 trials. Half of the blocks corresponded with a high reward and the other half corresponded with a neutral reward. The results for the different reward levels were included for a different project and are not analyzed here.

Participants were compensated 7 dollars and 50 cents for completing the task, which they were informed would take under 30 minutes. Participants received a bonus between $0 and $3 based on their task performance. They received $1 if their performance was at chance level (chance level was calculated as the number of points for equal correct and incorrect distribution evenly between blocks). They received between $1 and $3 for points above chance level, and between $0 and $1 for points below chance.

2.7. Data Analysis

Model-Agnostic Analysis

To determine whether the number of samples drawn (henceforth referred to as ‘draws-to-decision’) was significantly different than 1 under each condition, we ran a one-sample t-test between draws-to-decision under each uncertainty level and 1. To

compare whether draws-to-decision was significantly different between the two uncertainty levels, we ran paired samples t-tests. We consider an alpha level of 0.05 significant for these tests.

For our hypotheses, we were interested in sensitivity to uncertainty when participants were engaged in the task. We were therefore particularly interested in trials where participants made the correct choice. In the model-agnostic analyses, correct trials are defined based on the generating distribution (i.e. whether tall or short was correct on the given trial, irrespective of whether the mushrooms drawn supported this).

Model-Based Analysis

As the stimuli on each trial were generated from overlapping distributions, Bayes theorem was used to calculate the probability that each state was correct based upon the set of observations participants saw. That is to say, this analysis considers the evidence participants have seen and assumes they update their beliefs about the hidden state according to Bayes rule (Equation 1; Kaelbling et al., 1998).

In this simple model with only two possible hidden states, X_1 and X_2 ('tall' and 'short'), we define B = $P(X_2)$ and $P(X_1)$ = 1 − B. We use B′ to denote the subsequent belief state. In our experiment, each mushroom drawn is an observation *o* from the hidden state. Each time a participant samples a new mushroom, their posterior (B') is updated based on Bayes theorem and becomes the prior (B) for the next decision point. We set B to 0.5 at the beginning of each new trial.

$$B' = \frac{P(o|X_2)B}{P(o|X_1)(1-B)+P(o|X_2)B} \quad (1)$$

As before, we were interested in trials where participants made the correct choice, however, we ran the same analyses for correct-only trials. Since certainty about the correct label was dependent upon whether high or low-probability mushrooms were drawn throughout the trial, correct trials were defined in terms of what the participant observed.

We calculated the Bayesian posterior for each trial in terms of the stimulus that was most likely to be correct based on the set of observations the participant saw on that trial. We transformed the posterior for trials where participants were incorrect on the generating distribution but correct on evidence seen by subtracting the posterior for each of those trials from 1.

The final posterior before participants chose to decide on the state was used as a 'decision threshold.' We used one-sample t-tests to compare the decision thresholds for each uncertainty level with 0.5 (chance). We used paired samples t-tests to compare decision thresholds between uncertainty levels. We consider an alpha level of 0.05 significant for these tests.

Questionnaires

Questionnaires were scored according to the criteria outlined in Methods section 2.2. We report Spearman correlation coefficients for effect sizes of relationships between questionnaire scores. We consider an alpha level of 0.05 significant for these tests. We also report Spearman’s correlation coefficients for effect sizes of relationships between model-agnostic and model-based task measures and questionnaire scores. We consider an alpha level of less than 0.01 to be significant for these tests.

Exclusion Criteria

One strategy for completing the task faster is drawing fewer samples. Therefore, excluded from all analyses were all participants who chose 1 sample over 90% of the time for all 3 uncertainty levels. We determined these exclusion criteria after analyzing Pilot 2, and before collecting data for our full sample of participants.

3. RESULTS

3.1. Pilot 1

We first tested whether the draws-to-decision was greater than 1 in each condition, this was not significant in low (M=0.98; SD=0.07); [t = -0.59, p = 0.59], medium (M=1.08; SD=0.21); [t = 0.85, p = 0.44], or high uncertainty (M=1.07; SD=0.16); [t = 0.99, p = 0.37]. This suggests that participants were either not motivated or not cognizant of the option to sample multiple times. This supposition was confirmed by contacting Pilot 1 participants via Prolific, who confirmed they were unaware they could take more than 1 sample on each trial. Thus, to ensure awareness of the option to sample multiple times, the experiment was updated with an additional comprehension check and a message appearing on all practice trial screens explicitly stating the option to view another mushroom.

3.2. Pilot 2

One of the original 14 participants was excluded as determined by predefined exclusion criteria (for exclusion criteria see Methods section 2.7; for Pilot 1 figure see Appendix C1). The results of Pilot 2 confirmed the success of the changes made following Pilot 1, leading to no additional adjustments being made thereafter.

Model-Agnostic Analysis

We first tested whether the draws-to-decision were greater than 1 in each condition and found it was significant in high uncertainty (M = 3.71; SD = 1.34); [t = 7.31, p < 0.001], medium uncertainty (M=3.56; SD = 1.66); [t = 5.58, p < 0.001], and low uncertainty (M =

2.24; SD = 0.89); [t = 5.00, p < 0.001], thereby confirming that the changes implemented after Pilot 1 were successful (see Appendix C2).

We next tested whether low and medium uncertainty levels were significantly different. We found that participants took more samples in medium uncertainty (all trials: M = 3.56; SD = 1.59; correct only: M = 3.58; SD = 1.59) than low uncertainty (all trials: M = 2.24; SD = 0.86; correct only: M = 2.23; SD = 0.81), (all trials: [t = -4.29, p < 0.001]; correct only: [t = -4.61, p < 0.001]).

Next, we tested whether medium and high uncertainty levels were significantly different. We found that participants did not take significantly more samples in either high uncertainty (all trials: M = 3.71; SD = 1.28; correct only: M = 3.64; SD = 1.24) or medium uncertainty (all trials: M = 3.56; SD = 1.59; correct only: M = 3.58; SD = 1.59), (all trials: [t = -0.45, p = 0.66]; correct only: [t = -0.16, p = 0.87]).

Model-Based Analysis

We tested whether the Bayesian decision thresholds were greater than 0.5 in each condition and found it was significant in high uncertainty (M = 0.72; SD = 0.04); [t = 19.61, p < 0.001], medium uncertainty (M = 0.81; SD = 0.06); [t = 17.45, p < 0.001] and low uncertainty (M = 0.99; SD = 0.01); [t = 147.65, p < 0.001].

We next tested whether low and medium uncertainty levels were significantly different. We found that the decision threshold was greater in low uncertainty (all trials: M = 0.99; SD = 0.01; correct only: M = 1.00; SD < 0.001) than medium uncertainty (all trials: M = 0.81; SD = 0.06; correct only: M = 0.83; SD = 0.07), (all trials: [t = 10.64, p < 0.001]; correct only: [t = 8.38, p < 0.001]).

Next, we tested whether medium and high uncertainty levels were significantly different. We found that the decision threshold was greater in medium uncertainty (all trials: M = 0.81; SD = 0.06; correct only: M = 0.83; SD = 0.07) than high uncertainty (all trials: M = 0.72; SD = 0.04; correct only: M = 0.74; SD = 0.04), (all trials: [t = 6.73, p < 0.001]; correct only: [t = 5.99, p < 0.001]).

3.3. Main Experiment

Eight of the original 73 participants were excluded as determined by predefined exclusion criteria (see section 2.7).

Model-Agnostic Analysis

We first tested whether the draws-to-decision were greater than 1 in each condition and found it was significant in high uncertainty (M = 3.02; SD = 1.88); [t = 8.68, p < 0.001], medium uncertainty (M = 3.02; SD = 1.64); [t = 9.95, p < 0.001], and low uncertainty (M = 1.88; SD = 1.01); [t = 6.99, p < 0.001].

We next tested whether low and medium uncertainty levels were significantly different. We found that participants took more samples in medium uncertainty (all trials: M = 3.02; SD = 1.62; correct only: M = 2.98; SD = 1.54) than low uncertainty (all trials: M = 1.88; SD = 1.00; correct only: M = 1.89; SD = 0.98), (all trials: [t = -8.09, p < 0.001]; correct only: [t = -7.91, p < 0.001]). These results support our prediction that increasing perceptual overlap between the distributions increases draws-to-decision.

Next, we tested whether medium and high uncertainty levels were significantly different. We found that participants did not take significantly more samples in high uncertainty (all trials: M = 3.02; SD = 1.86; correct only: M = 3.03; SD = 1.80) than

medium uncertainty (all trials: M = 3.02; SD = 1.62; correct only: M = 2.98; SD = 1.54), (all trials: [t < 0.01, p = 0.99]; correct only: [t = -0.43, p = 0.67]).

Model-Based Analysis

We tested whether the Bayesian decision thresholds were greater than 0.5 in each condition and found it was significant in high uncertainty (M = 0.68; SD = 0.06); [t = 23.57, p < 0.001], medium uncertainty (M = 0.77; SD = 0.08); [t = 29.26, p < 0.001] and low uncertainty (M = 0.98; SD = 0.06); [t = 67.53, p < 0.001].

We next tested whether low and medium uncertainty levels were significantly different. We found that the decision threshold was higher in low uncertainty (all trials: M = 0.98; SD = 0.06; correct only: M = 1.00; SD = < 0.001) than medium uncertainty (all trials: M = 0.77; SD = 0.07; correct only: M = 0.80; SD = 0.06), (all trials: [t = 27.38, p < 0.001]; correct only: [t = 24.43, p < 0.001]).

Next, we tested whether medium and high uncertainty levels were significantly different. We found that the decision threshold was higher in medium uncertainty (all trials: M = 0.77; SD = 0.07; correct only: M = 0.80; SD = 0.06) than high uncertainty (all trials: M = 0.68; SD = 0.06; correct only: M = 0.70; SD = 0.05), (all trials: [t = 17.71, p < 0.001]; correct only: [t = 20.06, p < 0.001]).

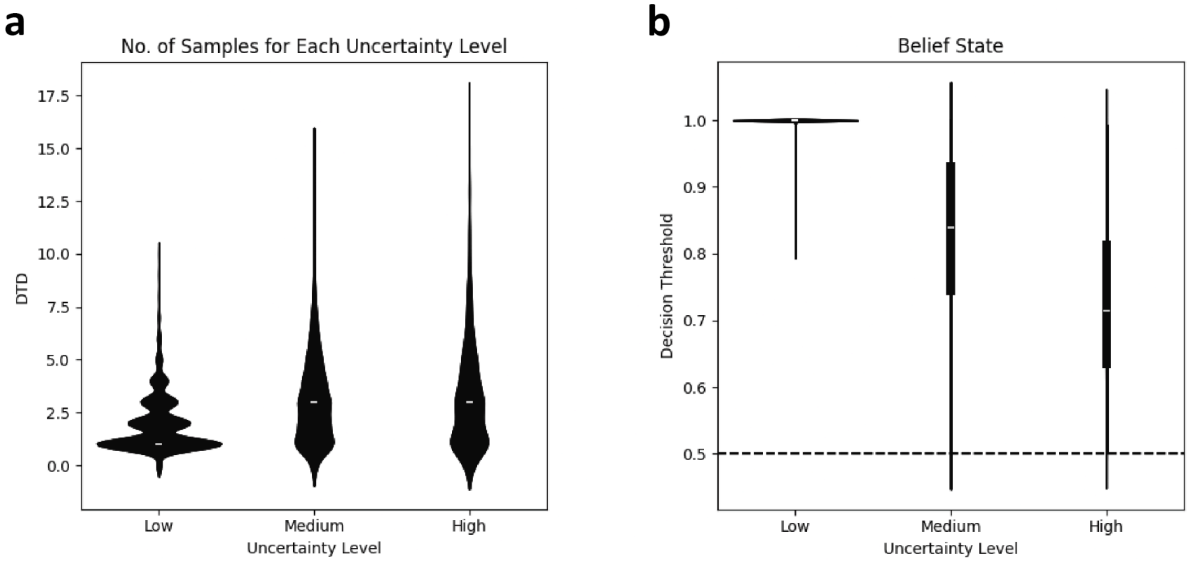

Fig. 5 | Draws-to-decision and decision threshold significantly different between medium and low uncertainty. Average across all 65 participants. **a**, Draws-to-decision at each uncertainty level. **b**, Decision threshold at each uncertainty level.

Questionnaires

Scores for the BDI-SF (M = 5.81; SD = 5.99), BAI (M = 8.06; SD = 9.82), and RRS (M = 34.97; SD = 13.00) were calculated following the criteria outlined in section 2.2.

For each participant, the mean number of draws-to-decision under low uncertainty was subtracted from the mean number of draws-to-decision under medium uncertainty, giving us the difference in draws-to-decision. We chose low and medium uncertainty for these analyses due to the large effect size. As before, we used data from correct trials only. We next tested whether the effect size of the relationship between the difference in draws-to-decision and questionnaire score was significant and that participants with higher depression scores exhibited less of a difference between medium and low uncertainty condition decision thresholds. We found the effect size was

not significant for the BDI-SF (rho = 0.266; p = 0.14), BAI (rho = 0.21; p = 0.238), or RRS (rho = -0.15; p = 0.407).

Next, the mean decision threshold under low uncertainty was subtracted from the mean decision threshold under medium uncertainty, giving us the difference in decision threshold. We next tested whether the effect size of the relationship between the difference in decision threshold and questionnaire score was significant. We found it was significant for the BDI-SF (rho = 0.47; $p < 0.01$) but not for the BAI (rho = 0.178; p = 0.328) or RRS (rho = -0.093; p = 0.61).

Finally, we tested whether the effect size of the relationship between questionnaire scores was significant. We found it was significant for the BDI-SF and BAI (rho = 0.446; p = 0.01), but not the BDI-SF and RRS (rho = -0.017; p = 0.92) or BAI and RRS (rho = -0.08; p = 0.66).

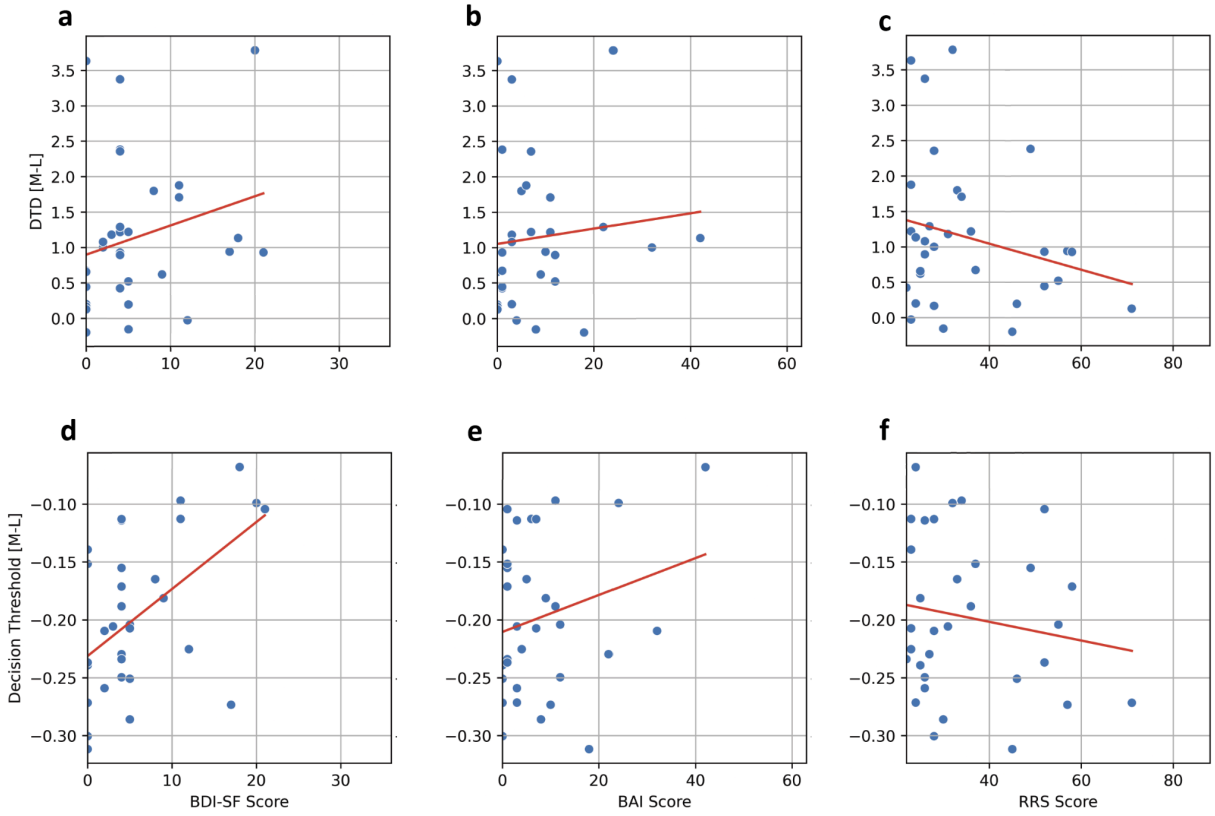

Fig. 6 | Relationship between depression score and decision threshold. a-c Difference in draws-to-decision as a function of questionnaire score. **d-f** Difference in decision threshold as a function of questionnaire score.

4. DISCUSSION

Using a novel online study designed to investigate sensitivity to uncertainty, we found an increase in draws-to-decision in all participants between low and medium uncertainty conditions. This result supports our hypothesis that increasing information ambiguity increases draws-to-decision.

Contrary to our main hypothesis, however, our results do not demonstrate a strong relationship between the difference in draws-to-decision and rumination symptom severity. This result is unlikely due to our sample population, as the adult sample in the literature that guided our battery of questionnaires exhibited generally lower scores on average (McLaughlin & Nolen-Hoeksema, 2010).

A key finding of our study is a relationship between depression score and decision threshold. We found that participants with higher depression scores exhibit less of a difference between medium and low uncertainty condition decision thresholds. That is to say, our results indicate that the mean decision threshold in the medium uncertainty condition was more similar to the mean decision threshold in the low uncertainty condition ($M = 0.98$) for participants with greater depression scores. This result suggests that more depressed participants required a higher degree of certainty before making a choice, possibly due to perceiving the samples as more ambiguous. We do, however, fail to see this result reflected in our model-agnostic approach. One potential explanation for this inconsistency is that the model-based approach is more sensitive to observed evidence.

Our results demonstrate no significant difference in sampling between medium and high uncertainty conditions. Our relatively small online sample size is one potential factor leading to this discrepancy. Another consideration is that there may not have

been a significant disparity between the medium and high uncertainty levels. Thus, an additional next step involves fine-tuning the standard deviation of the medium and hard distributions so that the difference between the two conditions is more perceptible.

In conclusion, when comparing low and medium uncertainty conditions, our results demonstrate some consistency with past studies that have found a need in clinical populations for a higher degree of certainty before making a choice (Banca et al., 2015). The specificity of the observed abnormalities in information sampling under uncertainty to trait depression suggests that the pathophysiological mechanisms underlying this symptom may be dissociable from those underlying comorbid disorders (e.g., generalized anxiety disorder), and thus selectively targetable by novel therapies.

REFERENCES

Baker, S. C., Konova, A. B., Daw, N. D., & Horga, G. (2019). A distinct inferential mechanism for delusions in schizophrenia. *Brain*, *142*(6), 1797–1812. https://doi.org/10.1093/brain/awz051

Banca, P., Vestergaard, M. D., Rankov, V., Baek, K., Mitchell, S., Lapa, T., Castelo-Branco, M., & Voon, V. (2015). Evidence Accumulation in Obsessive-Compulsive Disorder: The Role of Uncertainty and Monetary Reward on Perceptual Decision-Making Thresholds. *Neuropsychopharmacology*, *40*(5), 1192–1202. https://doi.org/10.1038/npp.2014.303

Beck, A. T. (1972). Measuring depression: The depression inventory. *Recent advances in the psychobiology of the depressive illnesses*, 299-302.

Beck, A. T., & Steer, R. A. (1990). Manual for the Beck Anxiety Inventory. San Antonio, TX: The Psychological Corporation.

Bedder, R. L., Pisupati, S., & Niv, Y. (2023). *Modelling Rumination as a State-Inference Process* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/tfjqn

Berg, M., Feldmann, M., Kirchner, L., & Kube, T. (2021). Oversampled and undersolved: depressive rumination from an active inference perspective.

Jacoby, R. J., Abramowitz, J. S., Reuman, L., & Blakey, S. M. (2016). Enhancing the ecological validity of the Beads Task as a behavioral measure of intolerance of uncertainty. *Journal of Anxiety Disorders*, *41*, 43–49. https://doi.org/10.1016/j.janxdis.2016.02.003

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. Artificial intelligence, 101(1-2), 99–134.

Lawrance, E. L., Gagne, C. R., O'Reilly, J. X., Bjisterbosch, J., & Bishop, S. J. (2022).

The Computational and Neural Substrates of Ambiguity Avoidance in Anxiety. *Computational Psychiatry*, *6*(1). https://doi.org/10.5334/cpsy.67

Maksimenko, V. A., Kuc, A., Frolov, N. S., Khramova, M. V., Pisarchik, A. N., & Hramov, A. E. (2020). Dissociating Cognitive Processes During Ambiguous Information Processing in Perceptual Decision-Making. *Frontiers in Behavioral Neuroscience*, *14*, 95. https://doi.org/10.3389/fnbeh.2020.00095

Martin, L. L., & Tesser, A. (1996). Some ruminative thoughts. In R. S. Wyer, Jr. (Ed.), *Ruminative thoughts* (pp. 1–47). Lawrence Erlbaum Associates, Inc.

McLaughlin, K. A., & Nolen-Hoeksema, S. (2011). Rumination as a transdiagnostic factor in depression and anxiety. *Behaviour research and therapy*, *49*(3), 186-193.

McLaughlin, K. A., & Nolen-Hoeksema, S. (2011). Rumination as a transdiagnostic factor in depression and anxiety. *Behaviour Research and Therapy*, *49*(3), 186–193. https://doi.org/10.1016/j.brat.2010.12.006

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. Perspectives on psychological science, 3(5), 400–424.

Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: the 1989 Loma Prieta earthquake. Journal of Personality and Social Psychology, 61,115e121.

Phillips, L. D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of experimental psychology*, *72*(3), 346.

Romero-Ferreiro, V., Susi, R., Sánchez-Morla, E. M., Marí-Beffa, P., Rodríguez-Gómez,

P., Amador, J., Moreno, E. M., Romero, C., Martínez-García, N., & Rodriguez-Jimenez, R. (2022). Bayesian reasoning with emotional material in patients with schizophrenia. *Frontiers in Psychology*, *13*, 827037. https://doi.org/10.3389/fpsyg.2022.827037

Schulz, L., Rollwage, M., Dolan, R. J., & Fleming, S. M. (2020). Dogmatism manifests in lowered information search under uncertainty. *Proceedings of the National Academy of Sciences*, *117*(49), 31527–31534. https://doi.org/10.1073/pnas.2009641117

Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, *90*(5), 927-939.

Sun, S., Yu, R., & Wang, S. (2017). A Neural Signature Encoding Decisions under Perceptual Ambiguity. *Eneuro*, *4*(6), ENEURO.0235-17.2017. https://doi.org/10.1523/ENEURO.0235-17.2017

Takano, K., & Tanno, Y. (2009). Self-rumination, self-reflection, and depression: Self-rumination counteracts the adaptive effect of self-reflection. *Behaviour research and therapy*, *47*(3), 260-264.

Ward, A., Lyubomirsky, S., Sousa, L., & Nolen-Hoeksema, S. (2003). Can't Quite Commit: Rumination and Uncertainty. *Personality and Social Psychology Bulletin*, *29*(1), 96–107. https://doi.org/10.1177/0146167202238375

Watkins, E. R., & Roberts, H. (2020). Reflecting on rumination: Consequences, causes, mechanisms and treatment of rumination. Behaviour Research and Therapy, 127, 103573

APPENDIX A

Berg et al. (2021) Model of Rumination

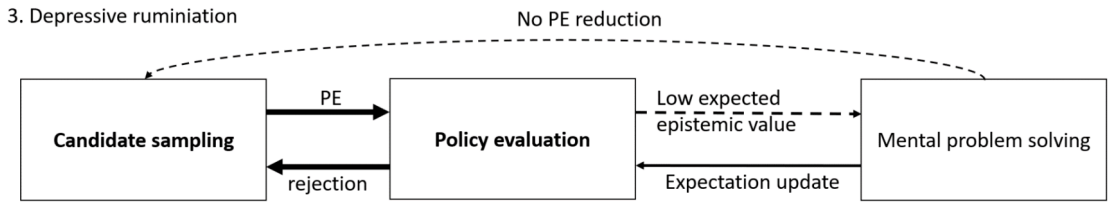

Fig. A | Berg et al. (2021) Model of rumination development and maintenance.

APPENDIX B

Participant Interface

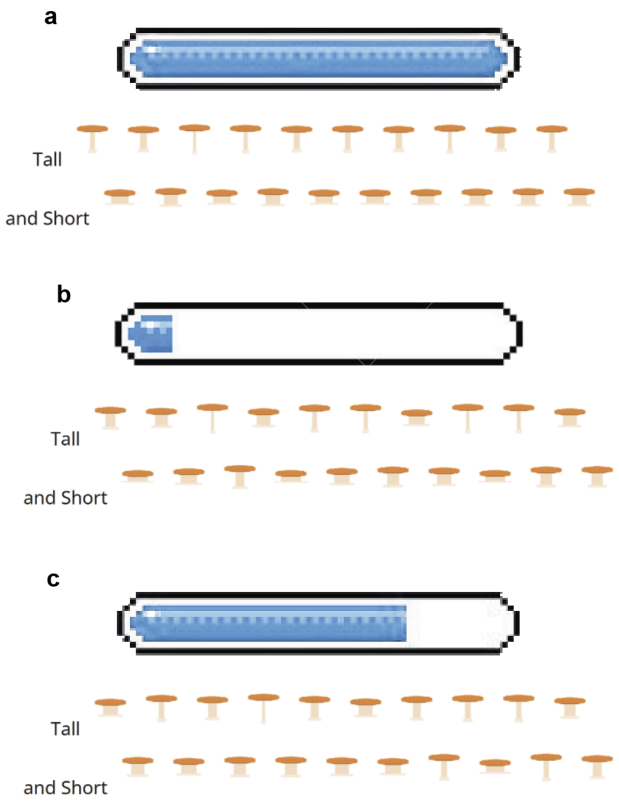

Fig. B1 | Maturity bars indicate maturity level and correspond to a particular degree of uncertainty. The three uncertainty levels, low, medium, and high, were communicated to participants as mushroom maturity, the options being either immature, mature, or elder, respectively. The maturity level for each block was indicated by a 'maturity bar,' which was visible to participants throughout the trial. **a**, A full bar indicates an elder (very mature) fungus. This maturity level corresponds to low uncertainty as the fungus species look very different. **b**, A short bar indicates an immature fungus. This maturity level corresponds to high uncertainty as the fungus species have similar-looking mushrooms. **c**, A two-thirds full bar indicates an immature fungus. This maturity level corresponds to medium uncertainty as the difference in mushroom size falls between the elder and immature species.

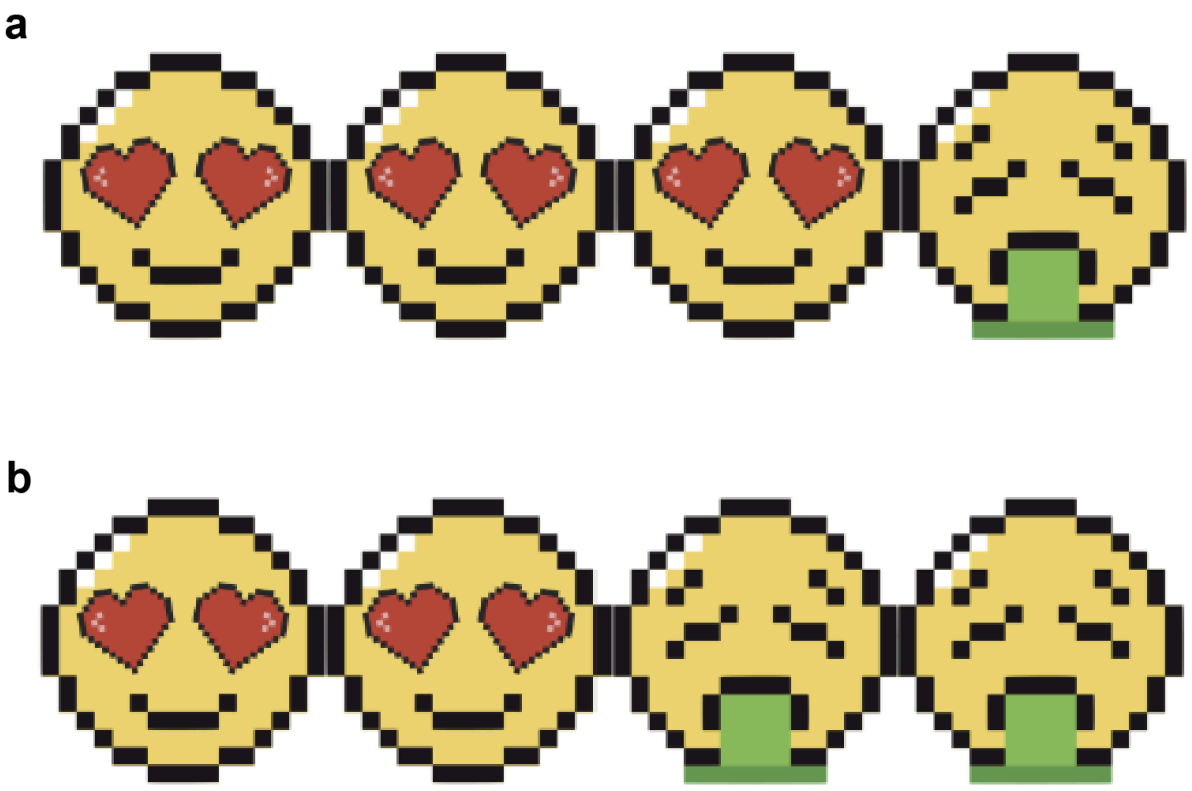

Fig. B2 | Potential rewards and punishment were communicated to participants using happy or sick emojis. Participants were informed that the fungi have different medicinal properties whose strength varies based on the forest (block). Reward magnitude was constant within each block, and was indicated by the amount of happy or sick emojis shown on screen, with the options being either 3,1 (high gain; a) or 2,2 (neutral reward; b).

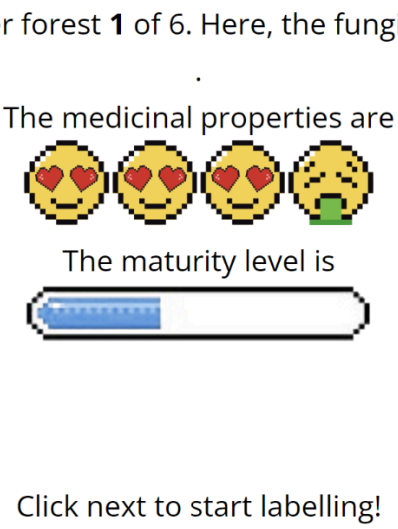

Fig. B3 | Screen indicating the beginning of a new block. Each block began with a description of the forest (i.e., the color of the stimuli), medicinal properties (i.e., potential reward magnitude), and maturity level (i.e., degree of uncertainty).

APPENDIX C

Pilot Figures

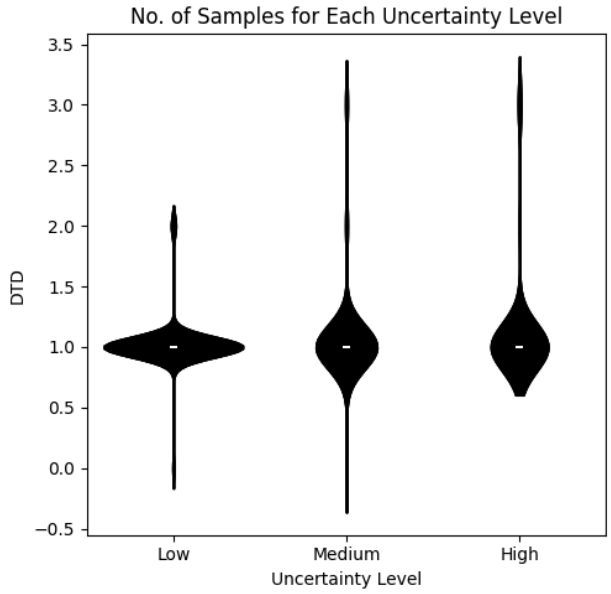

Fig. C1 | Pilot 1 results indicate participants were not cognizant of the option to sample multiple times. T-test run on Pilot 1 data (n = 5 participants) to determine whether draws-to-decision was significantly different than 1 revealed no significant difference for any uncertainty condition.

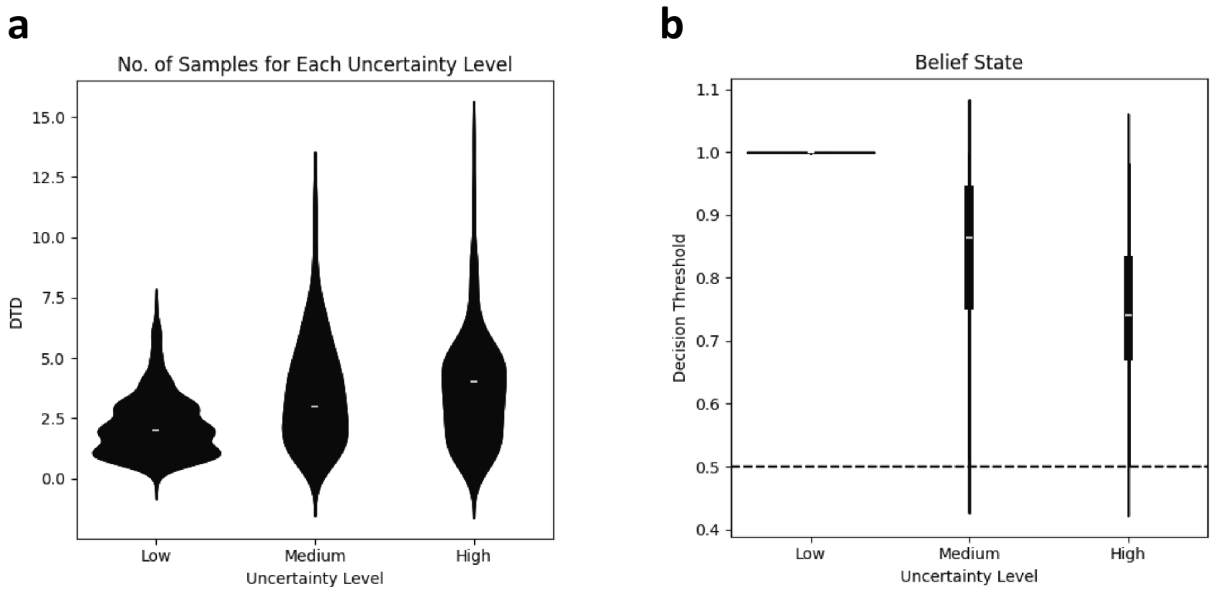

Fig. C2 | Changes implemented after the Pilot 1 study increased draws-to-decision. Average across all 13 participants in the Pilot 2 study. The horizontal axis represents the uncertainty level or perceptual overlap between the distributions. **a**, Draws-to-decision at each uncertainty level. **b**, Decision threshold at each uncertainty level.