A neural and behavioral tradeoff underlies exploratory decisions in normative anxiety

Kristoffer C. Aberg, Ido Toren, Rony Paz

Department of Neurobiology, Weizmann Institute of Science, Rehovot 7610001, Israel

Correspondence should be addressed to: K.A. (kc.aberg@gmail.com) or R.P (rony.paz@weizmann.ac.il)

Abstract

Exploration reduces uncertainty about the environment and improves the quality of future decisions, but at the cost of provisional uncertain and suboptimal outcomes. Although anxiety promotes intolerance to uncertainty, it remains unclear whether and by which mechanisms anxiety modulates exploratory decision-making. We use a dynamic three-armed-bandit task and find that higher trait-anxiety increases exploration, which in turn harms overall performance. We identify two distinct behavioral sources: first, decisions made by anxious individuals were guided towards reduction of uncertainty; and second, decisions were less guided by immediate value gains. Imaging (fMRI) revealed that anxiety correlated negatively with the representation of expected-value in the dorsal-anterior-cingulate-cortex, and in contrast, positively with the representation of uncertainty in the anterior-insula. Moreover, the findings were similar in both loss and gain domains, and demonstrated that affective trait modulates exploration and results in an inverse-U-shaped relationship between anxiety and overall performance. We conclude that a shift in balance towards representation of uncertainty in the insula prevails over reduced value representation in the dACC, and entails maladaptive decision-making in individuals with higher normal-range anxiety.

Introduction

Exploratory decision-making concerns the balance between exploiting options with known outcomes, such as dining in your favorite restaurant; and exploring options with uncertain but potentially better outcomes, such as sampling a novel restaurant. Exploration provides information about the environment and therefore provides a prospective benefit to decision-making by reducing the uncertainty associated with available options. For example, the benefits and costs of eating at a new restaurant are largely unknown until it has been explored, at which point the newly gathered information can be used to guide subsequent decisions (whether to return or not). Exploratory decisions are prevalent in everyday life and involve decisions ranging from small to large and even potentially life-changing ones, as in selecting a partner or employment (Cohen, McClure, and Yu 2007; Addicott et al. 2017; Mehlhorn et al. 2015). Because exaggerated exploration impairs the ability to exploit the good familiar options, whereas overly restrained exploration prevents the discovery of better options, the ability to maintain an appropriate balance has a significant influence on an individual's well-being and quality of life. In extreme cases, it might even contribute to psychopathologies (Addicott et al. 2017; Scholl and Klein-Flugge 2018).

Imaging studies in humans have shown that exploratory decisions engage the anterior insulae (aINS) as well as prefrontal regions involved in executive control, such as the frontopolar cortex (FPC) and the dorsal anterior cingulate cortex (dACC) (Daw et al. 2006; Badre et al. 2012; Cavanagh et al. 2012; Kayser et al. 2015; Laureiro-Martinez et al. 2015; Blanchard and Gershman 2018; Chakroun et al. 2020; Tomov et al. 2020), with recent research suggesting further differentiation between neuronal correlates related to 'directed' and 'random' exploration strategies (Tomov et al. 2020). Importantly, it was shown that anxious individuals find uncertainty more aversive and are also more intolerant thereof (Hartley and Phelps 2012; Buhr and Dugas 2009; Grupe and Nitschke 2013; Maner and Schmidt 2006; Maner et al. 2007; Charpentier et al. 2017), and recent studies have shown that aspects of uncertainty correlate with a hyper-responsive aINS in anxiety (Grupe and Nitschke 2013; Paulus and Stein 2006). In addition, anxiety also pertain to dysfunctional executive control processes in the prefrontal-cortex (PFC) which are hypothesized to induce attentional biases towards salient events (Etkin and Wager 2007; Bishop 2007).

These studies suggest that anxiety may have a particularly relevant influence on exploratory decision-making. On one hand, anxiety could decrease exploration because it requires selecting options with uncertain outcomes; yet on the other hand, anxiety could increase exploration by reducing uncertainty in the environment. Previous studies have examined the relationship between uncertainty and anxiety (Grupe and Nitschke 2013) and separately examined the relationship between uncertainty and exploration (Badre et al. 2012; Cavanagh et al. 2012; Kayser et al. 2015; Tomov et al. 2020). However, little is known about the direct relationship between anxiety and exploratory decision-making, the cognitive factors that contribute to it, and the underlying brain mechanisms. Moreover, although outcome valence (appetitive – aversive) was shown to interact with inter-individual factors and anxiety levels (Browning et al. 2015; Laufer, Israeli, and Paz 2016; Xu et al. 2013; Aberg, Doell, and Schwartz 2015, 2016b; Frank, Seeberger, and O'Reilly R 2004; Aberg, Muller, and Schwartz 2017), and activity in the dACC and aINS (Skvortsova, Palminteri, and Pessiglione 2014; Pessiglione and Delgado 2015), little is known about the interaction between these factors and exploratory decisions.

Taken together, we hypothesized that anxiety-levels would induce a bias in exploratory decisions, and that it can be mediated by shifts in saliency to uncertainty and to expected-outcomes, induced by a hyper-responsive aINS and/or by low recruitment of the FPC/dACC.

Results

Participants played a three-armed bandit task where in each trial they chose one of three slot machines (Fig.1A,B). After selection, the outcome currently associated with the selected machine was displayed. Exploration was promoted by inducing a dynamic environment in which the outcome associated with each machine varied across trials: each arm was associated with a sine-like wave of similar magnitude and mean, but with different period lengths (100, 67, or 33 trials), and phase-shifts (periods of 0.10, 0.45, and 0.90; Fig.1C). To examine differences in exploratory strategies between aversive and appetitive contexts, participants either accumulated points in a Gain condition where all outcomes are positive (Fig.1A), or avoided losing points in a Loss condition where all outcomes are negative (Fig.1B).

We first performed a separate behavioral study to test the hypothesis, and found a confirmative positive correlation between trait anxiety and exploration (Supplementary Note 1). Following this, we conducted the independent main behavioral-fMRI study, and examined the interaction across the range of anxiety (Trait-anxiety scores: mean=36.321, SD=9.982; see (Spielberger et al. 1983), Table 1). Normative anxiety is a major factor in modern lives affecting daily function, and enabled us to examine the gradual modulation of anxiety on exploration without the traditional focus and confounds imposed by clinical populations.

Exploration increases with anxiety

We examined the relationship between Trait anxiety and exploratory decision-making, and whether it interacts with outcome valence (Loss/Gain). As in previous studies (Tzovara et al. 2012; Daw et al. 2006), an exploratory decision was defined as the selection of a machine with a suboptimal expected value, if the selected machine was also different from the machine selected in the previous trial. Trial-by-trial estimates of expected values were estimated by a behavioral model which provided the most parsimonious fit to overall-behavior, as determined by confronting different types of behavioral models commonly used to model behavior and exploration in multi-armed bandit tasks (see below, Methods, and Supplementary Note 5). Infrequent explorers persevered in selecting a suboptimal machine (Fig.1D,E; *suboptimal persistence*), whereas frequent explorers more often rejected the optimal machine and chose a different one (Fig.1F,G; *suboptimal switching*).

We found that exploration, quantified as the proportion of exploratory decisions across all trials, was modulated by anxiety (Fig.2A; repeated measures two-way ANOVA, main effect of anxiety as linear covariate, $F(1,26)=24.98$, $p<0.001$, $\eta_p^2=0.490$, ANOVA; Loss: $r=0.707$, $p<0.001$; Gain: $r=0.573$, $p<0.001$, two-tailed Pearson correlations; no valence*anxiety interaction effect, $F(1,26)=1.85$, $p=0.19$, $\eta_p^2=0.066$, ANOVA). Therefore, more anxious individuals explored more.

Overall performance is inverse-U-shape modulated by exploration

In a framework of varying outcomes, it is expected that low exploration would prevent the discovery of changes and better alternatives (Infrequent explorers; Fig.1D,E), and on the other hand, high exploration would hinder the exploitation of the current optimal alternative (Frequent explorers; Fig.1F,G). Taken together, this should lead to an inverse-U-shape form of overall performance.

Indeed, both infrequent and frequent explorers made more suboptimal decisions, and as a result overall performance was non-linearly modulated by the proportion of exploratory decisions (Fig.2C). This was the

case when data was collapsed across conditions (Fig.2B, left panel; nonlinear regression with intercept and standardized quadratic term $b^2_{Explore}$=-0.231, p<0.001), and separately in both the Loss condition (Fig.2B, top right panel; nonlinear regression with intercept and standardized quadratic term $b^2_{Explore}$=-0.293, p<0.001) and in the Gain condition (Fig.2B, bottom right panel; standardized $b^2_{Explore}$=-0.200, p<0.05). There was no significant difference between conditions (Actual difference: $b^2_{Explore}$ Loss-Gain=-0.093; Null distribution difference: $b^2_{Explore}$ Loss-Gain=-0.040, p=0.198, Monte-Carlo randomization test, Supplementary Note 2). Therefore, our findings demonstrate that the individual tendency to explore modulates overall performance in an inverted U-shape fashion, as previously suggested (Addicott et al. 2017; Aston-Jones and Cohen 2005).

Exploration mediates the modulation of overall performance by anxiety

If exploration increases with anxiety and overall performance is non-linearly modulated by exploration, then one can hypothesize a relationship between anxiety and performance. We tested this directly and found that overall performance was also significantly and non-linearly modulated by anxiety across conditions (Fig.2C, left panel, nonlinear regression with intercept and standardized quadratic term $b^2_{Anxiety}$=-0.591, p<0.001), but only significantly so in the Loss condition (Fig.2C, top right panel, nonlinear regression with intercept and standardized quadratic term $b^2_{Anxiety}$=-0.712, p<0.001), and not in the Gain condition (Fig.2 C, bottom right panel, standardized $b^2_{Anxiety}$=-0.228, p=0.198), with a significant difference between conditions (Actual difference: $b^2_{Anxiety}$ Loss-Gain=-0.483; Null distribution difference: $b^2_{Anxiety}$ Loss-Gain=-0.258, p=0.032, Monte-Carlo permutation test, Supplementary Note 3).

To establish further that it is the exploration that mediated the relationship between anxiety and overall performance, we tested whether the relationship between anxiety and performance is reduced when controlling for the contribution of exploration (Hayes and Rockwood 2017). The relationship changed significantly when controlling for Exploration across conditions ($\Delta b^2_{Anxiety}$=-0.190, p<0.001, Monte-Carlo randomization test), and separately in both the Loss ($\Delta b^2_{Anxiety}$=-0.152, p<0.001, Monte-Carlo randomization test) and in the Gain condition ($\Delta b^2_{Anxiety}$=-0.260, p<0.001, Monte-Carlo randomization test; see Fig.2D and Supplementary Note 4 for details). Therefore, exploration contributes to the non-linear relationship between anxiety and overall performance.

Behavioral factors contributing to exploration

We examined the relative contribution of different behavioral factors by adapting a reinforcement-learning approach. These models track the expected-value of each machine and update it by the mismatch between the expected-value and the actual outcome (i.e. the prediction-error scaled by the learning rate. Methods, Equations 1,2,4). In addition, we included the following factors: First, because recent evidence suggests that exploration is aimed at reducing the outcome-variability associated with different options ('directed' exploration in static environments: (Badre et al. 2012; Frank et al. 2009; Tomov et al. 2020; Gershman 2018), we included an outcome-variability factor which was updated by the unsigned prediction-error scaled by a learning-rate (Equations 5-6). Second, in a dynamic environment the uncertainty associated with a machine's outcome increases with the time since it was last selected, and we therefore included an outcome-uncertainty factor as an additional way of modeling 'directed' exploration, defined as the number of trials since a machine was last selected (Equation 7). Third, an exploratory strategy that allows detecting changes in the environment is randomly switching machines between trials, and to capture such tendencies we included a random binary switch factor (Equation 8). Notice this is the inverse of 'stickiness' or 'perseveration', which models the tendency to repeat a previously chosen action (Chakroun et al. 2020; Gershman, Pesaran, and Daw 2009).

Finally, 'random' exploration in static environments was found to depend on the total uncertainty of the environment (Gershman 2018; Gershman and Tzovaras 2018), and we therefore included a term which scales the expected-value of each machine by the total outcome-variability of the environment (Equation 9).

In addition to the models described above, we also tested two common models: one with an ε-greedy decision rule in which suboptimal alternatives were selected with probability ε (Equation 11), as well as Kalman-filter model with dynamic learning rates that was previously used to disentangle 'directed' and 'random' exploration in static environments (Equations 12-15).

After the models were fitted individually to each participant's behavior, the probability of selecting machine *i* in trial t depends on weights β_Q, β_U, β_T, β_S, and β_{tU} reflecting the contribution of expected-value $Q_i(t)$, outcome-variability $U_i(t)$, outcome-uncertainty $T_i(t)$, random-switching $S_i(t)$, and 'random exploration' $Q_i(t)/tU_i(t)$, respectively. The goodness-of-fit of these models were compared using a Bayesian Model Selection (BMS) procedure (Stephan et al. 2009), with Akaike's Information Criterion corrected for small sample sizes as model evidence (AICc; (Burnham and Anderson 2004)), and further validated using the Bayesian Information Criterion (BIC) approximation of the log marginal likelihood (Bishop 2006; Gershman 2018), which provides a larger penalty term for more complex models.

We found that the most parsimonious model was the same when collapsed across valence conditions (Fig.3A, protected exceedance probabilities > 0.997; Supplementary Note 5) as well as for each loss/gain condition separately (Fig.3B). This model included decision weights for expected-value (β_Q), outcome-uncertainty (β_T), and random-switching (β_S) and it is therefore termed the QTS model hereafter. Importantly, the QTS model was found to be the most parsimonious one also in the independent behavioral pilot study (Supplementary Note 1). Next, to ensure that the fitted parameters were meaningful, we applied a parameter recovery procedure which tests whether model-parameters used to generate behavior can be successfully recovered when re-fitting the same model to the generated behavior (Wilson and Collins 2019). Indeed, all correlations between generating and recovered parameters were highly significant (Fig.3C; all two-sided Pearson's r>0.827, p-values<0.001).

As an additional validation, individually fitted parameters were used to calculate the utility of each machine in each trial and the machine with the highest utility was defined as the model's prediction. This revealed high similarity between actual and model-based selections (Fig.3D-F). Across participants, the model predicted actual selections well above chance-level (i.e. 0.333) in both the Loss (mean correct±SEM: 0.801±0.02) and the Gain condition (mean correct±SEM: 0.822±0.016).

Because the model was deliberately not fitted to exploratory behaviors per-se (to maintain independence), we tested that the model captured relevant aspects of exploration (Wilson and Collins 2019; Palminteri, Wyart, and Koechlin 2017), by calculating a model-derived probability of switching to a suboptimal machine across all trials and conditions. Reproducing the behavioral results, this model-derived probability of exploration was modulated by anxiety (repeated measures two-way ANOVA, main effect of anxiety as linear covariate, Fig.3G, left panel; F(1,26)=27.52, p<0.001, η_p^2=0.514, ANOVA and no valence*anxiety interaction effect, F(1,26)=0.16, p=0.696, η_p^2=0.006, ANOVA). Moreover, the model-derived probability of exploration correlated positively with anxiety in both the Loss (Fig.3G, top right panel; r=0.630, p<0.001, two-tailed Pearson correlation) and the Gain condition (Fig.3G, bottom right panel; r=0.628, p<0.001, two-tailed Pearson correlation). These results mirror the aforementioned behavioral results (Fig.2A).

Exploration in anxiety trades off value for uncertainty reduction

We tested how the different factors contribute to exploratory decisions and their modulation by anxiety. We repeated the analysis separately for each of the four different factors (namely, for their contribution as quantified by $\beta_Q, \beta_T, \beta_S, \text{and } \alpha_Q$; Supplementary Note 6). All four factors were significant predictors of behavior (Fig.4A; all p-values for their respective ANOVA intercept terms<0.001; Supplementary Table 3), but only two of them correlated with individual anxiety levels.

First, expected-value (β_Q) was significantly modulated by anxiety (Fig.4B, left panel; F(1,26)=8.41, p=0.008, η_p^2=0.24, ANOVA; Supplementary Table 3). Specifically, anxiety was negatively correlated with expected-value in both the Loss (Fig.4B, top right panel; r=-0.421, p=0.026, two-tailed Pearson-correlation) and the Gain condition (Fig.4B, bottom right panel; r=-0.424, p=0.025, two-tailed Pearson-correlation; Supplementary Table 3). Additionally, expected-value was negatively correlated with the proportion of exploratory decisions in both the Loss (Supplementary Fig.7A, left panel; r=-0.704, p<0.001, two-tailed Pearson-correlation) and the Gain condition (Supplementary Fig.7A, right panel; r=-0.585, p=0.005, two-tailed Pearson-correlation). This implies that the expected value of different options has a less pronounced impact on decisions made by more anxious individuals.

Second, outcome-uncertainty (β_T) was also significantly modulated by anxiety (Fig4C, left panel; F(1,26)=10.36, p=0.003, η_p^2=0.285, ANOVA; Supplementary Table 3), but in contrast to the negative relationship observed between anxiety and expected-value, outcome-uncertainty was positively correlated in both the Loss (Fig.4C, top right panel; r=0.446, p=0.018, two-tailed Pearson correlation) and the Gain condition (Fig.4C, bottom right panel; r=0.431, p=0.022, two-tailed Pearson correlation). Similarly, outcome-uncertainty correlated positively with the proportion of exploratory decisions in both Gain and Loss (Supplementary Fig.7B; Loss: r=0.412, p=0.029; Gain: r=0.589, p<0.001, two-tailed Pearson correlations). These results imply that more anxious individuals showed a stronger tendency to reduce uncertainty.

Notably, outcome-uncertainty was not correlated with expected-value (Fig.4D; Collapsed: r=-0.252, p=0.19; Loss: r=-0.075, p=0.71; Gain: r=-0.035, p=0.86, two-tailed Pearson correlation), but anxiety was significantly correlated with the difference between these two factors (namely, β_Q-β_T), across condition (Fig.4E, left panel; r=-0.650, p=0.001, two-tailed Pearson correlation), and separately in both the Loss (Fig.4E, top right panel; r=-0.583, p=0.001, two-tailed Pearson correlation) and the Gain condition (Fig.4E, bottom right panel; r=-0.593, p<0.001, two-tailed Pearson correlation).).

Finally, anxiety did not correlate with random-switching (β_S, Supplementary Table 3; Supplementary Fig.6A,B), or with the learning rate (α_Q, Supplementary Table 3; Supplementary Fig.6C,D), and neither factors (β_S or α_Q) correlated with the proportion of exploratory decisions (Supplementary Fig.7C,D).

To further assess the robustness of these results, we tested the fitted model-parameters of a Kalman-model with a dynamic learning rate, but with the same decision weights as the QTS model (i.e. β_Q, β_T, and β_S; see Methods). Although this model showed an inferior fit to behavior compared to the QTS model (Fig.3A; Supplementary Note 5), the main results were replicated revealing positive correlations between anxiety and the proportion of exploratory decisions, and positive correlations between trait anxiety and β_T (but not with β_Q; for a summary of these results, see Supplementary Note 7).

Taken together, the results suggest that increased levels of anxiety enhance exploration by trading off immediate expected-value gains for reduction in uncertainty.

Encoding of prediction-error is not modulated by anxiety

We first tested whether anxiety modulated the neuronal correlates of value learning. As expected from multiple previous studies, activity in the bilateral NAcc correlated significantly with prediction-errors (δ_Q; Supplementary Note 8; left NAcc: peak voxel MNI xyz = -9 11 -5, *t*(27)=5.468, p<0.001; FWE, SVC; right NAcc: peak voxel MNI xyz = 9 11 -5, *t*(27)=6.013, p<0.001; FWE, SVC). However, neither anxiety nor valence (gain/loss) impacted the neuronal correlate of predictions-errors (β_δ for NAcc ROIs, ANOVAs with within-subject factor valence and anxiety as continuous covariate; Supplementary Note 8). Importantly, these brain activations, in combination with the non-significant correlations between anxiety and learning rates (Supplementary Note 6), suggest that trait anxiety did not affect learning performance in the present task.

Functional localization of brain regions involved in exploration

To examine the neural representation of expected-value and outcome-uncertainty, we first derived functional ROIs by independently localizing brain regions engaged by exploration-exploitation decisions on a group-level (Vul et al. 2009). In accordance with previous studies (Blanchard and Gershman 2018; Daw et al. 2006; Laureiro-Martinez et al. 2015), exploration was associated with increased activity in the dACC (Fig.5A,B; peak voxel MNI xyz = 9 20 45, *t*(27)=9.717, p<0.001, one-tailed *t*-test FWE SVC), the bilateral aINS (Fig.5C-E; Left aINS: peak voxel MNI xyz = -30 23 3, *t*(27)=8.247, p<0.001; Right aINS: peak voxel MNI xyz = 33 23 3, *t*(27)=9.099, p<0.001, one-tailed *t*-test FWE SVC), and the right FPC (Fig.5F,G; peak voxel MNI xyz = 27 53 10, *t*(27)=5.724, p=0.004, one-tailed *t*-test FWE SVC). On the other hand, exploitation was associated with increased activity in the vmPFC (Fig.5H-I; peak voxel MNI xyz = 0 44 -15, *t*(27)=6.452, p<0.001, one-tailed *t*-test FWE SVC).

We then used individual trial-by-trial differences in expected-value (ΔQ) and outcome-uncertainty (ΔT), between the selected machine and the average of the two rejected machines, as parametric modulators to identify their respective brain correlates ($\beta_{\Delta Q}, \beta_{\Delta T}$). Activity estimates were extracted from 3mm radius spheres centered on the peak voxels identified above (for robustness and validation, analyses were repeated using spheres with different radii of 0, 3, 5, and 10mm, all producing the same results described below, see Methods), and were used in separate repeated measures ANOVAs with within-subject factor valence and anxiety as a continuous covariate.

Anxiety modulates the representation of expected-value in the dACC

In the dACC, anxiety correlated significantly with the representation of expected-value $\beta_{\Delta Q}$ (Fig.6B, left panel F(1,26)=10.242, p=0.004, η_p^2=0.283, ANOVA), but not in any other region (Supplementary Note 9). Specifically, anxiety correlated negatively with $\beta_{\Delta Q}$ in both the Loss (Fig.6B, top right panel; r=0.375, p=0.049, two-tailed Pearson correlation) and in the Gain condition (Fig.6B, bottom right panel; r=-0.603, p<0.001, two-tailed Pearson correlation). Across participants, the average value of $\beta_{\Delta Q}$ was positive (F(1,26)=60.594, p<0.001, ANOVA intercept term; Supplementary Table 5). Because $\beta_{\Delta Q}$ reflects the coupling between BOLD signal in the dACC and expected value differences ΔQ (Fig.6C), these results mean that participants with a large $\beta_{\Delta Q}$ (e.g. low anxiety) show an increased engagement of the dACC when ΔQ decreases, while the dACC of more anxious individuals is less responsive to changes in ΔQ ($\beta_{\Delta Q}$ close to zero).

We tested if the reduced tracking of expected-value differences in the dACC can be attributed to an overall increase in dACC activity, which could indicate a limited ability to encode task-related parameters due to dACC saturation. However, overall dACC activation did not correlate with anxiety, suggesting against this possibility

(Supplementary Note 10). Additionally, anxiety did not affect the representation of expected-value in the vmPFC (Supplementary Note 9), a brain region directly involved in value representations (Levy and Glimcher 2012; Domenech and Koechlin 2015).

Overall, anxiety modulates the neuronal representation of expected-value differences in the dACC. This result parallels and extends the behavioral finding that decisions made by more anxious individuals were less influenced by the immediate expected-value.

Anxiety modulates the representation of outcome-uncertainty in the aINS

In the aINS, anxiety correlated significantly with the representation of outcome-uncertainty difference $\beta_{\Delta T}$ (Fig.6E, left panel; left aINS: F(1,26)=15.432, p<0.001, η_p^2=0.372, ANOVA; right aINS: F(1,26)=11.855, p=0.002, η_p^2=0.313, ANOVA), but not in any other region (Supplementary Note 11). Here, anxiety was positively correlated with the representation of outcome-uncertainty in the aINS in both the Loss (Fig.6E, top right panel; left aINS: r=-0.587, p=0.001; right aINS: r=-0.492, p=0.008, two-tailed Pearson correlations) and the Gain condition (Fig.6E, bottom right panel; left aINS: r=-0.410, p=0.030; right aINS: r=-0.345, p=0.028, two-tailed Pearson correlation). To understand this result, consider that a positive value of $\beta_{\Delta T}$ indicates increased BOLD signal for an increased uncertainty ΔT (Fig.6F). Therefore, the aINS of more anxious individuals (large positive $\beta_{\Delta Q}$) is more responsive to increases in uncertainty.

Overall, anxiety modulates the neuronal representation of outcome-uncertainty differences in the bilateral aINS. This result parallels and extends the behavioral finding that decisions made by more anxious individuals were more influenced by outcome-uncertainty.

The neuronal trade-off mediates the behavioral trade-off

Finally, we tested whether the behavioral tradeoff between expected-value and uncertainty-reduction observed in more anxious individuals (i.e. $\beta_Q - \beta_T$; Fig.4E) is related to a differential representation of expected-value in the dACC (Fig.6B) and outcome-uncertainty in the aINS (Fig.6D). We found that the strength of the relationship between anxiety and the difference between expected-value and uncertainty-reduction, decreased significantly when controlling for the differential representation in the two brain regions. This was the case in both the Loss and the Gain condition (Loss: $\Delta b_{Anxiety}$=-0.254, p<0.001, Gain: $\Delta b_{Anxiety}$=-0.292, p<0.001; Mediation analysis with Monte-Carlo randomization test; Supplementary Note 12).

Overall, these results indicate that the behavioral tradeoff between expected-value and uncertainty-reduction, which determines the anxiety-related exploration rate, is mediated by how the brain represents expected-value in the dACC versus outcome-uncertainty in the aINS.

Discussion

We show here that individuals with higher levels of trait anxiety tend to explore more in a dynamic decision-making environment, and do so in a similar manner in both Gain and Loss conditions. This exploration-exploitation bias was associated with two distinct neuro-behavioral mechanisms. First, more anxious individuals were less influenced by immediate expected-values and showed a reduced correlation between activity in the dACC and expected-value differences. Second, more anxious individuals were more influenced by outcome-uncertainty and showed an increased correlation between outcome-uncertainty and activity in the aINS (see Fig.6G-I for a schematic summary of the results). Taken together, anxiety boosts exploration by trading off immediate value gains in order to reduce the uncertainty of future decisions, paralleled by a tradeoff in dACC-aINS activity.

Our results provide evidence that individual anxiety levels contribute to the trade-off between immediate benefits, namely gaining a reward or avoiding a risky/uncertain situation, and prospective decision making. In the current study, as in most real-life situations, exploration provides information that benefits future decision making. Our findings therefore extend on previous studies that found that anxiety increases aversion for uncertainty, both in terms of risk (i.e. known uncertainty) and ambiguity (i.e. unknown uncertainty; (Grupe and Nitschke 2013; Hartley and Phelps 2012). In these previous studies, however, there was no information to be gained from making a risky decision (e.g. (Charpentier et al. 2017)). Taken together, the evidence suggests that whereas anxiety decreases risk-taking in an environment where all information is provided and nothing can be learned, anxiety increases the willingness to confront uncertainty via increased exploration in a dynamic environment, where risky decisions provide prospective benefits in the form of information gains and/or uncertainty reductions.

Here, anxious individuals were more sensitive to a particular type of uncertainty, one that depends on the dynamics of the environment, and which is not present in static environments (e.g. where reward contingencies do not change across trials). Related to this finding, it has been shown that exploration depends on the overall uncertainty of the environment, such that exploration increases in more volatile and uncertain settings (Knox et al. 2011; Gershman 2018; Gershman and Tzovaras 2018). Combined with the present results, one possible hypothesis is that anxious individuals perceive the world as being more volatile and therefore needs to be explored more frequently. Previous studies reported that trait anxiety and internalizing psychopathology decreased the ability to adjust learning rates as a function of volatility (Browning et al. 2015; Gagne et al. 2020), and it has been suggested that anxious individuals misestimates uncertainty (Pulcu and Browning 2019). Indeed, one simple heuristic which can be used to maintain a low level of uncertainty in a decision-making environment is to explore frequently, independently of actual volatility levels.

Of note, neither outcome-variability nor total outcome-variability contributed to exploration in the present study. This finding is consistent with results obtained in previous studies using similar 'restless n-armed bandit tasks', e.g. dynamic decision making tasks (Daw et al. 2006; Payzan-LeNestour and Bossaerts 2011), but different from findings in other tasks where reward-contingencies were static and outcome-variability was suggested to drive directed exploration, such as in the 'clock-task' (Badre et al. 2012; Frank et al. 2009), or specific instances of two-armed bandit-tasks in which direct and random exploration was related to outcome-variability and total outcome-variability (Tomov et al. 2020; Gershman 2018; Gershman and Tzovaras 2018). One reason for the differences could be that dynamic environments require the continuous monitoring and updating of values, and their respective uncertainties, across all trials of an experiment, and therefore entails a heavier working memory load. By contrast, in static environments most of the learning occurs within the first few trials. Indeed, reinforcement learning has been shown to interact with working memory, such that the ability to make value-based decisions depends on working memory load (Collins and Frank 2012). As such,

the reduced influence of outcome-variability on decision making in dynamic environments could be either due to a degraded/noisier representation of uncertainty in the working memory, and/or the application of simpler decision-making strategies that reduces working memory load, such as keeping track of only expected-value but not outcome-variability.

While directed and random exploration was not related to outcome-variability in the present study, high trait anxiety increased exploration via more uncertainty-reduction and by decreasing the weighting of expected-value. An alternative interpretation of the present results therefore, framed in the context of directed and random exploration, is that increased uncertainty-reduction can be regarded as a form of 'directed' exploration because it reduces the uncertainty of a particular option, while a reduced focus on expected-value could be seen as a form of 'random' exploration because it allows deviations from purely greedy decisions. In other words, high trait anxiety may increase both directed and random exploration in dynamic environments.

The positive correlations between uncertainty-reduction and anxiety are in accordance with the previously described link between anxiety and intolerance to uncertainty (Buhr and Dugas 2009; Dar, Iqbal, and Mushtaq 2017), and the idea that anxiety is underpinned by the fear of the unknown (Carleton 2016). Yet, whether the fear of the unknown is relevant only in contexts with aversive outcomes (as opposed to also in positive contexts) is currently under debate (Einstein 2014; Grupe and Nitschke 2013). Our results, showing equal contributions from uncertainty in Gain and Loss conditions, suggest that it is the uncertainty itself, and not its interaction with outcome valence, that promotes exploration. Corroborating this notion, a recent study exposed participants to different types of written real-life scenarios with uncertain positive or negative outcomes, and revealed similar correlations between trait anxiety and aversion to uncertainty in both scenarios (Pepperdine, Lomax, and Freeston 2018), further supporting the notion that anxiety is associated with a general aversion to uncertainty. An interesting extension of these findings could be that changing an individual's intolerance to uncertainty in a positive context, for example via cognitive bias modification (Oglesby, Allan, and Schmidt 2017), may cause a general reduction in intolerance to uncertainty that transfers to negative contexts.

At the neuronal level, trait anxiety correlated negatively with the representation of differences in expected-value in the dACC, yet this activity increased overall across participants when expected-values became increasingly similar. This latter result is in accordance with findings obtained in foraging paradigms, where the dACC was found to track the relative difference in expected-value between continued exploitation and foraging (defined as a decision to abandon gradually depleting options; (Kolling et al. 2012; Shenhav et al. 2014)). Foraging theory posits that a decision to forage is optimal when the expected value difference between an exploited option and foraging options is small (Hayden, Pearson, and Platt 2011; Charnov 1976). Likewise, exploration is arguably optimal when the expected values of options are similar, because at this point additional information is needed to deduce which is the best option. A more accurate representation of relative expected-values in the dACC should therefore bias behavior towards exploitation; and correspondingly, a less accurate representation as seen in anxious individuals, should bias behavior towards exploration. Notably, it was recently argued that non-greedy decisions could result from noisy learning processes, where the amount of learning noise was tracked by the dACC well into the decision process (Findling et al. 2019). However, if decisions made by more anxious individuals in the present study was affected by such learning noise, we would expect to see an overall increase in dACC activation as a function of trait anxiety during decision making, but this was not the case. Moreover, we included an explicit factor in our model, outcome-uncertainty, which was uncorrelated with expected-value, and which significantly contributed to explain the proportion of non-greedy, exploratory decisions elicited by more anxious individuals.

Furthermore, anxious individuals showed a more positive coupling between outcome-uncertainty and BOLD signal in the bilateral aINS, as well as increased uncertainty reduction. These results parallels findings showing that the aINS responds to different types of uncertainty (Simmons et al. 2008), is involved in the aversion and the intolerance of uncertainty (Grupe and Nitschke 2013), and is hyper-responsive to uncertainty and distal threats in high anxiety (Etkin and Wager 2007; Grupe and Nitschke 2013; Fung et al. 2019b). Accordingly, a more rapid accumulation of uncertainty-related activation in the aINS facilitates decisions to explore, as displayed by more anxious individuals. Because the aINS also represents negative affective states and stimuli, including pain, disgust, anxiety, conditioned fear, negative pictures, and odors (Gogolla 2017; Alvarez et al. 2015; Ploghaus et al. 1999; Wright et al. 2004), a putative mechanism by which the aINS can increase exploration might be that uncertainty-related activation of the aINS is perceived as aversive, leading to decisions aimed at reducing uncertainty, namely exploration, and hence avoiding the negative state. This hypothesis needs to be tested directly.

These ideas are similar to recent findings in studies of over-generalization (Dunsmoor and Paz 2015; Lissek 2012). There as well, aversive contexts led to wider stimulus generalization (Laufer and Paz 2012; Resnik, Sobel, and Paz 2011), similar to our finding here of an overall higher exploration in aversive contexts; yet in anxious individual, these studies found wider generalization for both aversive and appetitive contexts (Laufer, Israeli, and Paz 2016), similar to what we find here that exploration is increased with anxiety for both types of valence. Because wider/over-generalization can also stem from uncertainty and less information about the stimulus, the two mechanisms might be linked. One possibility is an attentional-shift between top-down processes devoted to value-based decision making - evidenced by weaker dACC activity, to bottom-up processing of aversive information of increased uncertainty - evidenced by a responsive aINS (Eysenck et al. 2007; Bar-Haim et al. 2007). Supporting this idea, we found that the increased trade-off between expected-value and uncertainty-reduction, as seen in individuals with high trait anxiety, was mediated by the differential neuronal representation of expected-value in the dACC and uncertainty-reduction in the aINS.

Exploratory decision making is an integral part of everyday life, being involved in decisions ranging from small (e.g. where to eat lunch) to large (e.g. which partner to marry). Therefore, our results highlight two mechanisms which may be critical determinants of life-quality in anxious individuals, even if in the normal-range. A high exploratory drive could reduce the ability to engage in long-term commitments, and a reduced focus on expected values may lead to faster abandoning of good positions. In turn, maladaptive exploratory strategies could trigger a cascade of events that reinforces symptoms of anxiety and depression. For example, exaggerated exploration may lead to increased exposure to suboptimal outcomes which could trigger feelings of regret (Coricelli and Rustichini 2010) and reduce confidence in one's ability to make appropriate decisions (Bandura 1982). Indeed, anxiety and depression have been associated with both increased feelings of regret (Roese et al. 2009; Ehring and Watkins 2008) and reduced self-efficacy (Muris 2002; Zumberg, Chang, and Sanna 2008; Luszczynska, Gutierrez-Dona, and Schwarzer 2005). Moreover, high levels of trait anxiety is a risk-factor for chronic anxiety and depression (Chambers, Power, and Durham 2004; Weger and Sandi 2018), and their development and maintenance may depend on dysfunctional learning and decision making (Mineka and Oehlberg 2008; Paulus and Yu 2012; Lee 2013; Rosen and Schulkin 1998; Grupe 2017).

We estimated anxiety using the standard Spielberger's Trait-Anxiety Inventory (Spielberger et al. 1983), mainly because it provides a gradual scale also for the normal (below clinical) range. This provides two benefits: first, observing how exploration varies across the distribution, rather than comparing two populations which are somewhat arbitrarily divided (patients vs. controls); and second, examining the normal range, allowing us to examine how maladaptive decisions are mediated by exploration even in healthy individuals. Because such maladaptive decision have a huge impact on daily-life in all individuals, and definitely on societies and industry,

we argue that more studies should use such gradual scales over non-clinical populations (Browning et al. 2015; Fung et al. 2019a; Gagne et al. 2020). It should be noted that Spielberger's Trait-Anxiety Inventory has been debated for its lack of convergent and discriminant validity, suggesting that it estimates 'negative affectivity' rather than proneness to anxiety per-se (Balsamo et al. 2013). Yet, because negative affectivity is closely linked to psychopathology (Kotov et al. 2010; Stanton and Watson 2014), and has been noted as a vulnerability factor for developing anxiety and depression (Clark, Watson, and Mineka 1994), our results still bear significant relevance.

In summary, we identify two neuro-behavioral mechanisms whose tradeoff determines exploratory bias and is directly related to anxiety levels. Importantly, we show that anxiety within the 'normal' spectrum is a critical determinant of exploratory behaviors in humans, might increase maladaptive decision-making, and therefore affect life-quality in individuals dealing with everyday decisions. We suggest that in some individuals, these factors can develop further into a vicious cycle. Put simply, poor decision making caused by an excessive drive to explore may increase the frequency of poor real-life outcomes, which in turn may contribute to psychopathology.

Acknowledgments:

We thank Dr. Edna Furman-Haran and Fanny Attar for MRI procedures. K.C. Aberg was supported by the Swiss Society of Friends of the Weizmann Institute Postdoctoral Fellowship grant. The work was supported by a Joy-Ventures grant, an ISF #2352/19 and an ERC-2016-CoG #724910 grants to R. Paz.

Financial Disclosures section:

The authors declare no competing financial interest.

Methods

Participants

After having provided written informed consent, fifty-one (51) participants (n=20/31 for the behavioural pilot/main-fMRI study) joined the experiment. The study was approved by the Helsinki committee of the Sourasky Medical Center, under protocol number 0287–09-TLV (Ministry of health protocol no. HT5271) and further approved by the Institutional Review Board (IRB) of the Weizmann Institute.

All participants were right-handed and reported no previous history of psychiatric or neurological disorders. Handedness was assessed by asking participants to indicate their preferred hand for writing and drawing, on a scale ranging from 1 (always using the left hand) to 7 (always using the right hand). Only participants that indicated always using the right hand on both questions were recruited. Two participants who fell asleep in the scanner, and one participant who displayed excessive head movements during scanning were excluded. Therefore, data from 28 participants (eight males; average age ± SEM: 27.571 ± 0.730) were included in the subsequent data analyses. These participants showed a normative distribution of trait-anxiety scores (mean=36.321, SD=9.982; see (Spielberger et al. 1983), Table 1).

Task

In each trial, participants decided which of three slot machines to play (Fig.1A,B). The outcome of the selected machine was then revealed. Exploration was promoted by varying the outcome assigned to each slot machine across trials (Fig.1C). Thus, to optimize overall performance participants needed to track the expected outcome associated with each slot machine, and to make exploratory decisions to detect outcome changes.

Each trial started with a fixation cross displayed for 500 ms, followed by the display of the three slot machines (Fig.1A,B). Participants needed to elicit a response within 1500 ms, after which the selected machine became animated. The animation lasted for 1500 ms minus the response time plus a jittered time duration drawn from an exponential distribution with mean 3500 ms. The outcome associated with the selected machine was then presented for 1000 ms (Fig.1A,B), after which the next trial was initiated following a jittered delay drawn from an exponential distribution with mean 3500 ms. If no response had been registered within the 1500 ms, the words 'Too slow' were displayed for 1000 ms.

The outcome distribution of each slot machine was described by different sine-wave functions (some examples are shown in Fig.1C-G). The magnitude and mean of each sine-wave was always 20 and 50 points respectively. The period lengths of the three sine-waves in a session were 100, 67, or 33 trials, respectively, and the sine-waves were phase-shifted by 0.10, 0.45, and 0.90 period lengths. A pseudorandomization scheme was administered to ensure that no participant was exposed to the same stimulus configuration by assigning different phase-shifts to each sine-wave, different sine-waves based on Gain/Loss conditions, spatial position, and color. Moreover, the assignment of color to a machine's spatial position and Gain/Loss condition also varied pseudorandomly.

Participants engage more in risky decisions when facing certain losses, as compared to certain gains (De Martino et al. 2006), and this effect is exaggerated in more anxious individuals (Jepma and Lopez-Sola 2014; Xu et al. 2013). Here, to test for differences in exploratory decision making between gain and loss conditions, each participant performed two versions of the task. In the Gain condition, all outcomes were positive and participants acquired points (Fig.1D,F). By contrast, participants had to avoid losing points in a Loss condition

with only negative outcomes (Fig.1E,G). The points corresponded to a monetary bonus at the end of the experiment. In the Gain condition, the bonus (initially set to 0 Israeli new shekels) increased as more points were collected. In the Loss condition, the bonus (initially set to 10 shekels) decreased with the accumulation of negative points. At the end of the experiment, each participant's bonus was calculated as 10 * proportion of maximal and minimal points for the Gain and Loss condition, respectively. The order of Gain and Loss conditions was counterbalanced across participants.

Modelling approach

As in previous studies, a computational approach was used to determine whether a decision could be classified as exploratory or as an exploitation (Daw et al. 2006). Specifically, an exploratory decision was defined as the selection of a suboptimal machine which was different from the machine selected in the previous trial (Tzovara et al. 2012), where suboptimal referred to any machine with a less-than-highest expected value.

For all models, the expected value $Q_i(t)$ of the selected machine *i* in trial *t* was updated by the mismatch between the expected value and the actual outcome *R(t)*, i.e. the reward prediction error $\delta_Q(t)$, scaled by the learning rate α_Q.

$$Q_i(t+1) = Q_i(t) + \alpha_Q * \delta_Q(t) \tag{1}$$

$$\delta_Q(t) = R(t) - Q_i(t) \tag{2}$$

The probability *p* of selecting machine *i* in trial *t* was modeled using two approaches. For all models except one (see the ε-greedy model below), decisions were modeled using a soft-max choice probability function in which the probability of selecting a particular machine depends on its utility, u_i:

$$p_i(t) = \frac{e^{u_i(t)}}{\sum_j e^{u_j(t)}} \tag{3}$$

Here, u_i is defined as the sum of parameters that may contribute to a decision (e.g. expected value) multiplied by their respective decision weights (e.g. how much expected value contributes to a decision). For example, one common way to model exploratory tendencies is by the inverse temperature β_Q. Specifically, the decision weight β_Q determines the trade-off between exploitation and exploration such that small and large values of β_Q are associated with increased and decreased exploration, respectively, because β_Q determines the influence of expected value on a decision:

$$p_i(t) = \frac{e^{Q_i(t)*\beta_Q}}{\sum_j e^{Q_j(t)*\beta_Q}} \tag{4}$$

However, other factors unrelated to the expected values of stimuli may also contribute to decision making. For example, exploration may be directed towards options with large outcome-variability (Badre et al. 2012; Frank et al. 2009; Gershman 2018). To test this notion, the outcome-variability $U_i(t)$ was updated by the mismatch between the expected outcome-variability and the absolute reward prediction error $\delta_Q(t)$, i.e. the outcome-variability prediction error $\delta_U(t)$, scaled by the learning rate α_U:

$$U_i(t+1) = U_i(t) + \alpha_U * \delta_U(t) \tag{5}$$

$$\delta_U(t) = |\delta_Q(t)| = |R(t) - Q_i(t)| \tag{6}$$

Additionally, in a dynamic environment, the uncertainty associated with an option increases with the time since it was last visited. For example, an outcome is more likely to have changed for a slot machine that has

not been selected for many trials. To test the impact of this factor on decision making, we defined outcome-uncertainty $T_i(t)$ as the number of trials since a machine was last selected.

$$T_i(t+1) = \begin{cases} 0, & if\ selected = i \\ T_i(t) + 1, & otherwise \end{cases} \tag{7}$$

We also modeled the impact of an individual's tendency to randomly switch between different machines, a decision making strategy which could help to rapidly discover changes in the environment. Accordingly, $S_i(t)$ is set to 0 and 1 for slot machines that were selected and rejected, respectively. We refer to this factor here as 'random switching', whereas previous research denotes the reverse of this factor as 'choice stickiness' or 'perseverance' (Chakroun et al. 2020; Gershman, Pesaran, and Daw 2009).

$$S_i(t+1) = \begin{cases} 0, & if\ selected = i \\ 1, & otherwise \end{cases} \tag{8}$$

Given reports of more exploration with increased uncertainty in the environment (Gershman 2018; Wilson et al. 2014), we also included a factor which scales the impact of expected value with the total uncertainty of the environment, as estimated by the total outcome-variability (Gershman 2018; Gershman and Tzovaras 2018; Wilson et al. 2014). Specifically, $tU_i(t)$ is defined as:

$$tU(t) = \sum_i U_i(t) \tag{9}$$

These parameters were included in a soft-max choice probability function. Thus, in the model including all these parameters, the probability of selecting machine *i* depends on β_Q, β_U, β_T, β_S, β_{tU}, and β_{U*T}, which determine the influence of expected-value $Q_i(t)$, outcome-variability $U_i(t)$, outcome-uncertainty $T_i(t)$, random-switching $S_i(t)$, expected-value scaled by total outcome-variability $Q_i(t)/tU(t)$, and outcome-uncertainty scaled by outcome-variability $T_i(t) * U_i(t)$, respectively.

$$p_i(t) = \frac{e^{Q_i(t)*\beta_Q + U_i(t)*\beta_U + T_i(t)*\beta_T + S_i(t)*\beta_S + Q_i(t)/tU_i(t)*\beta_{tU} + U_i(t)*T_i(t)*\beta_{U*T}}}{\sum_j e^{Q_j(t)*\beta_Q + U_j(t)*\beta_U + T_j(t)*\beta_T + S_j(t)*\beta_S + Q_i(t)/tU_i(t)*\beta_{tU} + U_j(t)*T_j(t)*\beta_{U*T}}} \tag{10}$$

We also implemented an *ε*-greedy algorithm, in which the probability of selecting the machine with the highest expected value in a trial is 1-*ε*. Accordingly, small and large values of *ε* are associated with decreased and increased exploration, respectively:

$$p_i(t) = \begin{cases} 1 - \varepsilon, & if\ Q_i(t) = Q_{max}(t) \\ \varepsilon, & otherwise \end{cases} \tag{11}$$

Finally, we included two Kalman-filter models where the learning rate varies between trials as a function of a machine's current level of uncertainty. Models based on Kalman-filters have been used previously to disentangle random and directed exploration in human decision making (Bishop 2006; Gershman 2018; Gershman and Tzovaras 2018). In these models, the prediction error and expected-value update is determined by Equations 1-2, but the learning rate varies from trial-to-trial according to:

$$\alpha(t) = \frac{\sigma(t)^2}{\sigma(t)^2 + \sigma_0^2} \tag{12}$$

$\sigma(t)^2$ and σ_0^2 respectively denotes the selected machine's variance in trial t and prior variance. The variance of the selected machine is updated according to:

$$\sigma(t+1)^2 = \sigma(t)^2 - \alpha(t) * \sigma(t)^2 \tag{13}$$

In the first Kalman-filter type model, which is based on previous research (Gershman 2018; Gershman and Tzovaras 2018), decision weights were included for expected-value (β_Q), outcome-variability (β_U; here, the outcome variability is estimated in each trial by σ), and expected-value scaled by total outcome-variability (β_{tU}; here, tU is estimated in each trial by $\sqrt{\sum_i \sigma_i(t)^2}$):

$$p_i(t) = \frac{e^{Q_i(t)*\beta_Q+\sigma_i(t)*\beta_U+\frac{Q_i(t)}{tU(t)}*\beta_{uT}}}{\sum_j e^{Q_j(t)*\beta_Q+\sigma_j(t)*\beta_U+\frac{Q_j(t)}{tU(t)}*\beta_{uT}}} \tag{14}$$

In a second Kalman-filter model, decision weights were included for expected-value (β_Q), outcome-uncertainty (β_T; defined as in Equation 7), and random-switching (β_S; defined as in Equation 8).

$$p_i(t) = \frac{e^{Q_i(t)*\beta_Q+T_i(t)*\beta_T+S_i(t)*\beta_S}}{\sum_j e^{Q_j(t)*\beta_Q+T_j(t)*\beta_T+S_j(t)*\beta_S}} \tag{15}$$

All free parameters were fitted individually to each participant's behavior by minimizing the negative log-likelihood estimate:

$$-LLE = -\ln\left(\prod_1^n p_i(t)\right) \tag{16}$$

To avoid getting stuck in local minima, each fit was repeated 100 times with different random starting points for each free parameter. The two learning rates were limited by 0 and 1, while no bound was applied for the β decision weights (yet, the random starting points were limited by -5 and +5). For the Kalman-filter models, the prior variance σ_0^2 was set to 4.0 for reasons of model-degeneracy (Daw et al. 2006).

To obtain the most parsimonious model, i.e. the model that provides the best trade-off between its complexity and goodness of fit, different models incorporating different combinations of the above mentioned parameters were created and confronted. In total, 12 different models were tested: models 'Q', 'QU', 'QT', 'QUT', 'QS', 'QTS', 'QUTS', 'Q T*U S', and 'QTS Q/tU', all use constant learning rates and differed only in the factors included in the softmax choice-probability function, as indicated by their denotations. 'κ-QTS' and 'κ-QU Q/tU' are Kalman-filter type models using softmax choice probability function. In other words, the only difference between the QTS and the κ-QTS models is that they use constant and variable learning rates, respectively. Finally, the *ε*-greedy model uses a constant learning rate, but an *ε*-greedy choice probability function (Equation 11).

A version of Akaike's information criterion (AICc) that penalizes a model's goodness of fit based on the number of trials *n* and the number of fitted parameters *k* was used to compare the goodness of fit between models.

$$AICc = 2 * k - 2 * \ln(LLE) + \frac{2*k^2+2*k}{n-k-1} \tag{17}$$

To determine the most parsimonious model based on using the AICc as model evidence, we applied a Bayesian Model Selection (BMS) procedure (Stephan et al., 2009). The BMS assumes that the observed behavior of each participant may have been generated by a different model drawn from an unknown population distribution. The most parsimonious model can then be determined via the protected exceedance probability, which is defined as the posterior probability that the model has a higher frequency than the other tested models, while also accounting for the probability that differences were due to chance. For consistency, we performed the same BMS analysis using the Bayesian Information Criterion (BIC) as model evidence, because the BIC provides a larger penalty for more complex models, i.e. $\mathrm{k} * \ln(n)$ for BIC rather than $2 * k + \frac{2*k^2+2*k}{n-k-1}$ for AICc (Bishop 2006; Gershman 2018).

Parameter recovery

It needs to be shown that fitted model-parameters are meaningful via their extraction from a data set with known parameters (Wilson and Collins 2019). To accomplish this, the behaviors of 500 virtual participants were simulated using the QTS model by randomly sampling the model parameters from uniform distributions with boundaries [0 1], [0 1], [0 1], and [-5 5] for α_Q, β_Q, β_T, and β_S.

Statistical analyses

Behavior and computational results were analyzed using traditional statistical approaches, including repeated measures ANOVAs, *t*-tests, Pearson correlations, and nonlinear regression. A mediation analysis was conducted to test whether the nonlinear relationship between Trait anxiety (T) and Overall performance (O) was mediated by Exploration (E). In essence, this approach tests whether a nonlinear relationship between an independent variable (here, T) and a dependent variable (here, O) is significantly reduced by including a mediating variable (here, E) in the regression analysis (Hayes and Rockwood 2017). Critically, according to contemporary statistical thinking, it is sufficient to show that the predictive value of an independent variable is significantly reduced when controlling for a mediating variable (Hayes and Rockwood 2017). A Monte-Carlo randomization procedure was performed to assess statistical significance: First, the quadratic regression coefficients of T predicting O alone ($b^2_{Anxiety1}$), and when both T and E predicts O ($b^2_{Anxiety2}$ and $b^2_{Explore}$) were calculated. The change in T's predictive value was calculated as $\Delta b^2_{Anxiety}=b^2_{Anxiety2}-b^2_{Anxiety1}$. To assess the statistical significance of $\Delta b^2_{Anxiety}$, the above analysis was repeated 10'000 times by each time randomly shuffling the values of E. This procedure creates a distribution ($\Delta b^2_{Anxiety,\ null}$) for the null-hypothesis that Exploration does not change the relationship between Trait anxiety and Overall performance. Statistical significance of the observed $\Delta b^2_{Anxiety}$, and therefore also whether Exploration mediated the relationship between Trait anxiety and Overall performance, was assessed by comparing $\Delta b^2_{Anxiety}$ to the resulting null-distribution $\Delta b^2_{Anxiety,\ null}$.

MRI Data

Image Acquisition

MRI images were acquired using a 3T whole body MRI scanner (Trio TIM, Siemens, Germany) with a 12-channel head coil. Standard structural images were acquired with a T1 weighted 3D sequence (MPRAGE, Repetition time (TR)/Inversion delay time (TI)/Echo time (TE)=2300/900/2.98 ms, flip angle=9 degrees, voxel dimensions=1 mm isotropic, 32 slices). Functional images were acquired with a susceptibility weighted EPI sequence (TR/TE=2000/30 ms, flip angle=75 degrees, voxel dimensions=3x3x3.5 mm, 192 slices).

Data Analysis

Functional MRI data were preprocessed and then analyzed using the general linear model (GLM) for event-related designs in SPM8 (Welcome Department of Imaging Neuroscience, London, UK; http://www.fil.ion.ucl.ac.uk/spm). During preprocessing, all functional volumes were realigned to the mean image, co-registered to the structural T1 image, corrected for slice timing, normalized to the MNI EPI-template, and smoothed using an 8 mm FWHM Gaussian kernel. Statistical analyses were performed on a voxel wise basis across the whole-brain. At the first-level analysis, individual events were modelled by a standard synthetic hemodynamic response function (HRF) and six rigid-body realignment parameters were included as nuisance covariates when estimating statistical maps. Loss and Gain conditions were modeled as separate sessions within the same GLM. Two types of analysis were conducted to investigate the neural correlates of exploratory decision making.

fMRI analysis 1: Exploration versus exploitation

The purpose of this model was to define regions-of-interest (ROIs) to be used in subsequent analyses. To elucidate brain regions engaged by decisions to explore vs. exploit, an event-related design was created with four event-types (exploratory decision, exploitation, response, and feedback) for each Gain/Loss condition separately. The first two event-types were time-locked to the presentation of the slot machines (decision onset), the response was locked to the button press, and the feedback was time-locked to the presentation of the outcome (feedback onset). Trials which could not be categorized as exploration or exploitation were included in a regressor-of-no-interest together with trials in which no response was provided.

fMRI analysis 2: Parametric modulation of decisions by differences in expected value and outcome uncertainty, and by prediction error at feedback

To determine brain regions tracking trial-by-trial differences in expected value (ΔQ) and outcome uncertainty (ΔT), as well as prediction errors (δ), an event-related design was created with three event-types (decision onset, response onset, feedback onset) time-locked to the presentation of the slot machines, the button press, and the presentation of outcomes, respectively. In each trial, ΔQ was defined as the difference in expected value between the average expected value of the two rejected machines and the selected machine. Similarly, ΔT was defined as the difference in uncertainty between the average uncertainty for the two rejected machines and the selected machine. These two parameters were added as parametric modulators at decision onset, while δ was added as a parametric modulator of the feedback onset. Trials in which no response was provided were included in a regressor-of-no-interest.

Regions of interest (ROIs)

Exploratory decision making

The dorsal anterior cingulate cortex (dACC), the anterior insulae (aINS), and the frontopolar cortex (FPC) are engaged by different aspects of uncertainty and by decisions to explore (Badre et al. 2012; Blanchard and Gershman 2018; Daw et al. 2006; Laureiro-Martinez et al. 2015; Grupe and Nitschke 2013), while exploitation activates the ventromedial prefrontal cortex (vmPFC; (Laureiro-Martinez et al. 2015)). Our fMRI analyses therefore focus on these brain regions. A dACC ROI was created using the WFU Pickatlas toolbox (Maldjian et al. 2003) in combination with a previously described procedure (Cascio, Konrath, and Falk 2015; Wang et al. 2016): First, the union of Broadmann areas 24 and 32, and AAL masks of the cingulate cortex (anterior, middle, posterior; dilated to 2mm) was created. Next, Broadmann areas 8 and 9 were subtracted from this union, and the resulting ROI was further restricted by a cuboid with bounds x = [-16 16], y = [0 33], z = [6 52]. The bilateral aINS was defined by the left and right insula ROIs of the WFU pickatlas toolbox. The FPC ROI was defined as the bilateral Area Fp1 of the Anatomy toolbox (Eickhoff et al. 2005; Bludau et al. 2014). The vmPFC ROI was defined as a 10 mm radius sphere centered on x=-3, y=42, z=-15, calculated from a recent meta-analysis investigating the role of the vmPFC/orbitofrontal cortex in representing a neural common currency for choice (Levy and Glimcher 2012). These ROIs were combined into one ROI mask.

Prediction error

The Nucleus Accumbens (NAcc) encodes prediction errors (Aberg, Doell, and Schwartz 2016a; Cohen 2007). The NAcc ROI was obtained from a recent study trying to delineate the shell and core subregions of the NAcc (Baliki et al. 2013).

Statistical analysis

To investigate how the neural correlates of exploratory decision making are modulated by trait anxiety, the fMRI data was analyzed in two steps. This analysis follows recent recommendations for correlating inter-individual factors, such as personality traits, with fMRI data (Vul et al. 2009).

First, fMRI analysis 1 was used in combination with the dACC/aINS/FPC/vmPFC ROI mask to identify group-level peak-voxel activity for the contrast between exploratory decisions and exploitation (and vice versa). These activations were identified using a threshold of p=0.001 (uncorrected) and a minimum cluster size of five contiguous voxels. Small volume correction (SVC) using a threshold of p<0.05 Family-Wise Error Rate (FWER) for multiple comparisons were applied using the ROI mask. Importantly, because this analysis was performed across Gain/Loss condition and individuals, it allows an unbiased selection of brain coordinates for studying the differential contribution of Gain/Loss condition and modulations by trait anxiety.

Second, fMRI analysis 2 was used to extract beta parameter estimates (i.e. $\beta_{\Delta Q}$ and $\beta_{\Delta T}$) for the different parametric modulators (i.e. ΔQ and ΔT, respectively). Activity was extracted by averaging activity within 3 mm radius spheres centered on the coordinates of the peak voxels identified via fMRI analysis 1. The resulting data was then entered into repeated measures ANOVAs with one within-subject factor Condition (Gain, Loss) and continuous covariate Trait anxiety. Statistical thresholds were Bonferroni-corrected based on the number of ANOVAs conducted for each condition. To ensure that any significant result was not critically dependent on the 3mm radius of the sphere (i.e. by averaging the voxel intensity across five voxels), we repeated the same analyses using spheres with radiuses of 0mm (1 voxel), 5mm (19 voxels), and 10mm (137 voxels) respectively. The resulting p-values of these complementary analyses are presented together with the results of the main analyses that used 3mm radius spheres.

Figures and legends

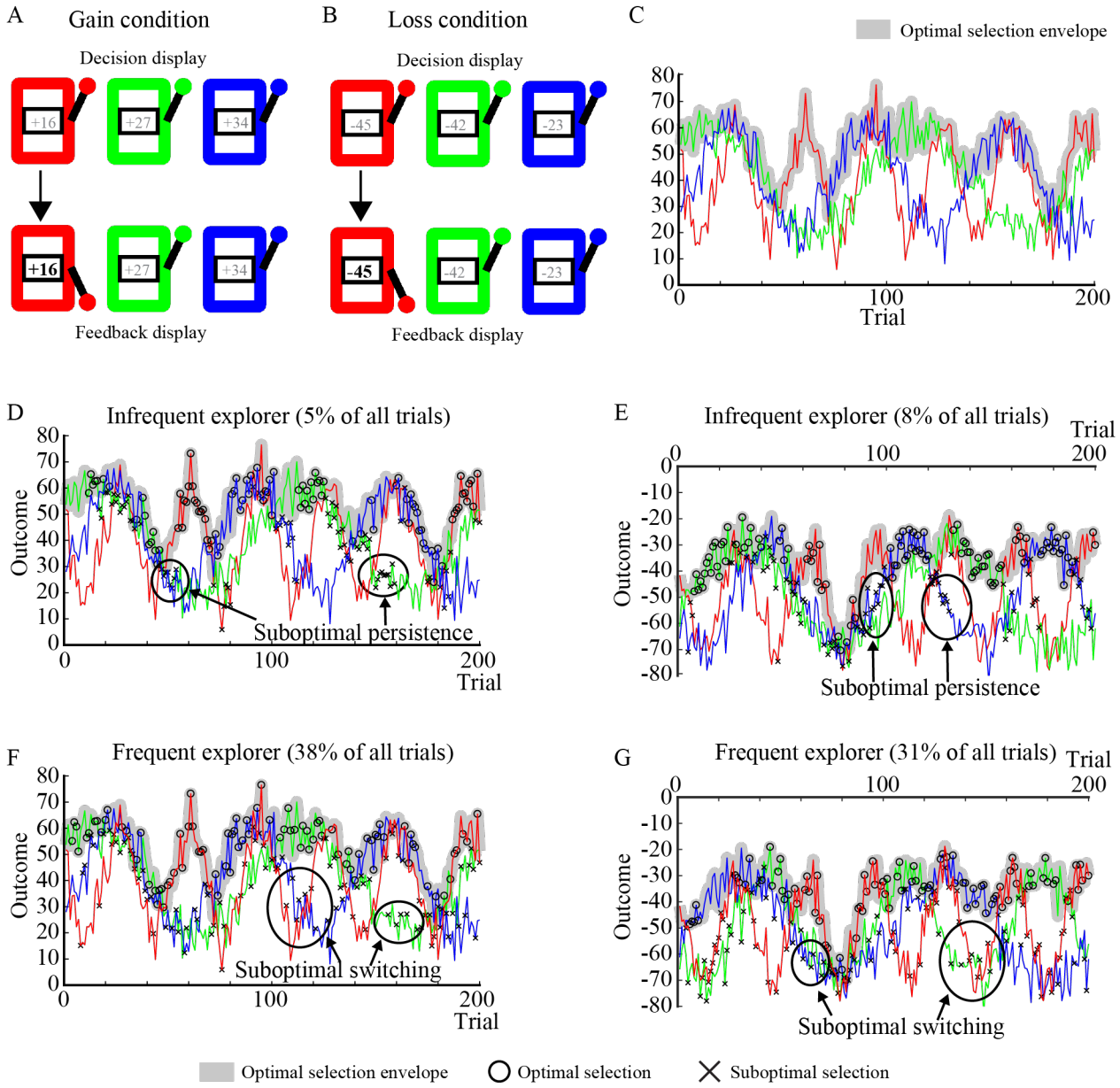

Figure 1. Paradigm and individual behavior. **A.** Illustration of Decision and Feedback displays in the Gain condition, where all outcomes are positive. During the decision display, the outcomes of each machine is hidden (grey characters), and participants decide to play one of the three slot machines by pressing the corresponding button. After selecting a slot machine its outcome is revealed (+16 points), while the outcomes of the non-selected machines remain hidden. **B.** In the Loss condition, all outcomes are negative and participants aim to avoid large losses. **C.** Exploration is promoted by varying the outcome associated with each machine across trials. The optimal potential outcome is displayed in grey. **D,E.** Examples of infrequent explorers in the Gain (**D**) and in the Loss (**E**) condition. **F,G.** Examples of frequent explorers in the Gain (**F**) and in the Loss (**G**) condition.

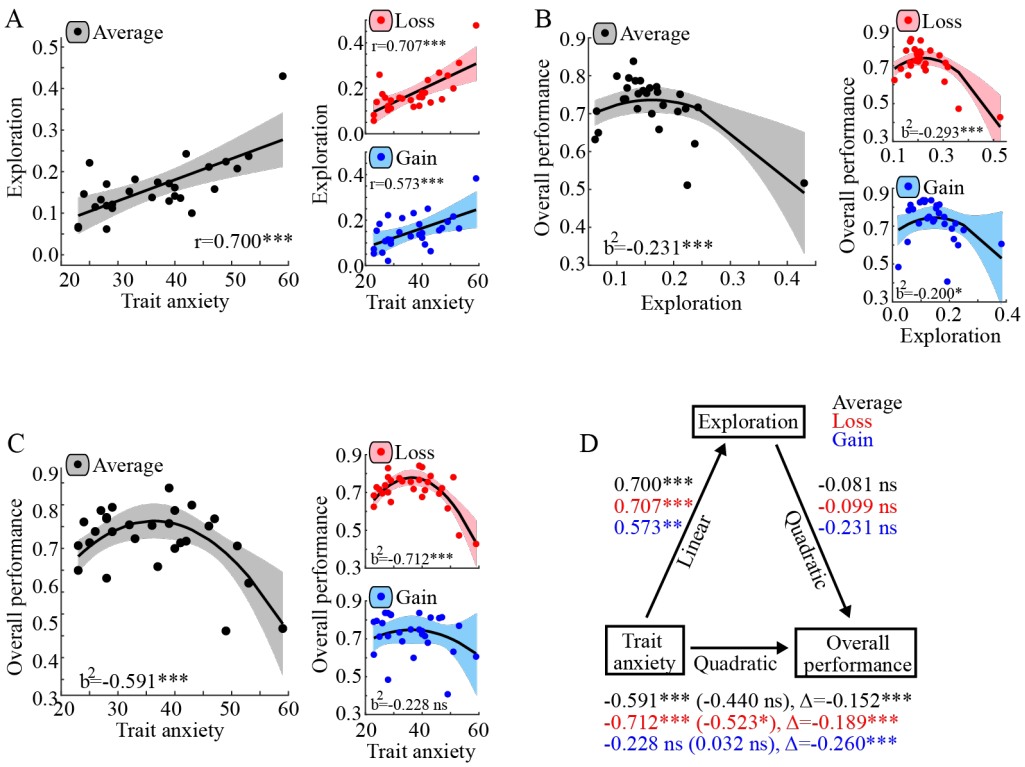

Figure 2. Behavioral results. **A.** Exploration, the proportion of exploratory decisions, increased linearly with increased levels of trait anxiety across conditions (left panel), and separately in the Loss (top right panel) and in the Gain (bottom right panel) condition. **B.** Overall performance, defined as the correlation between maximal and actual outcomes across all trials, varied non-linearly with exploration across conditions (left panel), and separately in the Loss (top right panel) and in the Gain condition (bottom right panel). **C.** Overall performance varied significantly and non-linearly with trait anxiety across conditions (left panel), and in the Loss (top right panel), but not significantly so in the Gain condition (bottom right panel). **D.** Standardized regression coefficients for the relationships between trait anxiety and overall performance as mediated by exploration. The standardized regression coefficient between trait anxiety and overall performance, when controlling for exploration, is in parenthesis. The relationship between trait anxiety and overall performance was mediated by exploration across conditions (black text), as well as in both the Loss (red text) and the Gain condition (blue text), as indicated by significantly reduced quadratic regression coefficients when controlling for exploration (both p-values <0.001). ***p<0.001; *p<0.05; ns p>0.05.

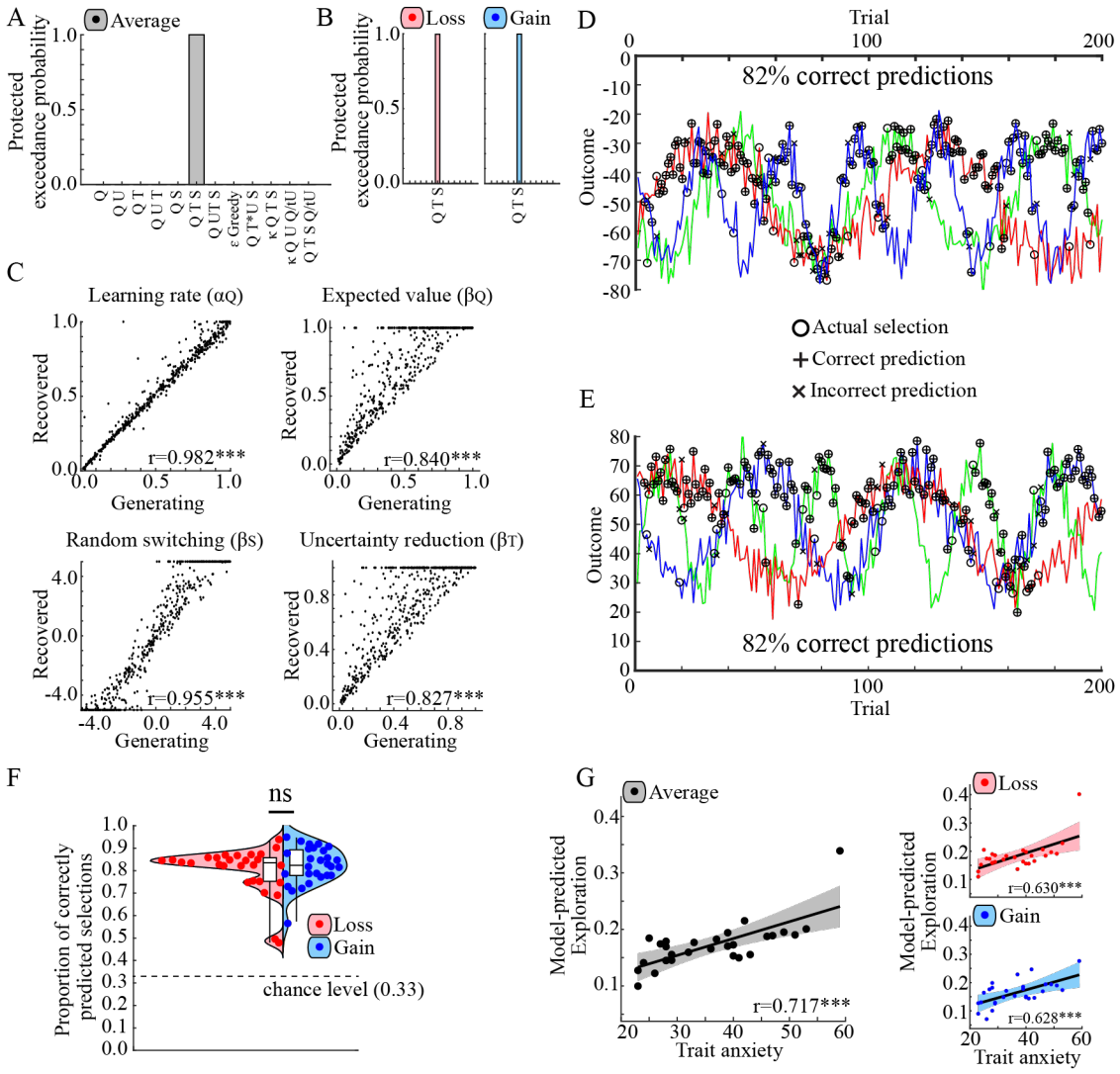

Figure 3. Modeling behavior. **A.** Protected exceedance probabilities obtained from the Bayesian Model Selection (BMS) procedure indicates the QTS model as the winning model when model evidence (here, AICc) was collapsed across conditions (top panel), and **B**. separately in the Loss (left panel) and in the Gain (right panel condition). **C.** The parameter-recovery procedure for the QTS model indicates significant correlations between the parameters used to generate behavior and the recovered parameters obtained from re-fitting the model to the generated behavior. **D,E.** Representative examples displaying the similarity between actual and model-predicted selections in the Loss (**D**) and in the Gain condition (**E**). **F.** The average proportion of correctly predicted selections by the QTS model was well above chance for all participants. **G.** The average model-predicted probability of exploring correlated positively with trait anxiety across conditions (left panel), and in the Loss (top right panel) and in the Gain condition (bottom right panel). ***p<0.001; *p<0.05; ns p>0.05.

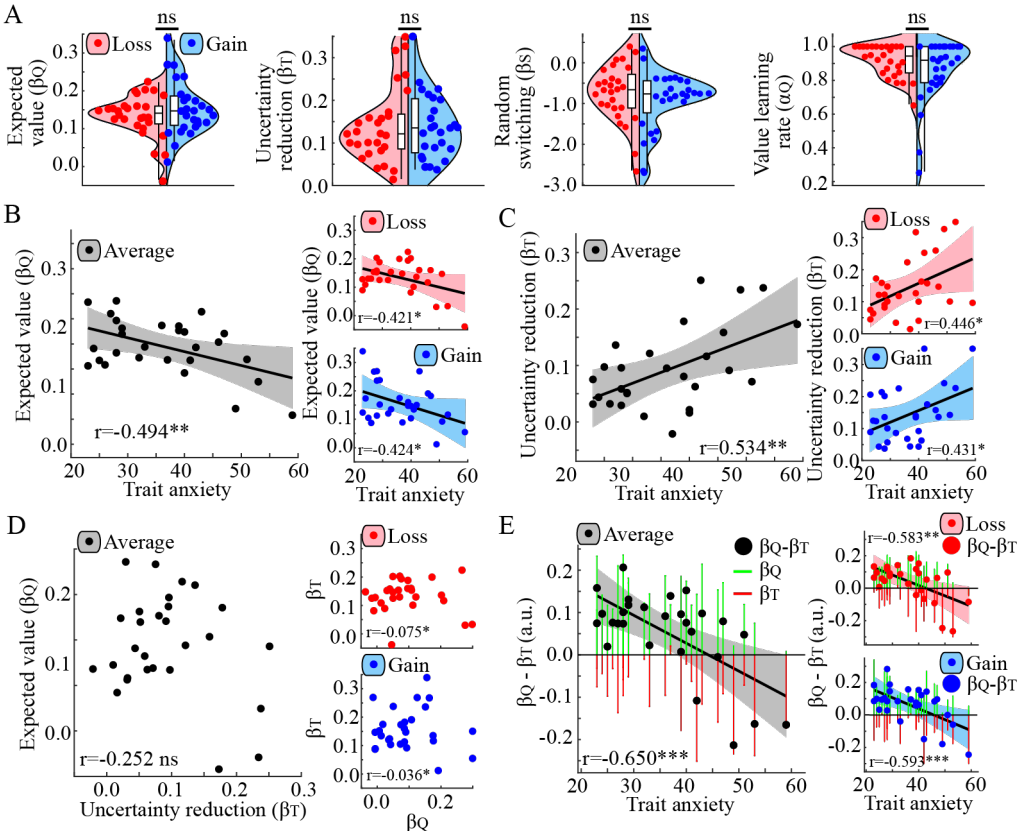

Figure 4. Individually fitted model parameters. **A.** Each of the four fitted parameters of the QTS model are relevant predictors of behavior across participants. **B.** Decision weights estimating the influence of expected value (β_Q) were negatively correlated with trait anxiety across conditions (left panel), as well as separately in the Loss (top right panel) and in the Gain condition (bottom right panel). **C.** Decision weights estimating reduction of outcome uncertainty (β_T) was positively correlated with trait anxiety across conditions (left panel), as well as separately in the Loss (top right panel) and the Gain condition (bottom right panel). **D.** β_Q did not correlate with β_T across conditions (left panel), nor in the Loss (top right panel) or in the Gain condition (bottom right panel). **E.** The difference between β_Q and β_T correlated negatively with trait anxiety across conditions (left panel), as well as separately in the Loss (top right panel) and in the Gain condition (bottom right panel). ***p<0.001; **p<0.01; *p<0.05; ns p>0.05.

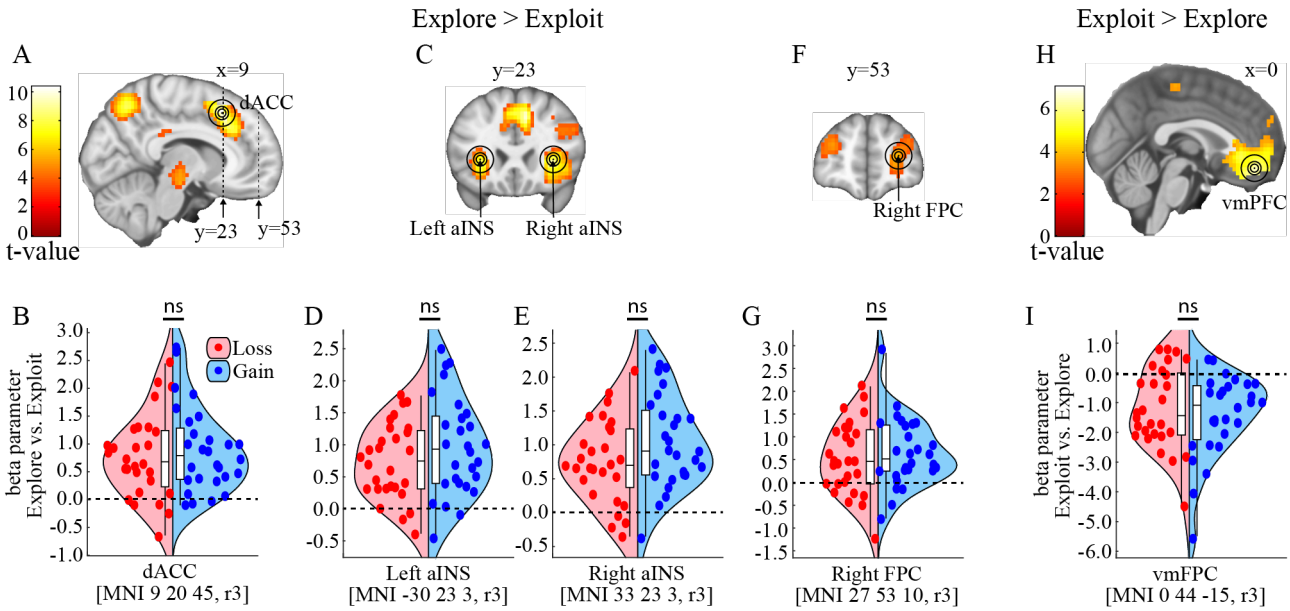

Figure 5. BOLD signal as a function of exploratory decision making. **A-G.** Decisions to explore (vs. exploit) increased activity in (**A,B)** the dorsal anterior cingulate cortex (dACC), (**C-E**) bilaterally in the anterior insula (aINS), and (**F,G**) in the right frontopolar cortex (FPC). Extracted beta parameters for decisions to explore (vs. exploit), extracted from spheres centered on the peak activation within a priori ROIs in the dACC (**B**), L aINS (**D**), R aINS (**E**), and the R FPC (**G**). **H.** Decisions to exploit (vs. explore) increased activity in the ventromedial prefrontal cortex (vmPFC). **I.** Extracted beta parameters for decisions to exploit (vs. explore), extracted from 3 mm radius spheres centered on the peak activation within the vmPFC ROI. Peak voxel coordinates presented here were used in subsequent analyses by extracting the average voxel intensity within 3 mm spheres centered on these coordinates. To ensure the robustness of any observed effects, all analyses were repeated using spheres with different radii (0, 3, 5, and 10mm), as illustrated in **A**, **C**, **F**, and **H**. For visualization purposes, the brain activation in **A**, **C**, **F**, and **H**, is displayed using an uncorrected threshold of p=0.001. ns p>0.05.

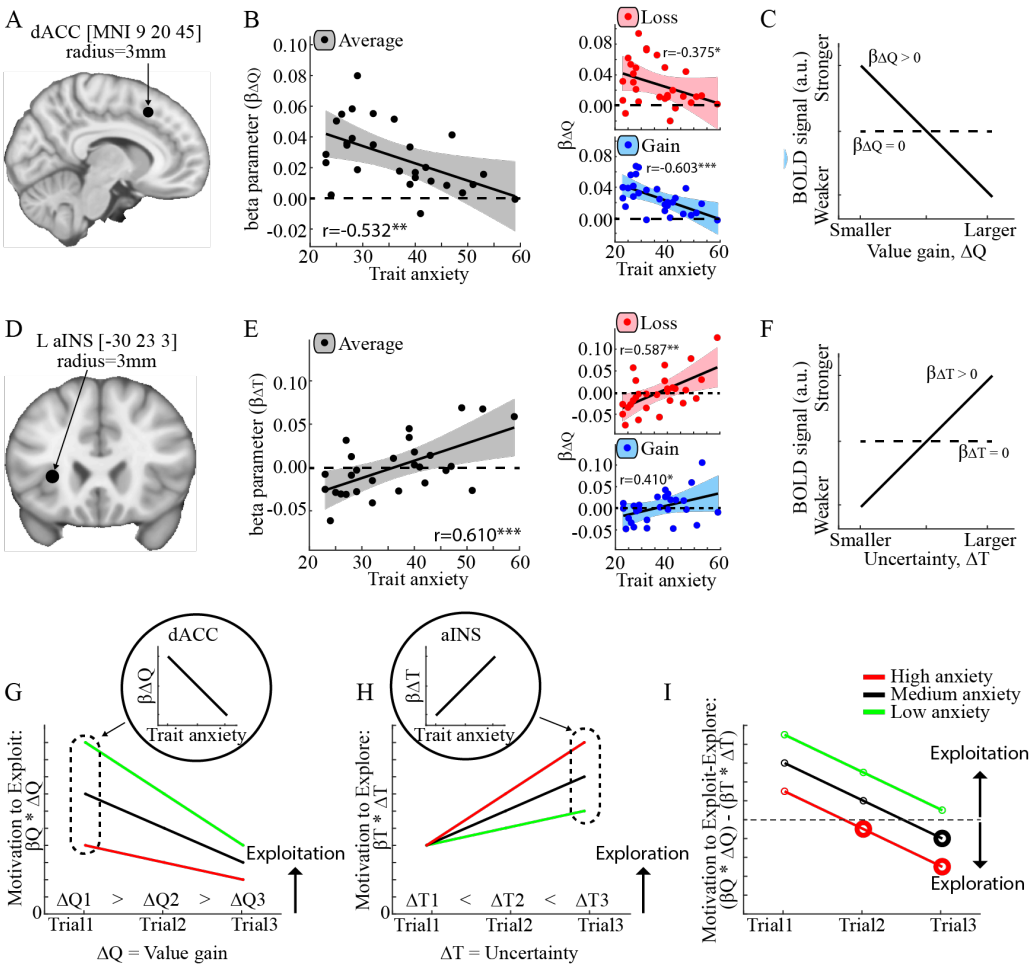

Figure 6. A-F. Brain correlates of differences in expected value (i.e. ΔQ) and uncertainty (i.e. ΔT) that is significantly modulated by Trait anxiety. Beta parameters were extracted from 3mm radius spheres (indicated by black dots in **A** and **D**) centered on the coordinates obtained from the independent localizer. **B.** $\beta_{\Delta Q}$ correlated negatively with trait anxiety in the dACC across conditions (left panel), and separately in the Loss (top right panel) and in the Gain condition (bottom right panel). **C.** To understand these results, consider that a positive $\beta_{\Delta Q}$ indicates that dACC activity increases when the expected values becomes increasingly similar, while a $\beta_{\Delta Q}$ close to zero (e.g. high anxiety) means that dACC activity does not change with ΔQ. **E.** $\beta_{\Delta T}$ correlated positively with trait anxiety in the bilateral aINS (here, only the left aINS is shown) across conditions (left panel), and separately in the Loss (top right panel) and in the Gain condition (bottom right panel). **F.** To understand these results, consider that a positive $\beta_{\Delta T}$ (e.g. high anxiety) indicates that aINS activity increases when the uncertainty increases, while a $\beta_{\Delta T}$ close to zero means that aINS activity does not change with ΔT. **G-I.** Schematic summary of the results. Consider a toy-example where the values of the machines are initially known (e.g. 40, 15, 25) and the best machine is selected in three consecutive trials which yield decreasing outcomes (e.g. 35-30-25). Accordingly, the differences in expected value (ΔQ) gradually decrease (**G**) while uncertainty of the non-selected machines (ΔT) increase (**H**). **G.** The motivation to exploit the best machine in a given trial is estimated by ΔQ scaled by an individual's weighting of expected-value gains (β_Q), i.e. $\Delta Q * \beta_Q$, where β_Q is smaller in high anxiety. The weighting of expected values may relate to the representation of ΔQ in the dACC, i.e. $\beta_{\Delta Q}$, which was close to zero in anxious individuals. **H.** The motivation to explore is estimated by ΔT scaled by an individual's weighting of uncertainty reduction (β_T), which was larger in high anxiety, i.e. $\Delta T * \beta_T$. The weighting of uncertainty may relate to the representation of ΔT in the aINS, i.e. $\beta_{\Delta T}$, which was larger in more anxious individuals. **I.** The exploit-explore trade-off is determined by the difference in the motivation to exploit versus explore, which is shifted towards exploration for more anxious individuals. ***p<0.001; **p<0.01; *p<0.05; ns p>0.05.

References

Aberg, K. C., K. C. Doell, and S. Schwartz. 2015. "Hemispheric Asymmetries in Striatal Reward Responses Relate to Approach-Avoidance Learning and Encoding of Positive-Negative Prediction Errors in Dopaminergic Midbrain Regions." *J Neurosci* 35 (43):14491-500. doi: 10.1523/JNEUROSCI.1859-15.2015.

Aberg, K. C., K. C. Doell, and S. Schwartz. 2016a. "The left hemisphere learns what is right: Hemispatial reward learning depends on reinforcement learning processes in the contralateral hemisphere." *Neuropsychologia* 89:1-13. doi: 10.1016/j.neuropsychologia.2016.05.023.

Aberg, K. C., K. C. Doell, and S. Schwartz. 2016b. "Linking Individual Learning Styles to Approach-Avoidance Motivational Traits and Computational Aspects of Reinforcement Learning." *PLoS One* 11 (11):e0166675. doi: 10.1371/journal.pone.0166675.

Aberg, K. C., J. Muller, and S. Schwartz. 2017. "Trial-by-Trial Modulation of Associative Memory Formation by Reward Prediction Error and Reward Anticipation as Revealed by a Biologically Plausible Computational Model." *Front Hum Neurosci* 11:56. doi: 10.3389/fnhum.2017.00056.

Addicott, M. A., J. M. Pearson, M. M. Sweitzer, D. L. Barack, and M. L. Platt. 2017. "A Primer on Foraging and the Explore/Exploit Trade-Off for Psychiatry Research." *Neuropsychopharmacology* 42 (10):1931-1939. doi: 10.1038/npp.2017.108.

Alvarez, R. P., N. Kirlic, M. Misaki, J. Bodurka, J. L. Rhudy, M. P. Paulus, and W. C. Drevets. 2015. "Increased anterior insula activity in anxious individuals is linked to diminished perceived control." *Transl Psychiatry* 5:e591. doi: 10.1038/tp.2015.84.

Aston-Jones, G., and J. D. Cohen. 2005. "An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance." *Annual Review of Neuroscience* 28:403-450. doi: 10.1146/annurev.neuro.28.061604.135709.

Badre, D., B. B. Doll, N. M. Long, and M. J. Frank. 2012. "Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration." *Neuron* 73 (3):595-607. doi: 10.1016/j.neuron.2011.12.025.

Baliki, M. N., A. Mansour, A. T. Baria, L. Huang, S. E. Berger, H. L. Fields, and A. V. Apkarian. 2013. "Parceling human accumbens into putative core and shell dissociates encoding of values for reward and pain." *J Neurosci* 33 (41):16383-93. doi: 10.1523/JNEUROSCI.1731-13.2013.

Balsamo, M., R. Romanelli, M. Innamorati, G. Ciccarese, L. Carlucci, and A. Saggino. 2013. "The State-Trait Anxiety Inventory: Shadows and Lights on its Construct Validity." *Journal of Psychopathology and Behavioral Assessment* 35 (4):475-486. doi: 10.1007/s10862-013-9354-5.

Bandura, A. 1982. "Self-Efficacy Mechanism in Human Agency." *American Psychologist* 37 (2):122-147. doi: Doi 10.1037/0003-066x.37.2.122.

Bar-Haim, Y., D. Lamy, L. Pergamin, M. J. Bakermans-Kranenburg, and IJzendoorn M. H. van. 2007. "Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study." *Psychol Bull* 133 (1):1-24. doi: 10.1037/0033-2909.133.1.1.

Bishop, C. M. 2006. *Pattern recognition and machine learning.*: Springer.

Bishop, S. J. 2007. "Neurocognitive mechanisms of anxiety: an integrative account." *Trends in Cognitive Sciences* 11:307-316. doi: 10.1016/j.tics.2007.05.008.

Blanchard, T. C., and S. J. Gershman. 2018. "Pure correlates of exploration and exploitation in the human brain." *Cogn Affect Behav Neurosci* 18 (1):117-126. doi: 10.3758/s13415-017-0556-2.

Bludau, S., S. B. Eickhoff, H. Mohlberg, S. Caspers, A. R. Laird, P. T. Fox, A. Schleicher, K. Zilles, and K. Amunts. 2014. "Cytoarchitecture, probability maps and functions of the human frontal pole." *Neuroimage* 93 Pt 2:260-75. doi: 10.1016/j.neuroimage.2013.05.052.

Browning, M., T. E. Behrens, G. Jocham, J. X. O'Reilly, and S. J. Bishop. 2015. "Anxious individuals have difficulty learning the causal statistics of aversive environments." *Nature Neuroscience* 18 (4):590-+. doi: 10.1038/nn.3961.

Buhr, K., and M. J. Dugas. 2009. "The role of fear of anxiety and intolerance of uncertainty in worry: an experimental manipulation." *Behav Res Ther* 47 (3):215-23. doi: 10.1016/j.brat.2008.12.004.

Burnham, K. P., and D. R. Anderson. 2004. "Multimodel inference - understanding AIC and BIC in model selection." *Sociological Methods & Research* 33 (2):261-304. doi: 10.1177/0049124104268644.

Carleton, R. N. 2016. "Into the unknown: A review and synthesis of contemporary models involving uncertainty." *J Anxiety Disord* 39:30-43. doi: 10.1016/j.janxdis.2016.02.007.

Cascio, C. N., S. H. Konrath, and E. B. Falk. 2015. "Narcissists' social pain seen only in the brain." *Soc Cogn Affect Neurosci* 10 (3):335-41. doi: 10.1093/scan/nsu072.

Cavanagh, J. F., C. M. Figueroa, M. X. Cohen, and M. J. Frank. 2012. "Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation." *Cereb Cortex* 22 (11):2575-86. doi: 10.1093/cercor/bhr332.

Chakroun, K., D. Mathar, A. Wiehler, F. Ganzer, and J. Peters. 2020. "Dopaminergic modulation of the exploration/exploitation trade-off in human decision-making." *Elife* 9. doi: 10.7554/eLife.51260.

Chambers, J. A., K. G. Power, and R. C. Durham. 2004. "The relationship between trait vulnerability and anxiety and depressive diagnoses at long-term follow-up of Generalized Anxiety Disorder." *Journal of Anxiety Disorders* 18 (5):587-607. doi: 10.1016/j.janxdis.2003.09.001.

Charnov, E. L. 1976. "Optimal foraging, the marginal value theorem." *Theor Popul Biol* 9 (2):129-36. doi: 10.1016/0040-5809(76)90040-x.

Charpentier, C. J., J. Aylward, J. P. Roiser, and O. J. Robinson. 2017. "Enhanced Risk Aversion, But Not Loss Aversion, in Unmedicated Pathological Anxiety." *Biological Psychiatry* 81 (12):1014-1022. doi: 10.1016/j.biopsych.2016.12.010.

Clark, L. A., D. Watson, and S. Mineka. 1994. "Temperament, Personality, and the Mood and Anxiety Disorders." *Journal of Abnormal Psychology* 103 (1):103-116. doi: Doi 10.1037/0021-843x.103.1.103.

Cohen, J. D., S. M. McClure, and A. J. Yu. 2007. "Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration." *Philos Trans R Soc Lond B Biol Sci* 362 (1481):933-42. doi: 10.1098/rstb.2007.2098.

Cohen, M. X. 2007. "Individual differences and the neural representations of reward expectation and reward prediction error." *Soc Cogn Affect Neurosci* 2 (1):20-30. doi: 10.1093/scan/nsl021.

Collins, A. G. E., and M. J. Frank. 2012. "How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis." *European Journal of Neuroscience* 35 (7):1024-1035. doi: 10.1111/j.1460-9568.2011.07980.x.

Coricelli, G., and A. Rustichini. 2010. "Counterfactual thinking and emotions: regret and envy learning." *Philosophical Transactions of the Royal Society B-Biological Sciences* 365 (1538):241-247. doi: 10.1098/rstb.2009.0159.

Dar, K. A., N. Iqbal, and A. Mushtaq. 2017. "Intolerance of uncertainty, depression, and anxiety: Examining the indirect and moderating effects of worry." *Asian J Psychiatr* 29:129-133. doi: 10.1016/j.ajp.2017.04.017.

Daw, N. D., J. P. O'Doherty, P. Dayan, B. Seymour, and R. J. Dolan. 2006. "Cortical substrates for exploratory decisions in humans." *Nature* 441 (7095):876-9. doi: 10.1038/nature04766.

De Martino, B., D. Kumaran, B. Seymour, and R. J. Dolan. 2006. "Frames, biases, and rational decision-making in the human brain." *Science* 313 (5787):684-687. doi: 10.1126/science.1128356.

Domenech, P., and E. Koechlin. 2015. "Executive control and decision-making in the prefrontal cortex." *Current Opinion in Behavioral Sciences* 1:101-106. doi: 10.1016/j.cobeha.2014.10.007.

Dunsmoor, J. E., and R. Paz. 2015. "Fear Generalization and Anxiety: Behavioral and Neural Mechanisms." *Biol Psychiatry*. doi: 10.1016/j.biopsych.2015.04.010.

Ehring, T., and E. R. Watkins. 2008. "Repetitive Negative Thinking as a Transdiagnostic Process." *International Journal of Cognitive Therapy* 1 (3):192-205. doi: DOI 10.1680/ijct.2008.1.3.192.

Eickhoff, S. B., K. E. Stephan, H. Mohlberg, C. Grefkes, G. R. Fink, K. Amunts, and K. Zilles. 2005. "A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data." *Neuroimage* 25 (4):1325-35. doi: 10.1016/j.neuroimage.2004.12.034.

Einstein, D. A. 2014. "Extension of the Transdiagnostic Model to Focus on Intolerance of Uncertainty: A Review of the Literature and Implications for Treatment." *Clin Psychol (New York)* 21 (3):280-300. doi: 10.1111/cpsp.12077.

Etkin, A., and T. D. Wager. 2007. "Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia." *Am J Psychiatry* 164 (10):1476-88. doi: 10.1176/appi.ajp.2007.07030504.

Eysenck, M. W., N. Derakshan, R. Santos, and M. G. Calvo. 2007. "Anxiety and cognitive performance: Attentional control theory." *Emotion* 7:336-353. doi: 10.1037/1528-3542.7.2.336.

Findling, C., V. Skvortsova, R. Dromnelle, S. Palminteri, and V. Wyart. 2019. "Computational noise in reward-guided learning drives behavioral variability in volatile environments." *Nature Neuroscience* 22 (12):2066-+. doi: 10.1038/s41593-019-0518-9.

Frank, M. J., B. B. Doll, J. Oas-Terpstra, and F. Moreno. 2009. "Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation." *Nat Neurosci* 12 (8):1062-8. doi: 10.1038/nn.2342.

Frank, M. J., L. C. Seeberger, and C. O'Reilly R. 2004. "By carrot or by stick: cognitive reinforcement learning in parkinsonism." *Science* 306 (5703):1940-3. doi: 10.1126/science.1102941.

Fung, B. J., S. Qi, D. Hassabis, N. Daw, and D. Mobbs. 2019a. "Slow escape decisions are swayed by trait anxiety." *Nature Human Behaviour* 3 (7):702-708. doi: 10.1038/s41562-019-0595-5.

Fung, B. J., S. Qi, D. Hassabis, N. Daw, and D. Mobbs. 2019b. "Slow escape decisions are swayed by trait anxiety." *Nat Hum Behav* 3 (7):702-708. doi: 10.1038/s41562-019-0595-5.

Gagne, C., O. Zika, P. Dayan, and S. J. Bishop. 2020. "Impaired adaptation of learning to contingency volatility in internalizing psychopathology." *Elife* 9. doi: 10.7554/eLife.61387.

Gershman, S. J. 2018. "Deconstructing the human algorithms for exploration." *Cognition* 173:34-42. doi: 10.1016/j.cognition.2017.12.014.

Gershman, S. J., B. Pesaran, and N. D. Daw. 2009. "Human reinforcement learning subdivides structured action spaces by learning effector-specific values." *J Neurosci* 29 (43):13524-31. doi: 10.1523/JNEUROSCI.2469-09.2009.

Gershman, S. J., and B. G. Tzovaras. 2018. "Dopaminergic genes are associated with both directed and random exploration." *Neuropsychologia* 120:97-104. doi: 10.1016/j.neuropsychologia.2018.10.009.

Gogolla, N. 2017. "The insular cortex." *Curr Biol* 27 (12):R580-R586. doi: 10.1016/j.cub.2017.05.010.

Grupe, D. W. 2017. "Decision-Making in Anxiety and Its Disorders." In *Decision Neuroscience - An Integrative Approach*, edited by J. C. Dreher and L. Tremblay, 327–338. Academic Press.

Grupe, D. W., and J. B. Nitschke. 2013. "Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective." *Nat Rev Neurosci* 14 (7):488-501. doi: 10.1038/nrn3524.

Hartley, C. A., and E. A. Phelps. 2012. "Anxiety and decision-making." *Biol Psychiatry* 72 (2):113-8. doi: 10.1016/j.biopsych.2011.12.027.

Hayden, B. Y., J. M. Pearson, and M. L. Platt. 2011. "Neuronal basis of sequential foraging decisions in a patchy environment." *Nat Neurosci* 14 (7):933-9. doi: 10.1038/nn.2856.

Hayes, A. F., and N. J. Rockwood. 2017. "Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation." *Behav Res Ther* 98:39-57. doi: 10.1016/j.brat.2016.11.001.

Jepma, M., and M. Lopez-Sola. 2014. "Anxiety and framing effects on decision making: insights from neuroimaging." *J Neurosci* 34 (10):3455-6. doi: 10.1523/JNEUROSCI.5352-13.2014.

Kayser, A. S., J. M. Mitchell, D. Weinstein, and M. J. Frank. 2015. "Dopamine, locus of control, and the exploration-exploitation tradeoff." *Neuropsychopharmacology* 40 (2):454-62. doi: 10.1038/npp.2014.193.

Knox, W. B., A. R. Otto, P. Stone, and B. C. Love. 2011. "The nature of belief-directed exploratory choice in human decision-making." *Front Psychol* 2:398. doi: 10.3389/fpsyg.2011.00398.

Kolling, N., T. E. Behrens, R. B. Mars, and M. F. Rushworth. 2012. "Neural mechanisms of foraging." *Science* 336 (6077):95-8. doi: 10.1126/science.1216930.

Kotov, R., W. Gamez, F. Schmidt, and D. Watson. 2010. "Linking "Big" Personality Traits to Anxiety, Depressive, and Substance Use Disorders: A Meta-Analysis." *Psychological Bulletin* 136 (5):768-821. doi: 10.1037/a0020327.

Laufer, O., D. Israeli, and R. Paz. 2016. "Behavioral and Neural Mechanisms of Overgeneralization in Anxiety." *Curr Biol* 26 (6):713-22. doi: 10.1016/j.cub.2016.01.023.

Laufer, O., and R. Paz. 2012. "Monetary loss alters perceptual thresholds and compromises future decisions via amygdala and prefrontal networks." *J Neurosci* 32 (18):6304-11. doi: 10.1523/JNEUROSCI.6281-11.2012.

Laureiro-Martinez, D., S. Brusoni, N. Canessa, and M. Zollo. 2015. "UNDERSTANDING THE EXPLORATION-EXPLOITATION DILEMMA: AN fMRI STUDY OF ATTENTION CONTROL AND DECISION-MAKING PERFORMANCE." *Strategic Management Journal* 36 (3):319-338. doi: 10.1002/smj.2221.

Lee, D. 2013. "Decision making: from neuroscience to psychiatry." *Neuron* 78 (2):233-48. doi: 10.1016/j.neuron.2013.04.008.

Levy, D. J., and P. W. Glimcher. 2012. "The root of all value: a neural common currency for choice." *Curr Opin Neurobiol* 22 (6):1027-38. doi: 10.1016/j.conb.2012.06.001.

Lissek, S. 2012. "Toward an account of clinical anxiety predicated on basic, neurally mapped mechanisms of pavlovian fear-learning: the case for conditioned overgeneralization." *Depress Anxiety* 29 (4):257-63. doi: 10.1002/da.21922 [doi].

Luszczynska, A., B. Gutierrez-Dona, and R. Schwarzer. 2005. "General self-efficacy in various domains of human functioning: Evidence from five countries." *International Journal of Psychology* 40 (2):80-89. doi: 10.1080/00207590444000041.

Maldjian, J. A., P. J. Laurienti, R. A. Kraft, and J. H. Burdette. 2003. "An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets." *Neuroimage* 19 (3):1233-9.

Maner, J. K., J. A. Richey, K. Cromer, M. Mallott, C. W. Lejuez, T. E. Joiner, and N. B. Schmidt. 2007. "Dispositional anxiety and risk-avoidant decision-making." *Personality and Individual Differences* 42 (4):665-675. doi: 10.1016/j.paid.2006.08.016.

Maner, J. K., and N. B. Schmidt. 2006. "The role of risk avoidance in anxiety." *Behavior Therapy* 37 (2):181-189. doi: DOI 10.1016/j.beth.2005.11.003.

Mehlhorn, K., B. R. Newell, P. M. Todd, M. D. Lee, K. Morgan, V. A. Braithwaite, D. Hausmann, K. Fiedler, and C. Gonzalez. 2015. "Unpacking the exploration-exploitation tradeoff: A synthesis of human and animal literatures." *Decision* 2 (3):191-215.

Mineka, S., and K. Oehlberg. 2008. "The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders." *Acta Psychologica* 127 (3):567-580. doi: 10.1016/j.actpsy.2007.11.007.

Muris, P. 2002. "Relationships between self-efficacy and symptoms of anxiety disorders and depression in a normal adolescent sample." *Personality and Individual Differences* 32 (2):337-348. doi: Doi 10.1016/S0191-8869(01)00027-7.

Oglesby, M. E., N. P. Allan, and N. B. Schmidt. 2017. "Randomized control trial investigating the efficacy of a computer-based intolerance of uncertainty intervention." *Behav Res Ther* 95:50-57. doi: 10.1016/j.brat.2017.05.007.

Palminteri, S., V. Wyart, and E. Koechlin. 2017. "The Importance of Falsification in Computational Cognitive Modeling." *Trends Cogn Sci* 21 (6):425-433. doi: 10.1016/j.tics.2017.03.011.

Paulus, M. P., and M. B. Stein. 2006. "An insular view of anxiety." *Biological Psychiatry* 60 (4):383-387. doi: 10.1016/j.biopsych.2006.03.042.

Paulus, M. P., and A. J. Yu. 2012. "Emotion and decision-making: affect-driven belief systems in anxiety and depression." *Trends Cogn Sci* 16 (9):476-83. doi: 10.1016/j.tics.2012.07.009.

Payzan-LeNestour, E., and P. Bossaerts. 2011. "Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings." *PLoS Comput Biol* 7 (1):e1001048. doi: 10.1371/journal.pcbi.1001048.

Pepperdine, E., C. Lomax, and M. H. Freeston. 2018. "Disentangling intolerance of uncertainty and threat appraisal in everyday situations." *J Anxiety Disord*. doi: 10.1016/j.janxdis.2018.04.002.

Pessiglione, M., and M. R. Delgado. 2015. "The good, the bad and the brain: Neural correlates of appetitive and aversive values underlying decision making." *Curr Opin Behav Sci* 5:78-84. doi: 10.1016/j.cobeha.2015.08.006.

Ploghaus, A., I. Tracey, J. S. Gati, S. Clare, R. S. Menon, P. M. Matthews, and J. N. Rawlins. 1999. "Dissociating pain from its anticipation in the human brain." *Science* 284 (5422):1979-81.

Pulcu, E., and M. Browning. 2019. "The Misestimation of Uncertainty in Affective Disorders." *Trends Cogn Sci* 23 (10):865-875. doi: 10.1016/j.tics.2019.07.007.

Resnik, J., N. Sobel, and R. Paz. 2011. "Auditory aversive learning increases discrimination thresholds." *Nat Neurosci* 14 (6):791-6. doi: 10.1038/nn.2802.

Roese, N. J., K. Epstude, F. Fessel, M. Morrison, R. Smallman, A. Summerville, A. D. Galinsky, and S. Segerstrom. 2009. "Repetitive Regret, Depression, and Anxiety: Findings from a Nationally Representative Survey." *Journal of Social and Clinical Psychology* 28 (6):671-688. doi: DOI 10.1521/jscp.2009.28.6.671.

Rosen, J. B., and J. Schulkin. 1998. "From normal fear to pathological anxiety." *Psychol Rev* 105 (2):325-50.

Scholl, J., and M. Klein-Flugge. 2018. "Understanding psychiatric disorder by capturing ecologically relevant features of learning and decision-making." *Behavioural Brain Research* 355:56-75. doi: 10.1016/j.bbr.2017.09.050.

Shenhav, A., M. A. Straccia, J. D. Cohen, and M. M. Botvinick. 2014. "Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value." *Nat Neurosci* 17 (9):1249-54. doi: 10.1038/nn.3771.

Simmons, A., S. C. Matthews, M. P. Paulus, and M. B. Stein. 2008. "Intolerance of uncertainty correlates with insula activation during affective ambiguity." *Neuroscience Letters* 430 (2):92-97. doi: 10.1016/j.neulet.2007.10.030.

Skvortsova, V., S. Palminteri, and M. Pessiglione. 2014. "Learning To Minimize Efforts versus Maximizing Rewards: Computational Principles and Neural Correlates." *Journal of Neuroscience* 34 (47):15621-15630. doi: 10.1523/Jneurosci.1350-14.2014.

Spielberger, C. D., R. L. Gorsuch, R. Lushene, P. R. Vagg, and G. A. Jacobs. 1983. *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Stanton, K., and D. Watson. 2014. "Positive and Negative Affective Dysfunction in Psychopathology." *Social and Personality Psychology Compass* 8 (9):555-567.

Stephan, K. E., W. D. Penny, J. Daunizeau, R. J. Moran, and K. J. Friston. 2009. "Bayesian model selection for group studies." *Neuroimage* 46 (4):1004-17. doi: 10.1016/j.neuroimage.2009.03.025.

Tomov, M. S., V. Truong, R. A. Hundia, and S. J. Gershman. 2020. "Dissociable neural correlates of uncertainty underlie different exploration strategies." *Nature Communications* 11 (1). doi: 10.1038/s41467-020-15766-z.

Tzovara, A., M. M. Murray, N. Bourdaud, R. Chavarriaga, R. Millan Jdel, and M. De Lucia. 2012. "The timing of exploratory decision-making revealed by single-trial topographic EEGanalyses." *Neuroimage* 60 (4):1959-69. doi: 10.1016/j.neuroimage.2012.01.136.

Vul, E., C. Harris, P. Winkielman, and H. Pashler. 2009. "Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition." *Perspect Psychol Sci* 4 (3):274-90. doi: 10.1111/j.1745-6924.2009.01125.x.

Wang, F., K. Peng, Y. Bai, R. Li, Y. Zhu, P. Sun, H. Guo, C. Yuan, P. Rotshtein, and J. Sui. 2016. "The Dorsal Anterior Cingulate Cortex Modulates Dialectical Self-Thinking." *Front Psychol* 7:152. doi: 10.3389/fpsyg.2016.00152.

Weger, M., and C. Sandi. 2018. "High anxiety trait: A vulnerable phenotype for stress-induced depression." *Neurosci Biobehav Rev* 87:27-37. doi: 10.1016/j.neubiorev.2018.01.012.

Wilson, R. C., and A. G. Collins. 2019. "Ten simple rules for the computational modeling of behavioral data." *Elife* 8. doi: 10.7554/eLife.49547.

Wilson, R. C., A. Geana, J. M. White, E. A. Ludvig, and J. D. Cohen. 2014. "Humans Use Directed and Random Exploration to Solve the Explore-Exploit Dilemma." *Journal of Experimental Psychology-General* 143 (6):2074-2081. doi: 10.1037/a0038199.

Wright, P., G. He, N. A. Shapira, W. K. Goodman, and Y. Liu. 2004. "Disgust and the insula: fMRI responses to pictures of mutilation and contamination." *Neuroreport* 15 (15):2347-51.

Xu, P., R. Gu, L. S. Broster, R. Wu, N. T. Van Dam, Y. Jiang, J. Fan, and Y. J. Luo. 2013. "Neural basis of emotional decision making in trait anxiety." *J Neurosci* 33 (47):18641-53. doi: 10.1523/JNEUROSCI.1253-13.2013.

Zumberg, K. M., C. Chang, and L. J. Sanna. 2008. "Does problem orientation involve more than generalized self-efficacy? Predicting psychological and physical functioning in college students." *Personality and Individual Differences* 45 (4):328-332. doi: 10.1016/j.paid.2008.04.017.