Behavioral elements and neural circuits underlying exploratory decisions in normative anxiety

Kristoffer C. Aberg, Ido Toren, Rony Paz

Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel

Abstract

Exploration reduces uncertainty about the environment and improves the quality of future decisions, but at the cost of uncertain and suboptimal outcomes. Although anxiety increases intolerance to uncertainty, it remains unclear whether and how anxiety modulates exploratory decision-making. We use a three-armed-bandit task in both loss and gain domains and find that higher trait-anxiety increases exploration, which in turn underlies an inverse-U-shaped relationship between anxiety and overall performance. We identify two distinct 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. We conclude that a shift in balance towards representation of uncertainty in the insula prevails over value representation in the dACC, and entails maladaptive decision-making in individuals with higher normal-range anxiety.

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

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.

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 [1,2]. 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 [2,3], although there is currently little evidence for a direct link between exploration in human decision making and disorders.

Exploration was shown to increase with environmental volatility [4], and evidence suggest that exploratory decisions are directed towards options with large outcome variabilities [5-8]. Imaging studies in humans have shown that exploratory decisions engage the anterior insulae (aINS) as well as prefrontal regions involved in executive control processes, such as the frontopolar cortex (FPC) and the dorsal anterior cingulate cortex (dACC) [5-11]. Interestingly, anxious individuals find uncertainty more aversive and are also more intolerant thereof [12-17], and recent studies have shown that various aspects of uncertainty correlate with a hyper-responsive aINS in anxiety [14,18]. In addition, anxiety also pertain to dysfunctional executive control processes in the prefrontal-cortex (PFC) which are believed to cause attentional biases towards salient events [19,20].

Together, 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 [14] and separately examined the relationship between uncertainty and exploration [5-8]. 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 [21-26], and activity in the dACC and aINS [27,28], little is known about the interaction between these factors and exploratory decisions. Even less is known about the interaction in normal-range anxiety, a major factor in modern lives affecting many individuals' daily function.

Here, we hypothesized that normative 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 (Figure 1A,B). After selection, the outcome currently associated with the selected machine was displayed. Exploration was promoted by varying the outcome associated with each machine 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) (Figure 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 (Figure 1A), or avoided losing points in a Loss condition where all outcomes are negative (Figure 1B).

Exploration increases with anxiety

We examined the direct relationship between Trait anxiety [29] and exploratory decision-making, and whether it interacts with outcome valence (Loss/Gain). An exploratory decision was defined as in previous studies [9,30]. 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 was modulated by anxiety (Fig.2A,B; 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). On average, exploration was higher in the Loss condition (Fig.2C; main effect of valence factor, $F(1,26)=4.81$, $p=0.037$, $\eta_p^2=0.156$, ANOVA; mean±SEM: Loss=0.175±0.016, Gain=0.149±0.014; no valence*anxiety interaction effect, $F(1,26)=1.85$, $p=0.19$, $\eta_p^2=0.066$, ANOVA). Therefore, even in the normal range, 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; Figure 1D,E), and on the other hand, high exploration would hinder the exploitation of the current optimal alternative (Frequent explorers; Figure 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.2D,E). This was the case in both the Loss condition (Fig.2D; nonlinear regression with intercept and standardized quadratic term $b^2_{Explore}=-0.293$, $p<0.001$) and in the Gain condition (Fig.2E; 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 1), and overall performance was similar for both valence conditions (Figure 2F; $t(27)=0.417$, $p=0.680$, paired *t*-test, mean±SEM: Loss=0.714±0.018; Gain=0.724±0.020, Cohens d=0.020). Therefore, our findings demonstrate that the individual tendency to explore modulates overall performance in an inverted U-shape fashion, as previously suggested [2,31].

Overall performance is inverse-U-shape modulated by anxiety

Although exploration increased with anxiety and overall performance was non-linearly modulated by exploration, the relationship between anxiety and performance still needs to be tested directly. We found that overall performance was also non-linearly (inverse U-shaped) modulated by anxiety, but only in the Loss condition (Fig.2G, nonlinear regression with intercept and standardized quadratic term $b^2_{Anxiety}$=-0.712, p<0.001), and not in the Gain condition (Fig.2H, standardized $b^2_{Anxiety}$=-0.228, p=0.198), with a significant difference between conditions (Actual difference: $b^2_{Explore}$ Loss-Gain=-0.483; Null distribution difference: $b^2_{Explore}$ Loss-Gain=-0.258, p=0.032, Monte-Carlo permutation test, Supplementary Note 2).

To examine further if it was 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 [32]. The relationship changed significantly when controlling for Exploration in both the Loss ($\Delta b^2_{Anxiety}$=-0.190, 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; Figure 1I; Supplementary Note 3). Therefore, the non-linear relationship between anxiety and overall performance is likely mediated by exploration.

Behavioral factors contributing to exploration

We examined the relative contribution of several behavioral factors by adapting a classical reinforcement-learning approach. All 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). We included three additional factors: First, because recent evidence suggests that exploration is aimed at reducing the outcome-variability associated with different options [5,8,33], we included an outcome-variability factor which was updated by a prediction-error scaled by a learning-rate (Equations 6-7). Second, an additional exploratory strategy that allows detecting changes in the environment is randomly switching machines between trials, and we therefore included a random-switching binary factor (Equation 9). Finally, 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, defined as the number of trials since a machine was last selected (Equation 8). After the models were fitted individually to each participant's behavior (using a soft-max choice probability function; Equation 10), the probability of selecting machine *i* in trial *t* depends on weights β_Q, β_U, β_T, and β_S, reflecting the contribution of expected-value $Q_i(t)$, outcome-variability $U_i(t)$, outcome-uncertainty $T_i(t)$, and random-switching $S_i(t)$, respectively. We then tested different combinations of these factors and found that the most parsimonious model included decision weights for expected-value (β_Q), outcome-uncertainty (β_T), and random-switching (β_S) (outcome-variance was not a significant contributor; Supplementary Note 4).

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.3A-C). Across participants, the model predicted actual selections well above chance-level in both the Loss (mean correct±SEM: 0.801±0.02) and the Gain condition (mean correct±SEM: 0.822±0.016), with no significant difference between valence conditions (Fig.3C; $t(27)$=1.314, p=0.2, paired *t*-test, Cohens d=0.248).

To further validate that the model captured relevant aspects of exploratory decisions [34], a model-derived probability of switching to a suboptimal machine was calculated across all trials and conditions. Replicating the aforementioned behavioral results, this model-derived probability of exploration was modulated by

anxiety (repeated measures two-way ANOVA, main effect of anxiety as linear covariate, F(1,26)=27.52, p<0.001, η_p^2=0.514, ANOVA), and by valence (main effect of valence, F(1,26)=4.35, p=0.047, η_p^2=0.143, 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.3D; r=0.630, p<0.001, two-tailed Pearson correlation) and the Gain condition (Fig.3E, r=0.628, p<0.001, two-tailed Pearson correlation), and was significantly larger on average in the Loss condition (Fig. 3F; mean±SEM: Loss=0.182±0.010; Gain=0.164±0.009). These results mirror the aforementioned behavioral results (Fig.2A-C).

Exploration in anxiety trades off value for uncertainty reduction

We used the set of parameters generated for each participant and condition to test how the different factors contribute to exploratory decisions and their modulation by anxiety and valence. We repeated the analysis separately for each of the four different factors (namely, for their contribution as quantified by $\beta_Q, \beta_T, \beta_S,$ and α_Q; Supplementary Note 5).

First, expected-value (β_Q) was significantly modulated by anxiety (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.4A; r=-0.421, p=0.026, two-tailed Pearson-correlation) and the Gain condition (Fig.4B; r=-0.424, p=0.025, two-tailed Pearson-correlation; but gain/loss did not differ across participants; Fig.4C; Supplementary Table 3). Additionally, β_Q was negatively correlated with the proportion of exploratory decisions in both the Loss (Supplementary Figure 5A; r=-0.704, p<0.001, two-tailed Pearson-correlation) and the Gain condition (Supplementary Figure 5B; 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 (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.4D; r=0.446, p=0.018, two-tailed Pearson correlation) and the Gain condition (Fig.4E; r=0.431, p=0.022, two-tailed Pearson correlation). Similarly, β_T correlated positively with the proportion of exploratory decisions in both Gain and Loss (Supplementary Figure 5C,D; Loss: r=0.412, p=0.029; Gain: r=0.589, p<0.001, two-tailed Pearson correlations; but Gain/Loss did not differ across participants, Fig.4F; Supplementary Table 3). These results imply that more anxious individuals showed a stronger tendency to reduce uncertainty.

Importantly, outcome-uncertainty was not correlated with expected-value (Fig.4D,E insets; 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), in both the Loss (Fig.4G; r=-0.583, p=0.001, two-tailed Pearson correlation) and the Gain condition (Fig.4H; r=-0.593, p<0.001, two-tailed Pearson correlation); and this difference was similar across participants in both conditions (Fig.4I; *t*(27)=1.273, p=0.214, two-tailed paired *t*-test; mean±SEM: Loss=0.063±0.022; Gain=0.041±0.021).

Finally, anxiety did not correlate with random-switching (β_S, Supplementary Table 3; Supplementary Figure 5A-C), or with the learning rate (α_Q, Supplementary Table 3; Supplementary Figure 5D-F), and neither factors (β_S or α_Q) correlated with the proportion of exploratory decisions (Supplementary Figure 6C,D).

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. Across participants and Gain/Loss condition, activity in the bilateral NAcc correlated significantly with prediction-errors (δ_Q; Supplementary Note 6; 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). Neither anxiety nor valence 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 6). These results are in accordance with previous literature and with the behavioral results suggesting no differences in learning (Supplementary Note 5).

Functional localization of brain regions involved in exploration versus exploitation

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 [35]. In accordance with previous studies [9-11], 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 trial-by-trial differences in expected-value between the selected machine and the average of the two rejected machines (ΔQ), and similarly for outcome-uncertainty (ΔT), as parametric modulators to identify their respective brain correlates ($\beta_{\Delta Q}, \beta_{\Delta T}$). Beta estimates were extracted from 3mm radius spheres centered on the peak voxels identified above 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.6A-C, $F(1,26)=10.242$, $p=0.004$, $\eta_p^2=0.283$, ANOVA), but not in any other region (Supplementary Note 7). Specifically, anxiety correlated negatively with $\beta_{\Delta Q}$ in both the Loss (Fig.6B; $r=0.375$, $p=0.049$, two-tailed Pearson correlation) and in the Gain condition (Fig.6C; $r=-0.603$, $p<0.001$, two-tailed Pearson correlation). Across participants, the average value of $\beta_{\Delta Q}$ was positive (Fig.6D; $F(1,26)=60.594$, $p<0.001$, ANOVA intercept term; Supplementary Table 4). Because $\beta_{\Delta Q}$ reflects the coupling between BOLD signal in the dACC and expected value differences ΔQ (Figure 6E), 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 anxious individuals is less responsive to changes in ΔQ ($\beta_{\Delta Q}$ close to zero).

We also tested if the inability to track expected-value differences can be attributed to anxious people perceiving the task as more difficult overall, as the dACC is engaged with increased task difficulty [36]. However, overall dACC activation did not correlate with anxiety, suggesting this is not the case (Supplementary Note 8). Additionally, anxiety did not affect the representation of expected-value in the vmPFC (Supplementary Note 7), a brain region directly involved in value representations [37,38].

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 $\beta_{\Delta T}$ (Figure 6F-I; 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 9). In contrast to expected-value in the dACC, anxiety was positively correlated with the representation of outcome-uncertainty in the aINS in both the Loss (Fig.6G; 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.6H; 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 (Figure 6J). Therefore, the aINS of 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.

Finally, we tested whether the tradeoff between expected-value and uncertainty-reduction observed in more anxious individuals (i.e. $\beta_Q - \beta_T$; Figure 4G,H) is related to a differential representation of expected-value in the dACC (Figure 6B,C) and outcome-uncertainty in the aINS (Figure 6G,H). We found that the strength of the relationship between anxiety and $\beta_Q - \beta_T$ decreased significantly when controlling for $\beta_{\Delta Q,dACC} - \beta_{\Delta T,aINS}$ in both the Loss (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 11).

This suggests that the 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 anxiety explored more in both Gain and Loss conditions. This exploration-exploitation bias was associated with two distinct neurocomputational mechanisms. First, anxious individuals were less influenced by immediate expected-values and showed a reduced correlation between activity in the dACC and expected value differences. Second, anxious individuals increased the reduction of outcome-uncertainty and showed an increased correlation between outcome-uncertainty and activity in the aINS (see Fig.6K,L,M 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.

It might seem surprising at first that anxious individuals explore more, as exploration entails selecting options with largely uncertain outcomes, and anxiety increases uncertainty aversion, both in terms of risk (i.e. known uncertainty) and ambiguity (i.e. unknown uncertainty; Grupe, 2013 #138; Hartley, 2012 #38}). However, most studies that assessed the impact of anxiety on decision making under uncertainty used paradigms where decisions were independent, i.e. one decision's outcome did not provide information that could benefit subsequent decisions [17]. Yet in most real-life situations, exploration provides information that benefits future decision making. For example, whether a new restaurant is good or bad and therefore worthy of a re-visit, can only be determined by first gathering information via an exploratory decision to eat there. The present study provides evidence that individual anxiety levels contribute to the trade-off between immediate benefits (i.e. gaining a reward or avoiding a risky/uncertain situation) and prospective decision making.

At the neuronal level, 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 reminiscent of those 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; [36,39]). Foraging theory posits that a decision to forage is optimal when the expected value difference between an exploited option and foraging options is small [40]. 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.

In contrast, anxious individuals showed a more positive coupling between 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 [41], is involved in the aversion and the intolerance of uncertainty [14], and is hyper-responsive to uncertainty in high anxiety [14,19]. Moreover, the aINS represents negative affective states and stimuli, including pain, disgust, anxiety, conditioned fear, negative pictures, and odors [42-45]. Therefore, one interpretation of the present results is that uncertainty-related activation of the aINS is perceived as aversive, and can be immediately avoided by decisions aimed at reducing uncertainty, namely exploration. Accordingly, a more rapid accumulation of uncertainty-related activation in the aINS facilitates decisions to explore, as displayed by more anxious individuals.

This idea is in accordance with the previously described link between anxiety and intolerance to uncertainty [13,46], and the idea that anxiety is underpinned by the fear of the unknown [47]. Yet, whether the fear of the unknown is relevant only in contexts with aversive outcomes (as opposed to also positive contexts) is currently under debate [14,48]. 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 is aversive. 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 [49], further supporting the notion that anxiety is associated with a general aversion to uncertainty. An interesting extension of these findings is that changing an individual's intolerance to uncertainty in a positive context, for example via cognitive bias modification [50], may cause a general reduction in intolerance to uncertainty that transfers to negative contexts.

Interestingly, this is similar to recent findings in studies of over-generalization [51,52]. There as well, aversive contexts led to wider stimulus generalization [53,54], similar to our finding here of an overall higher exploration in aversive contexts; but on the other hand, anxious individual showed wider generalization for both aversive and appetitive contexts [22], 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 [55].

It is important to highlight that anxious individuals were more sensitive to a particular type of uncertainty, namely uncertainty which depends on the dynamics of the environment, and which is not present in static environments. It has been shown that exploration depends on the volatility of the environment, such that exploration increases in more volatile settings [56]. 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. A previous study reported that anxiety decreased the ability to adjust learning rates as a function of volatility [21], and it has been suggested that anxious individuals misestimates uncertainty [57]. 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. How anxiety interacts with volatility to determine exploration remains to be tested.

Notably, outcome-variability, the variability of a machine's outcome, did not contribute to decision making/exploration in the present study. This finding is in accordance with previous studies using restless n-armed bandit tasks [9,58], but contrasts with other studies reporting that exploration was directed towards options with large outcome variability, as the 'clock-task' [5,33] or a two-armed bandit-task [8]. Because reinforcement learning interacts with working memory, such that the ability to make value-based decisions depends on working memory load [59], the ability to make directed exploratory decisions towards options with large outcome variability may similarly depend on working memory load. It is worth noting that most studies reporting directed exploration towards large outcome variability used tasks in which reward contingencies were static, i.e. one option provided the optimal outcome throughout the experiment, whereas studies in which reward contingencies change dynamically across trials do not report any influence from outcome variability. Therefore, outcome variance might impact differently on decision making in static and dynamic environments, and/or interact with working memory load.

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 [60] and reduce confidence in one's ability to make appropriate decisions [61]. Indeed, anxiety and depression have been associated with both increased feelings of regret [62,63] and reduced self-efficacy [64-66]. Moreover, high levels of trait anxiety is a risk-factor for chronic anxiety and depression [67,68], and their development and maintenance may depend on dysfunctional learning and decision making [69-73].

In summary, we identify two neurocomputational 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 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.

Methods

Participants

After having provided written consent according to the ethical regulations of the Weizmann Institute of Science, thirty-one participants joined the experiment, which was performed in accordance with the Declaration of Helsinki. 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.

Task

In each trial, participants decided which of three slot machines to play (Figure 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 (Figure 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 (Figure 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 (Figure 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 Figures 1C, E-H). 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 [74], and this effect is exaggerated in more anxious individuals [23,75]. 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 (Figure 1D,F). By contrast, participants had to avoid losing points in a Loss condition with only negative outcomes (Figure 1 E,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

A computational approach was used to determine whether a decision could be classified as exploratory or as an exploitation. 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. First, we 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:

$$p_i(t) = \begin{cases} 1-\varepsilon, & if\ Q_i(t) = Q_{max}(t) \\ \varepsilon, & otherwise \end{cases} \tag{3}$$

In the second approach, 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:

$$p_i(t) = \frac{e^{u_i(t)}}{\sum_j e^{u_j(t)}} \tag{4}$$

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:

$$p_i(t) = \frac{e^{Q_i(t)*\beta_Q}}{\sum_j e^{Q_j(t)*\beta_Q}} \tag{5}$$

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 [5,33]. 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:

$$U_i(t+1) = U_i(t) + \alpha_U * \delta_U(t) \tag{6}$$

$$\delta_U(t) = |\delta_Q(t)| = |R(t) - Q_i(t)| \tag{7}$$

Additionally, in a changing 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 is larger 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.

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

Finally, we modeled the impact of an individual's tendency to randomly switch between different options, 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.

$$S_i(t+1) = \begin{cases} 0, & if\ selected = i \\ 1, & otherwise \end{cases} \tag{9}$$

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, and β_S, which determine the influence of expected value $Q_i(t)$, outcome variability $U_i(t)$, outcome uncertainty $T_i(t)$, and random switching between trials $S_i(t)$, respectively.

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

The free parameters, i.e. α_Q, α_U, β_Q, β_U, β_T, and β_S, were fitted individually to each participant's behavior by minimizing the negative log-likelihood estimate:

$$-LLE = -\ln\left(\prod_1^n p_i(t)\right) \tag{11}$$

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).

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. For example, the full model (m_{QUTS}) incorporates all of the above parameters while model (m_Q) only incorporates the expected value $Q_i(t)$.

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.

$$AICc = 2 * k - 2 * \ln(LLE) + \frac{2*k^2 + 2*k}{n-k-1} \tag{12}$$

A Bayesian Model Selection procedure (Stephan et al., 2009), with AICc as model evidence was used to determine the model most likely to explain the observed data.

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 [32]. 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 [32]. 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 [5,9-11,14], while exploitation activates the ventromedial prefrontal cortex (vmPFC; [10]). Our fMRI analyses therefore focus on these brain regions. A dACC ROI was created using the WFU Pickatlas toolbox [76] in combination with a previously described procedure [77,78]: 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 [79,80]. 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 [37]. These ROIs were combined into one ROI mask.

Prediction error

The Nucleus Accumbens (NAcc) encodes prediction errors [81,82]. The NAcc ROI was obtained from a recent study trying to delineate the shell and core subregions of the NAcc [83]. Because only the right NAcc ROI has been made available for download by the authors (http://apkarianlab.northwestern.edu/publications/supplemental_materials.php), the left NAcc ROI was created by mirroring the right NAcc ROI to the left hemisphere.

Statistical analysis

To investigate how the neural correlates of exploratory decision making are modulated by trait anxiety and Gain/Loss contexts, 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 [35].

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 (see above paragraph). 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.

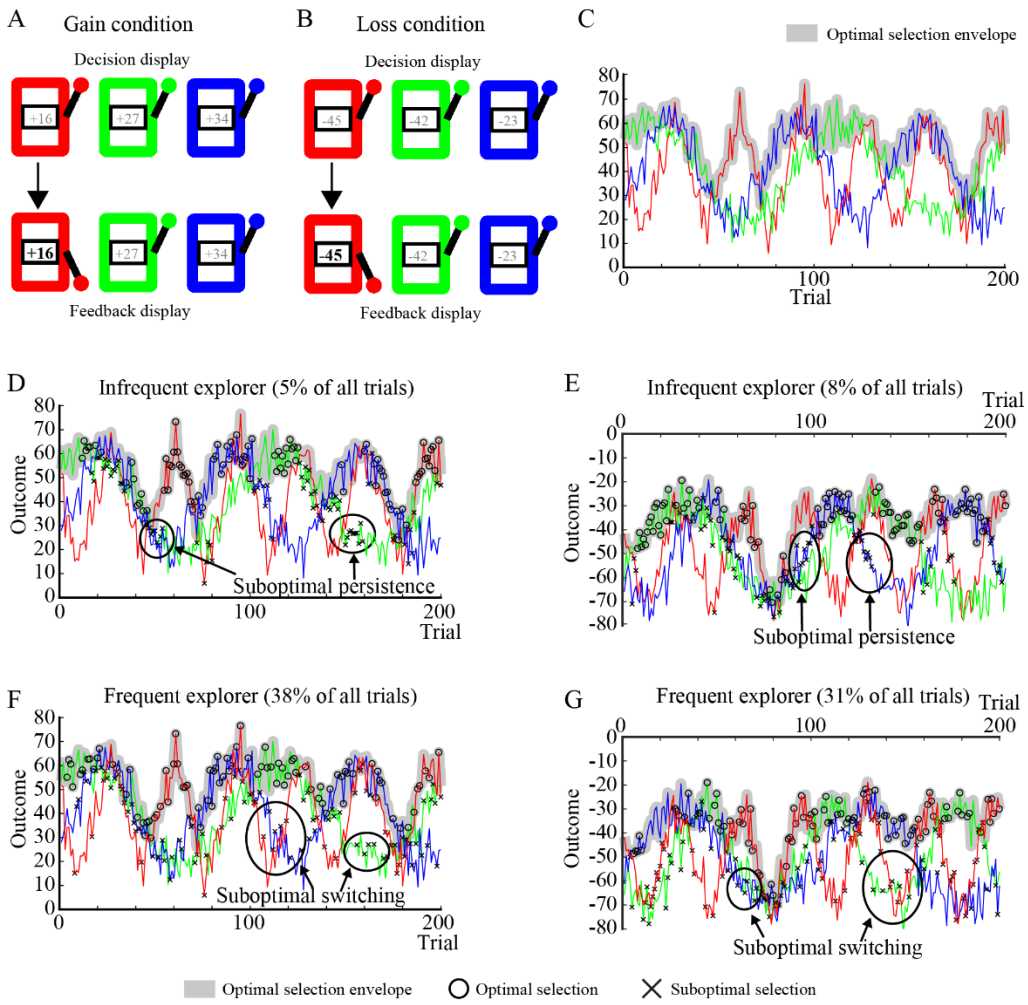

Figure 1. Paradigm and individual behavior. **A.** Illustration of Decision and Feedback displays in the Gain condition, where all outcomes are positive. During the decision display, the outcomes of each machine is hidden (grey characters), and participants decide to play one of the three slot machines 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.

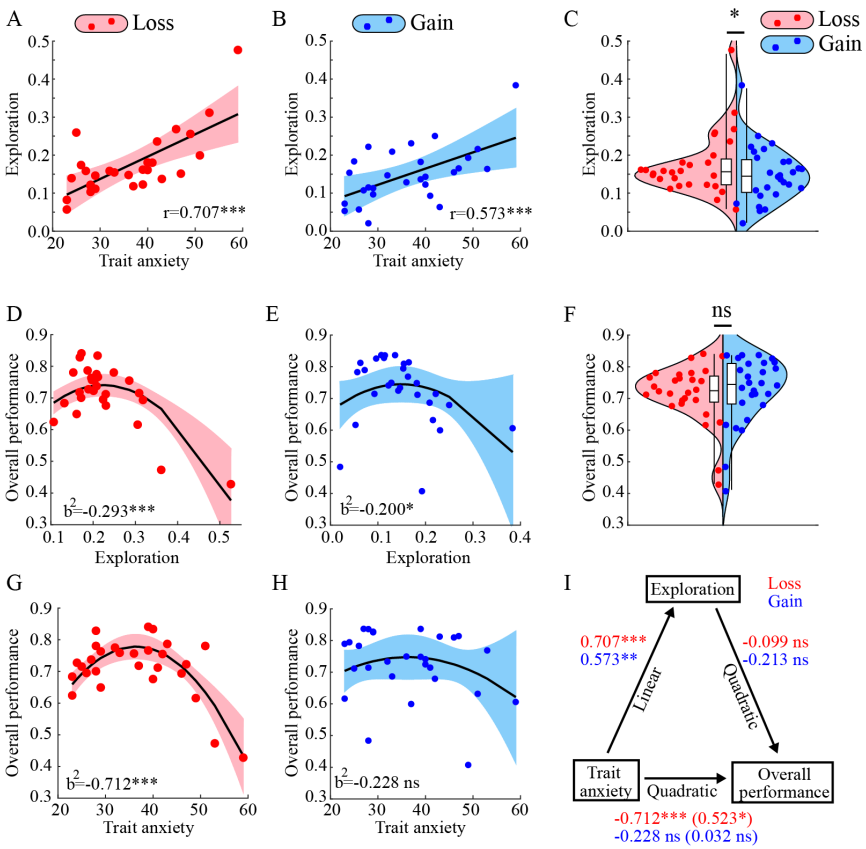

Figure 2. Behavioral results. **A-C.** Exploration, the proportion of exploratory decisions, increased linearly with increased levels of trait anxiety in the Loss (**A**) and in the Gain condition (**B**). Across participants, exploration was more frequent in the Loss (vs. the Gain) condition (**C**). **D-F.** Overall performance, defined as the correlation between maximal and actual outcomes across all trials, varied non-linearly with exploration in the Loss (**D**) and in the Gain condition (**E**). Overall performance did not differ between the Loss and the Gain condition (**F**). **G,H.** Overall performance varied significantly and non-linearly with trait anxiety in the Loss (**G**), but not in the Gain condition (**H**). **I.** 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 in both the Loss and the Gain condition, 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.

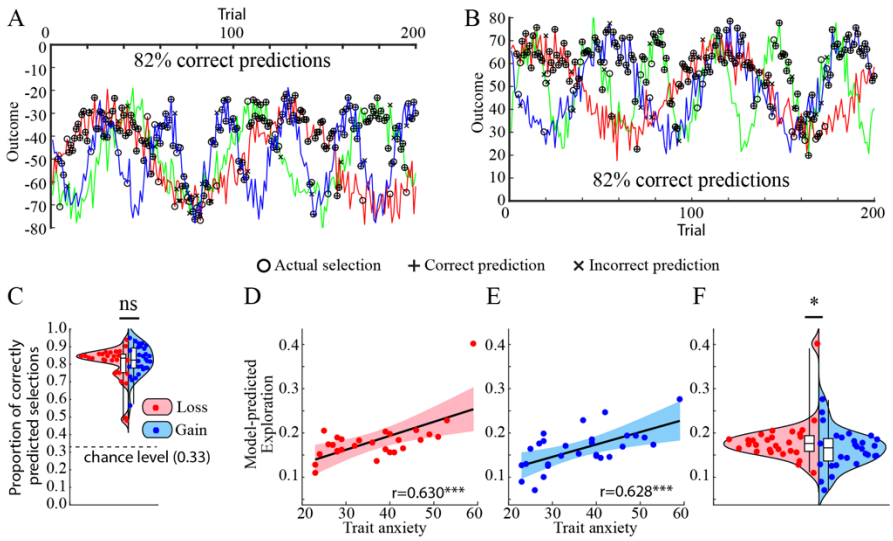

Figure 3. Modeling behavior. **A,B.** Representative examples displaying the similarity between actual and model-predicted selections in the Loss (**A**) and in the Gain condition (**B**). **C.** The average proportion of correctly predicted selections by the model was well above chance for all participants. **D-E.** The average model-predicted probability of exploring correlated positively with trait anxiety in the Loss (**D**) and in the Gain condition (**E**), and was higher in the Loss (vs. the Gain) condition (**F**). ***p<0.001; *p<0.05; ns p>0.05.

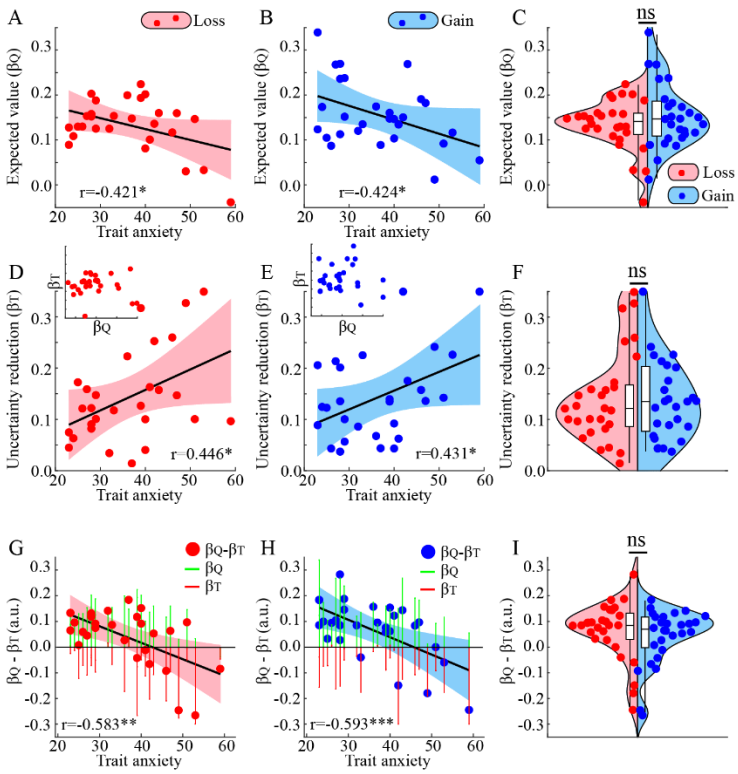

Figure 4. Individually fitted model parameters versus Trait anxiety and Gain/Loss condition. **A-C.** Decision weights estimating the influence of expected value (β_Q) were negatively correlated with trait anxiety in the Loss (**A**) and in the Gain condition (**B**), and did not differ between Gain/Loss conditions (**C**). **D-F.** Decision weights estimating reduction of outcome uncertainty (β_T) was positively correlated with trait anxiety in the Loss (**D**) and the Gain condition (**E**), but did not differ between Gain/Loss conditions (**F**). β_Q did not correlate with β_T in the Loss (**D, inset**) or in the Gain condition (**E, inset**). **G-H.** The difference between β_Q and β_T correlated negatively with trait anxiety in the Loss (**G**) and in the Gain condition (**H**), but did not differ between conditions (**I**). ***p<0.001; **p<0.01; *p<0.05; ns p>0.05.

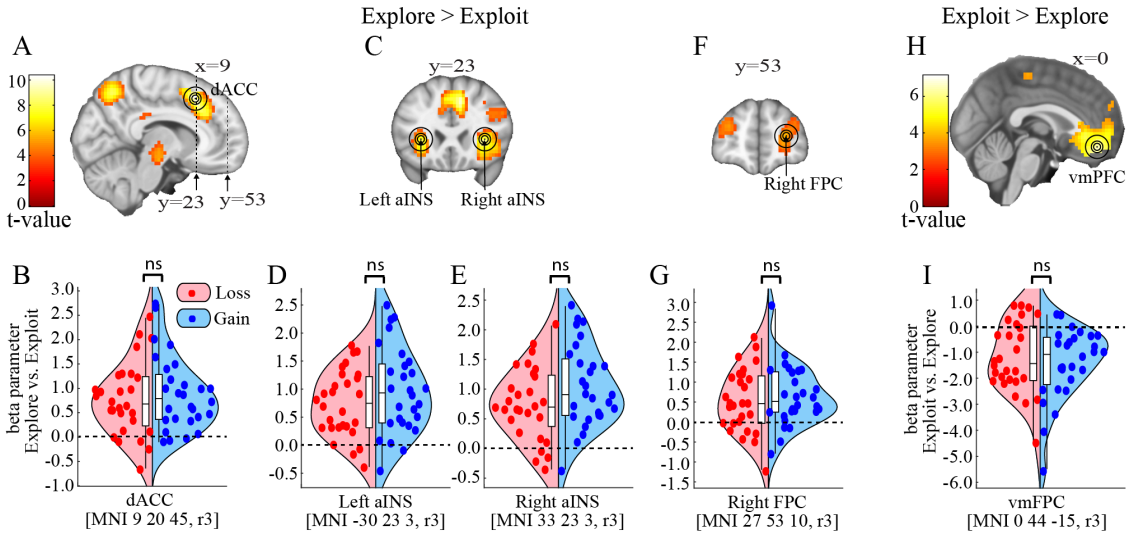

Figure 5. BOLD signal as a function of exploratory decision making. **A-G.** Decisions to explore (vs. exploit) increased activity in (**A,B)** the dorsal anterior cingulate cortex (dACC), (**C-E**) bilaterally in the anterior insula (aINS), and (**F,G**) in the right frontopolar cortex (FPC). Extracted 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.

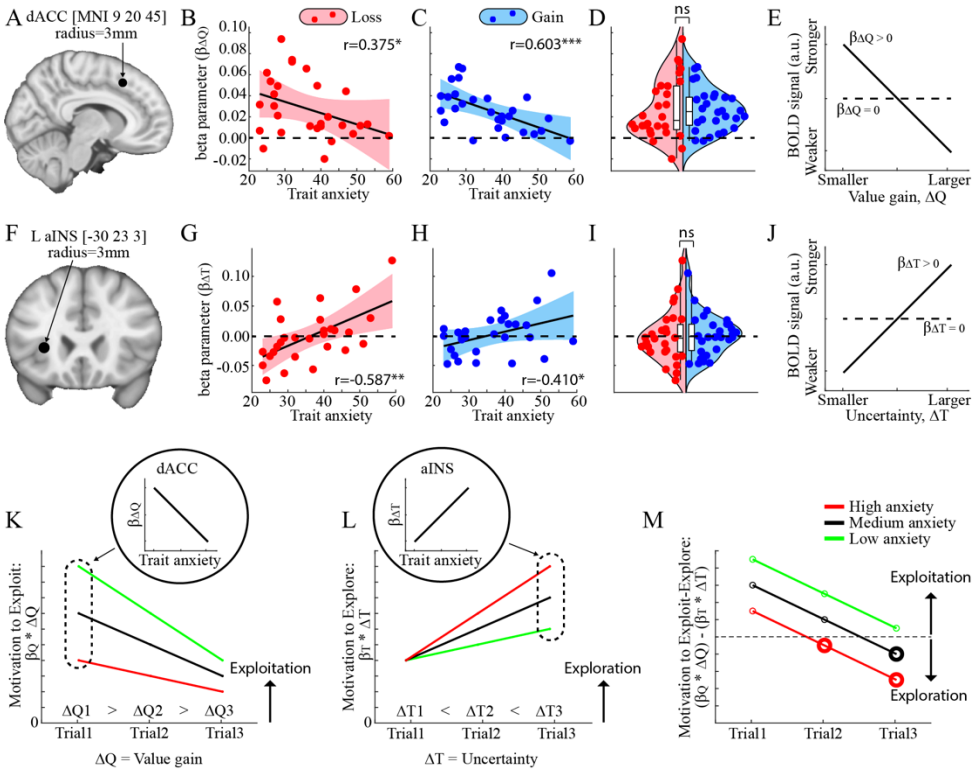

Figure 6. A-J. Brain correlates of differences in expected value (i.e. ΔQ) and uncertainty (i.e. ΔT) that is significantly modulated by Trait anxiety. Beta parameters were extracted from 3mm radius spheres (indicated by black dots) centered on the coordinates obtained from the independent localizer. **A-D.** $\beta_{\Delta Q}$ correlated negatively with trait anxiety in the dACC in the Loss (**B**) and in the Gain condition (**C**), but did not differ between Gain/Loss conditions (**D**). **E.** 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. **F-J.** $\beta_{\Delta T}$ correlated positively with trait anxiety in the bilateral aINS (here, only the left aINS is shown) in the Loss (**G**) and in the Gain condition (**H**), but did not differ between Gain/Loss conditions (**I**). **J.** 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. **K-M.** 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 (**K**) while uncertainty of the non-selected machines (ΔT) increase (**L**). **K.** The motivation to exploit the best machine in a given trial is estimated by ΔQ scaled by an individual's weighting of 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. **L.** 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. **M.** 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

1. Cohen, J.D., McClure, S.M. & Yu, A.J. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **362**, 933-942 (2007).
2. Addicott, M.A., Pearson, J.M., Sweitzer, M.M., Barack, D.L. & Platt, M.L. A Primer on Foraging and the Explore/Exploit Trade-Off for Psychiatry Research. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* **42**, 1931-1939 (2017).
3. Scholl, J. & Klein-Flugge, M. Understanding psychiatric disorder by capturing ecologically relevant features of learning and decision-making. *Behavioural brain research* **355**, 56-75 (2018).
4. Knox, B.W., Otto, A.R., Stone, P. & Love, B.C. The nature of belief-directed exploratory choice in human decision-making. *Frontiers in psychology* **3** (2012).
5. Badre, D., Doll, B.B., Long, N.M. & Frank, M.J. Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron* **73**, 595-607 (2012).
6. Cavanagh, J.F., Figueroa, C.M., Cohen, M.X. & Frank, M.J. Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cereb Cortex* **22**, 2575-2586 (2012).
7. Kayser, A.S., Mitchell, J.M., Weinstein, D. & Frank, M.J. Dopamine, locus of control, and the exploration-exploitation tradeoff. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* **40**, 454-462 (2015).
8. Tomov, M.S., Truong, V., Hundia, R.A. & Gershman, S.J. Dissociable neural correlates of uncertainty underlie different exploration strategies. *Nat Commun* **11** (2020).
9. Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B. & Dolan, R.J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876-879 (2006).
10. Laureiro-Martinez, D., Brusoni, S., Canessa, N. & Zollo, M. UNDERSTANDING THE EXPLORATION-EXPLOITATION DILEMMA: AN fMRI STUDY OF ATTENTION CONTROL AND DECISION-MAKING PERFORMANCE. *Strategic Manage J* **36**, 319-338 (2015).
11. Blanchard, T.C. & Gershman, S.J. Pure correlates of exploration and exploitation in the human brain. *Cognitive, affective & behavioral neuroscience* **18**, 117-126 (2018).
12. Hartley, C.A. & Phelps, E.A. Anxiety and decision-making. *Biological psychiatry* **72**, 113-118 (2012).
13. Buhr, K. & Dugas, M.J. The role of fear of anxiety and intolerance of uncertainty in worry: an experimental manipulation. *Behaviour research and therapy* **47**, 215-223 (2009).
14. Grupe, D.W. & Nitschke, J.B. Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective. *Nat Rev Neurosci* **14**, 488-501 (2013).
15. Maner, J.K. & Schmidt, N.B. The role of risk avoidance in anxiety. *Behav Ther* **37**, 181-189 (2006).
16. Maner, J.K.*, et al.* Dispositional anxiety and risk-avoidant decision-making. *Pers Indiv Differ* **42**, 665-675 (2007).
17. Charpentier, C.J., Aylward, J., Roiser, J.P. & Robinson, O.J. Enhanced Risk Aversion, But Not Loss Aversion, in Unmedicated Pathological Anxiety. *Biological psychiatry* **81**, 1014-1022 (2017).
18. Paulus, M.P. & Stein, M.B. An insular view of anxiety. *Biological psychiatry* **60**, 383-387 (2006).
19. Etkin, A. & Wager, T.D. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *The American journal of psychiatry* **164**, 1476-1488 (2007).
20. Bishop, S.J. Neurocognitive mechanisms of anxiety: an integrative account. *Trends Cogn Sci* **11**, 307-316 (2007).
21. Browning, M., Behrens, T.E., Jocham, G., O'Reilly, J.X. & Bishop, S.J. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature neuroscience* **18**, 590-+ (2015).
22. Laufer, O., Israeli, D. & Paz, R. Behavioral and Neural Mechanisms of Overgeneralization in Anxiety. *Curr Biol* **26**, 713-722 (2016).
23. Xu, P.*, et al.* Neural basis of emotional decision making in trait anxiety. *J Neurosci* **33**, 18641-18653 (2013).

24. Aberg, K.C., Doell, K.C. & Schwartz, S. 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**, 14491-14500 (2015).
25. Aberg, K.C., Doell, K.C. & Schwartz, S. Linking Individual Learning Styles to Approach-Avoidance Motivational Traits and Computational Aspects of Reinforcement Learning. *PloS one* **11**, e0166675 (2016).
26. Frank, M.J., Seeberger, L.C. & O'Reilly R, C. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* **306**, 1940-1943 (2004).
27. Skvortsova, V., Palminteri, S. & Pessiglione, M. Learning To Minimize Efforts versus Maximizing Rewards: Computational Principles and Neural Correlates. *J Neurosci* **34**, 15621-15630 (2014).
28. Pessiglione, M. & Delgado, M.R. The good, the bad and the brain: Neural correlates of appetitive and aversive values underlying decision making. *Curr Opin Behav Sci* **5**, 78-84 (2015).
29. Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R. & Jacobs, G.A. *Manual for the State-Trait Anxiety Inventory* (Consulting Psychologists Press, Palo Alto, CA, 1983).
30. Tzovara, A.*, et al.* The timing of exploratory decision-making revealed by single-trial topographic EEGanalyses. *NeuroImage* **60**, 1959-1969 (2012).
31. Aston-Jones, G. & Cohen, J.D. An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience* **28**, 403-450 (2005).
32. Hayes, A.F. & Rockwood, N.J. Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation. *Behaviour research and therapy* **98**, 39-57 (2017).
33. Frank, M.J., Doll, B.B., Oas-Terpstra, J. & Moreno, F. Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nature neuroscience* **12**, 1062-1068 (2009).
34. Palminteri, S., Wyart, V. & Koechlin, E. The Importance of Falsification in Computational Cognitive Modeling. *Trends Cogn Sci* **21**, 425-433 (2017).
35. Vul, E., Harris, C., Winkielman, P. & Pashler, H. Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. *Perspectives on psychological science : a journal of the Association for Psychological Science* **4**, 274-290 (2009).
36. Shenhav, A., Straccia, M.A., Cohen, J.D. & Botvinick, M.M. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nature neuroscience* **17**, 1249-1254 (2014).
37. Levy, D.J. & Glimcher, P.W. The root of all value: a neural common currency for choice. *Current opinion in neurobiology* **22**, 1027-1038 (2012).
38. Domenech, P. & Koechlin, E. Executive control and decision-making in the prefrontal cortex. *Curr Opin Behav Sci* **1**, 101-106 (2015).
39. Kolling, N., Behrens, T.E., Mars, R.B. & Rushworth, M.F. Neural mechanisms of foraging. *Science* **336**, 95-98 (2012).
40. Hayden, B.Y., Pearson, J.M. & Platt, M.L. Neuronal basis of sequential foraging decisions in a patchy environment. *Nature neuroscience* **14**, 933-939 (2011).
41. Simmons, A., Matthews, S.C., Paulus, M.P. & Stein, M.B. Intolerance of uncertainty correlates with insula activation during affective ambiguity. *Neurosci Lett* **430**, 92-97 (2008).
42. Gogolla, N. The insular cortex. *Current biology : CB* **27**, R580-R586 (2017).
43. Alvarez, R.P.*, et al.* Increased anterior insula activity in anxious individuals is linked to diminished perceived control. *Translational psychiatry* **5**, e591 (2015).
44. Ploghaus, A.*, et al.* Dissociating pain from its anticipation in the human brain. *Science* **284**, 1979-1981 (1999).
45. Wright, P., He, G., Shapira, N.A., Goodman, W.K. & Liu, Y. Disgust and the insula: fMRI responses to pictures of mutilation and contamination. *Neuroreport* **15**, 2347-2351 (2004).
46. Dar, K.A., Iqbal, N. & Mushtaq, A. Intolerance of uncertainty, depression, and anxiety: Examining the indirect and moderating effects of worry. *Asian journal of psychiatry* **29**, 129-133 (2017).
47. Carleton, R.N. Into the unknown: A review and synthesis of contemporary models involving uncertainty. *J Anxiety Disord* **39**, 30-43 (2016).

48. Einstein, D.A. Extension of the Transdiagnostic Model to Focus on Intolerance of Uncertainty: A Review of the Literature and Implications for Treatment. *Clinical psychology : a publication of the Division of Clinical Psychology of the American Psychological Association* **21**, 280-300 (2014).
49. Pepperdine, E., Lomax, C. & Freeston, M.H. Disentangling intolerance of uncertainty and threat appraisal in everyday situations. *J Anxiety Disord* (2018).
50. Oglesby, M.E., Allan, N.P. & Schmidt, N.B. Randomized control trial investigating the efficacy of a computer-based intolerance of uncertainty intervention. *Behav Res Ther* **95**, 50-57 (2017).
51. Dunsmoor, J.E. & Paz, R. Fear Generalization and Anxiety: Behavioral and Neural Mechanisms. *Biol Psychiatry* (2015).
52. Lissek, S. Toward an account of clinical anxiety predicated on basic, neurally mapped mechanisms of pavlovian fear-learning: the case for conditioned overgeneralization. *Depress Anxiety* **29**, 257-263 (2012).
53. Laufer, O. & Paz, R. Monetary loss alters perceptual thresholds and compromises future decisions via amygdala and prefrontal networks. *J Neurosci* **32**, 6304-6311 (2012).
54. Resnik, J., Sobel, N. & Paz, R. Auditory aversive learning increases discrimination thresholds. *Nature neuroscience* **14**, 791-796 (2011).
55. Eysenck, M.W., Derakshan, N., Santos, R. & Calvo, M.G. Anxiety and cognitive performance: Attentional control theory. *Emotion* **7**, 336-353 (2007).
56. Knox, W.B., Otto, A.R., Stone, P. & Love, B.C. The nature of belief-directed exploratory choice in human decision-making. *Frontiers in psychology* **2**, 398 (2011).
57. Pulcu, E. & Browning, M. The Misestimation of Uncertainty in Affective Disorders. *Trends Cogn Sci* **23**, 865-875 (2019).
58. Payzan-LeNestour, E. & Bossaerts, P. Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS computational biology* **7**, e1001048 (2011).
59. Collins, A.G.E. & Frank, M.J. How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *Eur J Neurosci* **35**, 1024-1035 (2012).
60. Coricelli, G. & Rustichini, A. Counterfactual thinking and emotions: regret and envy learning. *Philos T R Soc B* **365**, 241-247 (2010).
61. Bandura, A. Self-Efficacy Mechanism in Human Agency. *American Psychologist* **37**, 122-147 (1982).
62. Roese, N.J.*, et al.* Repetitive Regret, Depression, and Anxiety: Findings from a Nationally Representative Survey. *J Soc Clin Psychol* **28**, 671-688 (2009).
63. Ehring, T. & Watkins, E.R. Repetitive Negative Thinking as a Transdiagnostic Process. *Int J Cogn Ther* **1**, 192-205 (2008).
64. Muris, P. Relationships between self-efficacy and symptoms of anxiety disorders and depression in a normal adolescent sample. *Pers Indiv Differ* **32**, 337-348 (2002).
65. Zumberg, K.M., Chang, C. & Sanna, L.J. Does problem orientation involve more than generalized self-efficacy? Predicting psychological and physical functioning in college students. *Pers Indiv Differ* **45**, 328-332 (2008).
66. Luszczynska, A., Gutierrez-Dona, B. & Schwarzer, R. General self-efficacy in various domains of human functioning: Evidence from five countries. *Int J Psychol* **40**, 80-89 (2005).
67. Chambers, J.A., Power, K.G. & Durham, R.C. The relationship between trait vulnerability and anxiety and depressive diagnoses at long-term follow-up of Generalized Anxiety Disorder. *Journal of Anxiety Disorders* **18**, 587-607 (2004).
68. Weger, M. & Sandi, C. High anxiety trait: A vulnerable phenotype for stress-induced depression. *Neurosci Biobehav Rev* **87**, 27-37 (2018).
69. Mineka, S. & Oehlberg, K. The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. *Acta psychologica* **127**, 567-580 (2008).
70. Paulus, M.P. & Yu, A.J. Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends Cogn Sci* **16**, 476-483 (2012).
71. Lee, D. Decision making: from neuroscience to psychiatry. *Neuron* **78**, 233-248 (2013).

72. Rosen, J.B. & Schulkin, J. From normal fear to pathological anxiety. *Psychological review* **105**, 325-350 (1998).
73. Grupe, D.W. Decision-Making in Anxiety and Its Disorders. in *Decision Neuroscience - An Integrative Approach* (ed. J.C. Dreher & L. Tremblay) 327–338 (Academic Press, 2017).
74. De Martino, B., Kumaran, D., Seymour, B. & Dolan, R.J. Frames, biases, and rational decision-making in the human brain. *Science* **313**, 684-687 (2006).
75. Jepma, M. & Lopez-Sola, M. Anxiety and framing effects on decision making: insights from neuroimaging. *J Neurosci* **34**, 3455-3456 (2014).
76. Maldjian, J.A., Laurienti, P.J., Kraft, R.A. & Burdette, J.H. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage* **19**, 1233-1239 (2003).
77. Cascio, C.N., Konrath, S.H. & Falk, E.B. Narcissists' social pain seen only in the brain. *Soc Cogn Affect Neurosci* **10**, 335-341 (2015).
78. Wang, F.*, et al.* The Dorsal Anterior Cingulate Cortex Modulates Dialectical Self-Thinking. *Frontiers in psychology* **7**, 152 (2016).
79. Eickhoff, S.B.*, et al.* A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* **25**, 1325-1335 (2005).
80. Bludau, S.*, et al.* Cytoarchitecture, probability maps and functions of the human frontal pole. *NeuroImage* **93 Pt 2**, 260-275 (2014).
81. Aberg, K.C., Doell, K.C. & Schwartz, S. The left hemisphere learns what is right: Hemispatial reward learning depends on reinforcement learning processes in the contralateral hemisphere. *Neuropsychologia* **89**, 1-13 (2016).
82. Cohen, M.X. Individual differences and the neural representations of reward expectation and reward prediction error. *Soc Cogn Affect Neurosci* **2**, 20-30 (2007).
83. Baliki, M.N.*, et al.* Parceling human accumbens into putative core and shell dissociates encoding of values for reward and pain. *J Neurosci* **33**, 16383-16393 (2013).