Archival Report

Biological Psychiatry: CNNI

Keeping an Eye Out for Change: Anxiety Disrupts Adaptive Resolution of Policy Uncertainty

Amrita Lamba, Michael J. Frank, and Oriel FeldmanHall

ABSTRACT

BACKGROUND: Human learning unfolds under uncertainty. Uncertainty is heterogeneous with different forms exerting distinct influences on learning. While one can be uncertain about what to do to maximize rewarding outcomes, known as policy uncertainty, one can also be uncertain about general world knowledge, known as epistemic uncertainty (EU). In complex and naturalistic environments such as the social world, adaptive learning may hinge on striking a balance between attending to and resolving each type of uncertainty. Prior work illustrates that people with anxiety—those with increased threat and uncertainty sensitivity—learn less from aversive outcomes, particularly as outcomes become more uncertain. How does a learner adaptively trade-off between attending to these distinct sources of uncertainty to successfully learn about their social environment?

METHODS: We developed a novel eye-tracking method to capture highly granular estimates of policy uncertainty and EU based on gaze patterns and pupil diameter (a physiological estimate of arousal).

RESULTS: These empirically derived uncertainty measures revealed that humans (N = 94) flexibly switched between resolving policy uncertainty and EU to adaptively learn about which individuals can be trusted and which should be avoided. However, those with increased anxiety (n = 49) did not flexibly switch between resolving policy uncertainty and EU and instead expressed less uncertainty overall.

CONCLUSIONS: Combining modeling and eye-tracking techniques, we show that altered learning in people with anxiety emerged from an insensitivity to policy uncertainty and rigid choice policies, leading to maladaptive behaviors with untrustworthy people.

https://doi.org/10.1016/j.bpsc.2024.07.015

Adaptive social functioning requires individuals to efficiently resolve a multitude of uncertainty signals when interacting with others, especially in repeated interactions in which our dynamics with others are constantly evolving (1–5). Imagine, for instance, deciding whether to have a cup of coffee with a friend whom you recently had a disagreement with. In such scenarios, one may focus on reducing policy uncertainty (6)—figuring out which set of actions produce desirable outcomes (i.e., what should I do?). You might weigh the relative costs and benefits of talking to your friend: Would meeting for coffee allow both parties to settle their disagreement, or conversely would interacting simply escalate the conflict? To resolve policy uncertainty, one might rely on a history of observed outcomes to reach a decision. One can also focus on resolving epistemic uncertainty (EU) (7,8)—acquiring detailed information about others for the sake of knowledge. For example, a learner may also be motivated to gather additional information about their friend, either directly or indirectly, to estimate the exact value (i.e., benefits) of meeting for coffee. While this type of value-based knowledge can increase the precision of one's beliefs (8,9), acquiring detailed epistemic knowledge is cognitively taxing (7,10) and may not always offer additional benefit for optimizing outcomes (11).

While efficiently resolving each uncertainty signal provides distinct advantages for learning, overreliance on one signal at the expense of another can maladaptively bias behavior. If we focus too much on minimizing policy uncertainty, our behavior might become too rigid, and we may not learn sufficient information that would allow us to generalize across contexts (12,13). If we focus too much on reducing EU, we might end up investing too many resources gathering irrelevant information, which can ultimately slow learning (7,10). While it seems people can strike a delicate balance between optimizing rewarding outcomes and gathering additional knowledge when the opportunity arises, it is unknown how humans effectively manage such trade-offs.

Reinforcement learning (RL) frameworks elegantly illustrate this learning dichotomy. While value-based (e.g., Q-learning) models iteratively learn the expected values of each action (14,15), a form of epistemic knowledge, policy-based models (e.g., actor-critic) directly optimize choice policies that maximize rewards without explicitly learning the expected values (15–19). Although both strategies facilitate learning, it is often more expedient to directly optimize a policy by identifying the best set of actions, given that value-based methods exhibit slowed convergence to reliable expected values (12,20), and yet solely relying on policy optimization can prevent people from gathering value-based information that can be highly useful if one suddenly needs to transfer knowledge to a novel

ISSN: 2451-9022

problem (12,13). The tension between policy optimization and value-based learning is also observed in classic explore-exploit trade-offs (10,21). Across species and circumstances, agents are often confronted with the dilemma of repeating tried and true choices or selecting a new option that could provide useful information and potentially better outcomes.

Thus, despite the efficiency of policy optimization, humans should also assign utility to epistemic knowledge (11,21–23)—especially in the social world where epistemic information can help learners distinguish between the value of others (4,24).

The distinct advantages associated with reducing policy uncertainty and EU suggest adaptive behavior may indeed emerge from a combination of policy optimization and value-based learning strategies (25,26), yet how these strategies are combined to guide learning remains largely unknown (27). One way to effectively manage this inherent tension is by dynamically orienting attention toward value- or policy-based information as new task demands arise (28), such that information sampling patterns might unveil how distinct uncertainty signals are directly prioritized for learning (29–31).

In the current study, we tested the hypothesis that adaptive social learning would be characterized by flexible and frequent attention switching between policy and epistemic information to regulate learning rates. In particular, we evaluated whether people would first reduce policy uncertainty to improve task performance and then flexibly switch to gathering value-based information to further minimize EU. This requires tracking subjective experiences of policy uncertainty and EU in real time as people manage these competing demands. Prior work suggests that eye movements can provide a reliable readout of uncertainty (32,33), revealing what information is being attended to, i.e., expected values or information related to the choice policy (34,35). Moreover, using gaze patterns to measure uncertainty offers the advantage of sidestepping issues with existing measures (36,37) that constrain the granularity of subjective uncertainty estimates to a single point estimate (Likert scales) and thereby omit critical details about the degree and type of uncertainty experienced. Given that fluctuations in uncertainty are also accompanied by increased physiological arousal (38–40), we can further index pupil-based arousal to the rate of learning adjustment (41–44).

Finally, we also explored whether increased uncertainty sensitivity disrupts one's ability to leverage policy uncertainty and EU to effectively guide learning. It is well known that individuals with increased trait anxiety experience increased distress and intolerance toward uncertainty (45–49), and this hypersensitivity may affect one's ability to swiftly adjust behavior in uncertain environments (42,50–52). Although prior work hints that individuals with anxiety fail to expediently adjust their behavior when policy uncertainty increases (52), altered learning could potentially emerge because individuals with anxiety invest disproportionate cognitive resources gathering social knowledge. Given that the coupling between uncertainty and physiological arousal is blunted in people with increased anxiety (42), we can additionally leverage gaze patterns and pupillometry to directly test whether failures in tracking policy uncertainty or EU impinge on learning.

In the current research, we constructed empirically derived estimates of policy uncertainty and EU from gaze patterns, allowing us to examine how individuals direct their attention toward policy uncertainty and EU signals as social interactions unfold. To dissociate between policy uncertainty and EU, we implemented a novel eye-tracking procedure in which participants indicated trial-level predictions about another's trustworthiness using their eye gaze. We found that task performance was predicted by how quickly an individual first resolved policy uncertainty and was then able to flexibly switch to resolving residual value-based EU. Adaptive switching was yoked to how much a partner's behavior changed during the task (i.e., becoming increasingly untrustworthy) and was reflected in pupil-based arousal. However, the behavioral and physiological fingerprint of flexibly switching between resolving different types of uncertainties was altered in highly anxious individuals. We fit a Bayesian RL (BRL) model that further revealed that people with increased anxiety were slower to adjust their behavior as partners became increasingly untrustworthy because of a tendency to perseverate on reward history, even when learned values no longer reflected the statistics of the environment. We found that reduced behavioral adjustment to untrustworthy partners was related to optimistic beliefs that partners would return a greater sum of the investment. Finally, when partners were untrustworthy, highly anxious individuals expressed less policy uncertainty, as assessed via both gaze- and model-derived measures.

METHODS AND MATERIALS

Participants

Participants (N = 100, n_{female} = 47, n_{male} = 53, mean age = 20.41 years) were recruited from the subject research pool managed by the Department of Cognitive and Psychological Sciences at Brown University in Providence, Rhode Island. All participants received either monetary compensation ($15/hour) or course credit, including additional performance-based bonus payment of up to $20. Six participants were excluded from final analyses because they did not adequately perform the task (i.e., no behavioral variability in the choice data) or due to poor gaze and pupil readout from the eye tracker, resulting in a final sample of 94 participants.

Evaluating Anxiety

Participants were grouped as a function of low and high anxiety levels based on their responses from the Generalized Anxiety Disorder (GAD-7) scale and the State-Trait Anxiety Inventory (STAI) (see the Supplement for distribution of anxiety scores). Our low- and high-anxiety groups were based on established clinical guidelines, in which a score of 10+ on the GAD-7 scale (53,54) and a score of 40+ on the trait component of the STAI scale (55) are reliable predictors of pathological anxiety (i.e., disruptive to one's daily functioning and well-being). For our statistical analyses of group differences, participants were placed in the high-anxiety group if they were above the significance cutoff on one or both anxiety inventories and in the low-anxiety group if they were below the cutoff on both the GAD-7 and STAI scales (see the Supplement for behavioral analyses using continuous anxiety scores). Based on these criteria, our sample was split into roughly evenly sized high- and low-anxiety groups ($n_{\text{low anxiety}}$ = 45, $n_{\text{high anxiety}}$ = 49). Of note, assessed anxiety levels in our study did not evaluate whether participants met the criteria for a DSM-5 generalized anxiety disorder.

Task Design

Participants completed 96 trials of the dynamic trust game based on the task design developed in our prior work (52) and interacted with 3 distinct partners that varied in their trustworthiness. Unbeknownst to participants, their partners were preprogrammed, slowly drifting in their reward rate over the course of the task, thereby requiring participants to continually adjust their choice policy to optimize rewards, the amount of money a partner reciprocated back to the participant (Figure 1A, B; see the Supplement for details). The task consisted of 2 distinct trial types that influenced which type of uncertainty should be more salient at a particular moment. Adjust policy trials occur when the amount of money returned by the partner has just crossed the outcome boundary (Figure 1C). On these trials, one's prior choice policy no longer maximizes their earnings, which naturally increases policy uncertainty until the policy is revised. In contrast, exploit policy trials comprise trials where we expect learning to have stabilized (i.e., at the end of a window in which partners were consistently trustworthy or untrustworthy). In these time

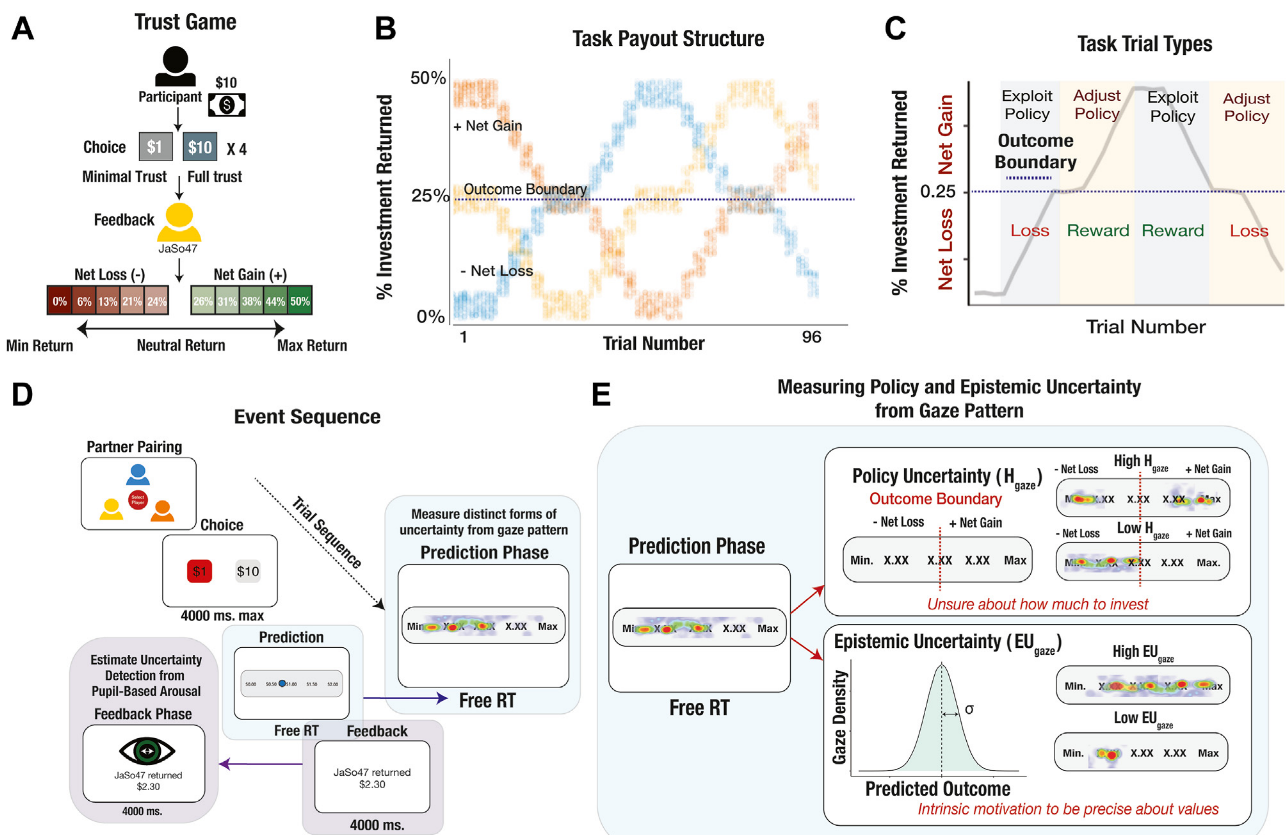

Figure 1. Experimental design and eye-tracking method to empirically estimate uncertainty. **(A)** Trust game. At the start of each trial, participants were paired with 1 of 3 presumed online partners and could invest $1 or $10. The invested money was then quadrupled in value, and the partner received the quadrupled sum. Partners then decided to return anywhere from 0% to 50% of the investment such that participants could lose all of the initial investment (0% return), double their investment (50% return), or receive any outcome in between. **(B)** Partner payout structure. Social partners gradually reversed their payouts and drifted toward increasingly fair (trustworthy) or selfish (untrustworthy) strategies, requiring participants to continually adjust their choice policy. The dotted black line denotes the outcome boundary determining the optimal policy to maximize returns: participants should invest maximally ($10) when partners return more than 25% and minimally ($1) otherwise to avoid a monetary loss. **(C)** Task trial types. During adjust policy trials, the partner's return rate crossed the 25% outcome boundary, requiring an adjustment in one's current choice policy. In contrast, during exploit policy trials, participants could continue using their current choice policy with no consequence to their current payoffs (assuming choice was already optimized). Trial types were orthogonalized to dissociate the effects of each type of uncertainty and outcome valence on learning. **(D)** Task event sequence. Trials commenced with a partner pairing phase in which the computer selected a partner for the current round (see Methods and Materials). Participants were then given up to 4 seconds to indicate their investment and then made a prediction about how much money they believed their partner would return, before observing trial outcomes. Gaze measures of uncertainty were collected during the prediction phase, and pupil-linked arousal was measured during the feedback phase. **(E)** Measuring policy uncertainty and EU from gaze patterns. During the prediction phase, participants were instructed to align their eye gaze (signaled to them with a blue dot) to the anticipated trial outcome. The gaze pattern over values was used to construct estimates of policy uncertainty and EU. Policy uncertainty was assessed as the extent to which gaze patterns spanned both sides of the boundary determining the optimal policy, quantified by entropy (*H*-gaze) (see Methods and Materials). EU was quantified by the variance in gaze patterns irrespective of the outcome boundary. RT, response time.

windows, participants have generally learned the optimal choice policy and can therefore use this opportunity to resolve residual EU about partners. Critically, adjust and exploit policy trials were crossed with reward and loss blocks (i.e., when partners were trustworthy or untrustworthy, respectively) to dissociate the effects of policy reliability and outcome valence on behavior (see the Supplement for in-depth task details).

Gaze Measures of Uncertainty

To obtain trial-level estimates of uncertainty, we asked participants to predict their partner's behavior (amount of money reciprocated) before observing trial outcomes (Figure 1D). Participants indicated their predictions using a response bar that displayed all possible monetary returns given the amount invested (Figure 1E). A blue dot on the screen corresponded to the participant's gaze, allowing participants to lock in their predictions of how much money their partners would return by moving the blue dot with their eyes to the predicted monetary outcome. This enabled us to use gaze patterns to evaluate participants' trial-by-trial expectations and their experienced uncertainty about anticipated outcomes, which is thought to govern the rate of learning. Thus, by leveraging gaze patterns, we derived distinct policy uncertainty and EU estimates. Furthermore, if individuals flexibly switch between resolving policy uncertainty and EU, then this attentional reorientation should also be reflected in physiological correlates of arousal. To quantify trial-level estimates of physiological arousal, we measured the percent change in pupil diameter from baseline at the time of feedback (Figure 1D; see Methods and Materials).

To estimate policy uncertainty, we borrowed insights from prior computational models that adjust learning as a function of policy uncertainty, quantified by entropy (H), over choice probabilities (6), which reliably captures human choice data in the current task (52). We derived an analogous measure of policy uncertainty based on gaze patterns, quantified by computing the entropy of the proportion of gazes on either side of the outcome boundary (H-gaze) (see the Supplement for entropy computation details). The midpoint—which we refer to as the outcome boundary—is not explicitly marked but acts as a psychological boundary indexing whether participants expect to earn or lose money on the current trial, thus determining whether they should invest or not (Figure 1E). An increase in gaze fixations on both sides of the boundary indicates greater uncertainty about the optimal choice policy and should thus increase learning rates. However, even when one might be relatively certain about what they should do in the task (e.g., invest maximally or minimally), they may still experience residual uncertainty about the specific outcomes on a given trial (i.e., exactly how much their partner will return), motivating the pursuit of epistemic knowledge. EU about how much money would be returned was quantified as the standard deviation in gaze patterns over the range of outcomes (EU-gaze), which captures the precision of one's predictions. See example mappings between gaze trajectories and uncertainty estimates for a prototypical participant in Figure 2A. Eye-tracking procedures and preprocessing steps are detailed in the Supplement.

RESULTS

People Dynamically Reoriented Their Attention Toward Policy Uncertainty and EU Signals

Using our gaze-derived uncertainty measures, we examined whether knowledge about the optimal policy (i.e., invest minimally or maximally) captured the expression of distinct sources of uncertainty. H-gaze was significantly greater on trials with suboptimal investments (i.e., choices that were inconsistent with the optimal choice policy) than optimal investments ($t = -8.05$, $p < .001$). In contrast, EU-gaze was greater on trials in which participants selected the optimal response, suggesting that participants instead expressed uncertainty about exact monetary outcomes once they knew the optimal choice policy ($t = 7.07$, $p < .001$) (Supplement; Figure S1). We next tested whether the physiological expression of policy uncertainty and EU depends on policy reliability (i.e., whether one can exploit or must adjust their current policy) and the trustworthiness of their partner. We observed just this: Policy uncertainty was greater during the adjust policy trials (main effect of trial type: $t = 4.47$, $p < .001$), whereas EU was greater during the exploit policy trials (main effect of trial type: $t = -2.89$, $p < .001$; uncertainty type $\times$ trial type interaction $t = 4.82$, $p < .001$) (Figure 2B, C). Moreover, participants expressed greater policy uncertainty and EU when partners were untrustworthy (main effect of outcome valence on H-gaze: $t = 12.94$, $p < .001$ and EU-gaze: $t = 8.44$, $p < .001$), suggesting that participants experienced more uncertainty about what to do when partners were selfish, and were more motivated to gather precise knowledge about untrustworthy people. In sum, although expressed policy uncertainty and EU were correlated ($t = 30.29$, $p < .001$), we also observed that policy uncertainty and EU signals are dissociable when choice policies need to be adjusted.

As an independent validation of the mappings between our theoretically informed gaze metrics and task demands, we evaluated the role of model-free gaze signatures that were analogous to H-gaze and EU-gaze calculations. Using the pattern of eye movements from the prediction phase, we quantified the number of times the participant's gaze traversed the midpoint of predicted outcomes on each trial. Indeed, the number of switches across the outcome boundary increased during adjust policy trials ($t = 5.57$, $p < .001$) and when partners were untrustworthy ($t = 9.50$, $p < .001$), recapitulating H-gaze patterns (Figure 2D). Information sampling patterns should also vary as participants adeptly learn to predict their partner's behavior, particularly during exploit policy trials when choice policies were presumably optimized. In particular, as participants learned to anticipate their partner's return, they may be motivated to expediently gather information and refine their predictions; thus, the rate of sampling may increase during periods of optimized behavior. To test this, we computed the sampling rate (values sampled/s) as a model-free metric of information sampling expediency. Consistent with this hypothesis, the sampling rate was greater during exploit versus adjust trials ($t = -2.14$, $p < .032$) (Figure 2D), indicating that the way participants sampled information in the task was indeed sensitive to policy reliability, recapitulating our EU findings.

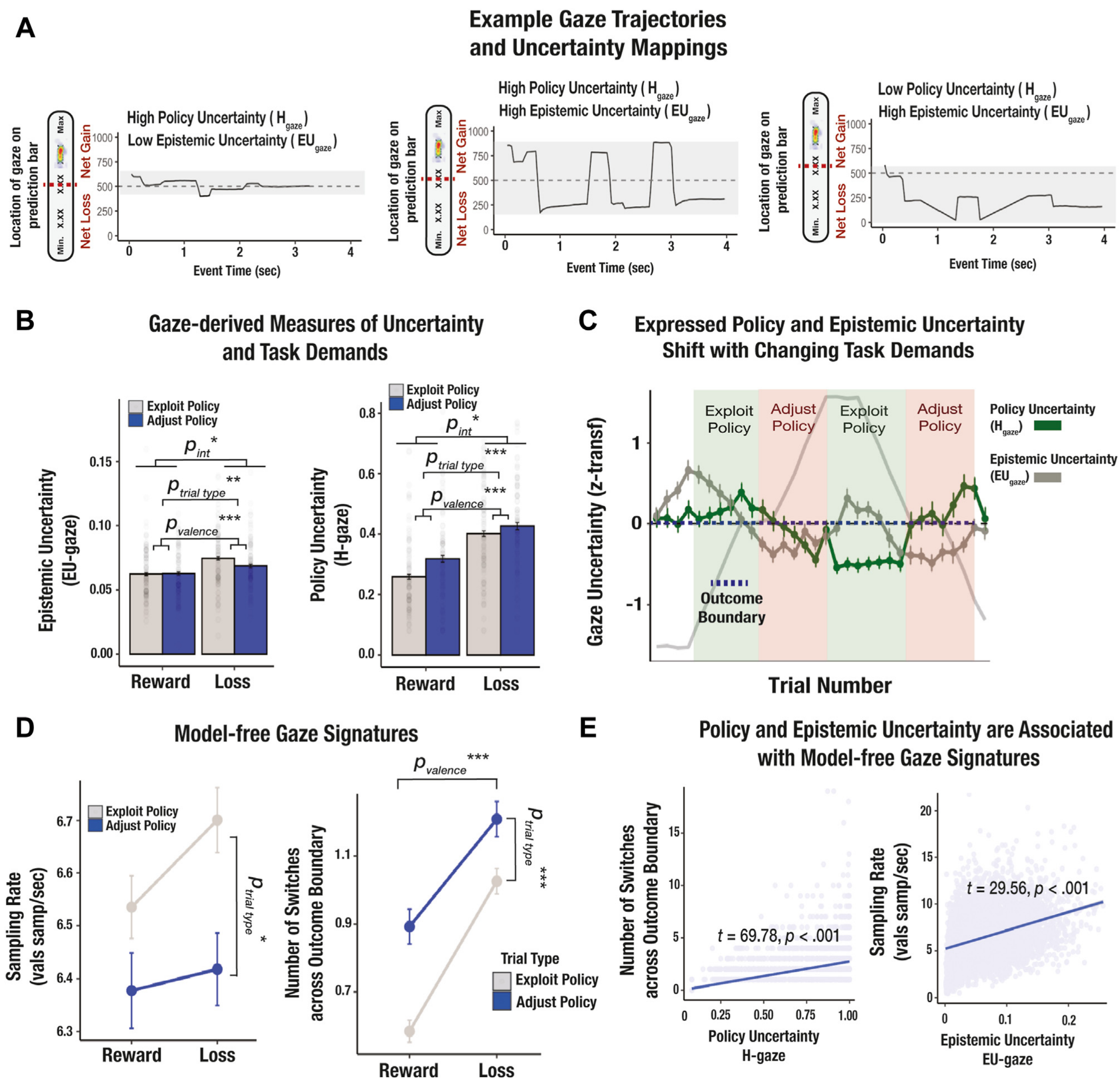

Figure 2. Expressions of uncertainty dynamically adjusted with changing task demands. **(A)** Example gaze trajectories. Panels show gaze trajectories from prototypical participant trials. High policy uncertainty was characterized by increased switching across the outcome boundary, whereas high EU trials were characterized by greater dispersion of eye movements. **(B)** Gaze-derived measures of uncertainty. Policy uncertainty and EU estimates were sensitive to valence and trial type. Participants showed greater EU on exploit policy trials and when partners were untrustworthy. In contrast, untrustworthy partners elicited greater policy uncertainty on adjust policy trials. **(C)** Expressed policy uncertainty and EU shift with changing task demands. Left panel shows mean estimates of policy uncertainty and EU dynamically oscillate depending on whether prior choice policies could be exploited or needed to be adjusted. Results for 1 example partner type are shown. Right panel shows mean estimates of each type of uncertainty aggregated across all partner types. **(D)** Model-free gaze signatures. Gaze traversed the outcome boundary (i.e., the net gain vs. net loss side of the prediction bar) more frequently on loss than gain trials, and on adjust vs. exploit trials. Sampling rate indicated that participants more expediently sampled the search space on exploit than adjust trials, suggesting a tendency to narrow one's predictions once the policy was optimized. **(E)** Policy uncertainty and EU were associated with model-free gaze signatures. *H*-gaze tracked the number of switches across the outcome boundary, whereas EU-gaze reliably captured the rate of sampling across multiple possible outcomes. $^{*}p < .05$, $^{**}p < .01$, and $^{***}p < .001$. Errors bars indicate the standard error of the mean.

Thus, simple, model-free gaze signatures reliably captured uncertainty about whether anticipated outcomes were worth the investment (policy uncertainty) (Figure 2E) but also how participants accumulated additional knowledge about specific outcomes when choice policies were already optimized (EU).

Anxiety Affected the Ability to Detect Policy Uncertainty

Individual variability in tolerance of uncertainty is likely to affect how humans attend to and expediently resolve uncertainty signals. Replicating our prior results (52), we observed that highly anxious participants invested significantly more money on loss blocks (i.e., when partners were untrustworthy) than the low-anxiety group (untrustworthy-start partner: valence × group interaction, t = 2.21, p = .027; neutral-start partner: valence × group interaction, t = 2.18, p = .030; trustworthy-start partner: valence × group interaction, t = 2.05, p = .040) (Figure 3A), resulting in greater monetary losses (valence × group interaction: t = −2.51, p = .012) (Figure 3B). We next tested whether investments with untrustworthy partners were accompanied by the belief partners were indeed trustworthy. This hypothesis diverges from social signaling accounts that would predict continued investment, despite having knowledge that partners will not reciprocate—a mechanism for restoring trust. Predicted outcomes derived from gaze measures indicated that participants with high anxiety also expressed more optimistic beliefs about partner returns, particularly when their partners were untrustworthy (valence × group interaction: t = 3.16, p = .002) (Figure 3C). Thus, our findings suggest that altered learning profiles did not emerge

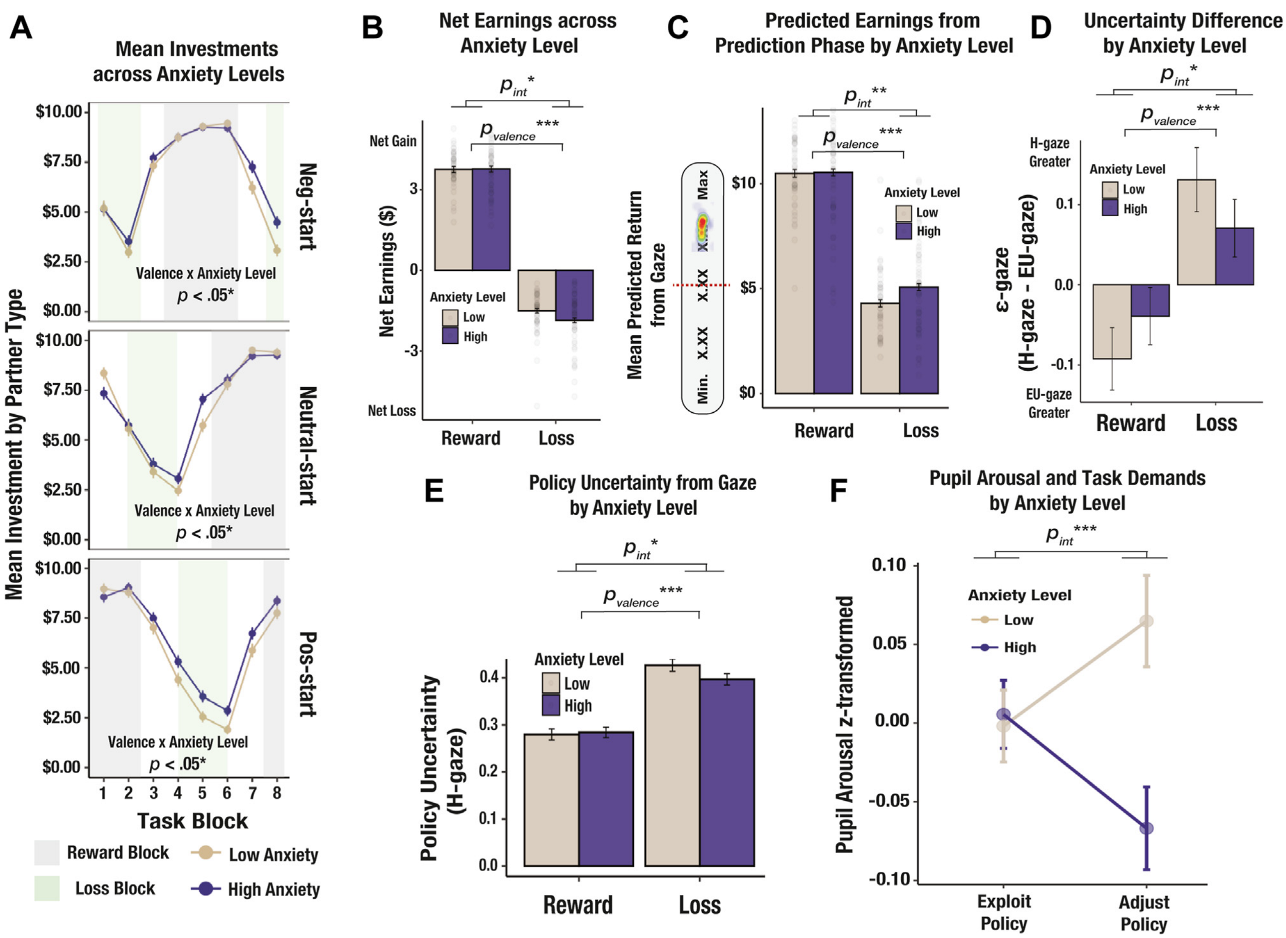

Figure 3. Impact of anxiety levels on expressions of uncertainty and pupil-based arousal. **(A)** Mean investment differences across anxiety groups. The high-anxiety group invested significantly more money with partners that were untrustworthy leading to greater monetary losses than the low-anxiety group. Asterisks correspond to valence × group interaction effect for each partner type. **(B)** Net earnings across anxiety level. Participants with high anxiety lost significantly more money when interacting with untrustworthy partners (loss trials) than the low-anxiety group. Groups did not differ on earnings when partners were trustworthy. *p_{int} denotes the interaction effect between anxiety level and outcome valence. **(C)** Predicted earnings from gaze by anxiety level. In addition to investing more money with untrustworthy partners, participants with high anxiety also expressed more optimistic beliefs about the amount of money that would be returned on loss trials. **(D)** Uncertainty difference by anxiety level. Participants with low anxiety expressed greater epistemic uncertainty with trustworthy partners and greater policy uncertainty with untrustworthy partners. In contrast, participants with high anxiety exhibited reduced valence-asymmetrical expressions of uncertainty. *p_{int} denotes the interaction effect between anxiety level and outcome valence where the effect of valence on ε-gaze only reaches statistical significance in the low-anxiety group. **(E)** Expressions of policy uncertainty from gaze. Participants with high anxiety expressed less policy uncertainty when interacting with untrustworthy partners compared with participants with low anxiety, indicating reduced sensitivity to untrustworthy partners. *p_{int} denotes the interaction effect between anxiety level and outcome valence. **(F)** Pupil arousal and task demands. The low-anxiety group showed increased arousal during adjust policy trials after partners changed their behavior, whereas the high-anxiety group showed suppressed arousal when choice policies needed to be adjusted. *$p < .05$, **$p < .01$, and ***$p < .001$. Errors bars indicate the standard error of the mean.

from a desire to restore trust, but rather from biased beliefs that partners would reciprocate.

We next evaluated whether anxiety affects how effectively individuals resolve uncertainty when learning about others. To formally capture the relative weighting between policy uncertainty and EU, we quantified the relative difference between gaze-derived policy uncertainty and EU estimates on each trial (ε-gaze = H-gaze − EU-gaze), yielding a signed variable that indicates whether policy or EU dominates at that moment ($+\varepsilon$-gaze: greater H-gaze; $-\varepsilon$-gaze: greater EU-gaze). Comparing ε-gaze differences between high- and low-anxiety groups, we found that the low-anxiety group was more unsure of the optimal choice policy when partners were untrustworthy and conversely expressed greater uncertainty about precise values (epistemic knowledge) with trustworthy partners (valence × group interaction: $t = -2.16$, $p = .031$) (Figure 3D). This indicates that participants with low anxiety differentially prioritized which type of uncertainty signal to resolve depending on their partner's trustworthiness. Although participants with high anxiety expressed a similar profile, valence-dependent expressions of uncertainty did not reach statistical significance.

Next, we tested whether anxiety levels affected the expression of distinct sources of uncertainty. Although both groups expressed greater policy uncertainty when interacting with untrustworthy partners (main effect of valence: $t = 12.13$, $p < .001$) (Figure 3E), participants with high anxiety expressed less policy uncertainty compared with the low-anxiety group (valence × group interaction: $t = -2.38$, $p = .017$). Groups did not vary in their expressions of EU. Collectively, these findings reveal key differences between high- and low-anxiety individuals: People with low anxiety exhibit distinct prioritization of uncertainty signals depending on their partner's trustworthiness, a process that was blunted in the high- anxiety group. Furthermore, participants with high anxiety showed reduced policy uncertainty about partners, suggesting more entrenched policies when interacting with untrustworthy individuals. The tendency toward fixed policies (and thus reduced policy uncertainty) was mirrored by more optimistic beliefs, indicating that reduced sensitivity to policy uncertainty in the high-anxiety group emerged in part from harder-to-undo beliefs once participants learned that a partner was trustworthy.

Individuals With Anxiety Showed Blunted Arousal to Policy Uncertainty

Two common threads emerged across our behavioral and gaze-based analyses. First, we found that highly anxious participants were slower to behaviorally adapt to increasingly untrustworthy partners. Second, we found that the relative expression of policy uncertainty was altered in individuals with anxiety, particularly in the loss domain. However, it remains unclear whether slower adaptation to untrustworthy people occurred because participants with anxiety failed to detect policy uncertainty writ large or because they maladaptively responded to policy uncertainty. With the latter, maladaptive responding could be a function of persistently enhanced physiological arousal. To test these competing hypotheses, we compared the relative magnitude of pupil-linked arousal during feedback across exploit versus adjust policy trials. As demonstrated in prior work, increased pupil-based arousal during feedback indexes the magnitude of surprise elicited from outcomes and the rate of learning adjustment (43). Thus, by linking distinct policy reliability periods with pupil-based arousal, we can identify the extent to which increased arousal served as a distinct physiological update signal in the high- and low-anxiety groups. Comparing arousal profiles, participants with low anxiety demonstrated the predicted pattern of increased arousal after partners' behaviors crossed the outcome boundary, indicating that the choice policy should be updated, whereas highly anxious participants demonstrated reduced arousal during this same time period (trial type × anxiety level interaction: $t = -3.13$, $p = .002$) (Figure 3F). These effects were further modulated by valence, such that the low-anxiety group exhibited increased arousal when engaging with increasingly untrustworthy partners (policy period × valence interaction: $t = 2.24$, $p = .025$), whereas the high-anxiety group generally reduced arousal when partners were untrustworthy (main effect of valence: $t = -2.52$, $p = .012$). Thus, the arousal pattern from our pupillometry measures suggests that reduced learning in the high-anxiety group arose, in part, from a failure to physiologically respond to unstable choice policies.

Computational Modeling Revealed Asymmetrical Reductions in Learning From Social Losses Versus Rewards in the High-Anxiety Group

Flexibly resolving policy uncertainty and EU guides successful learning in a dynamic environment—a process that is disrupted by anxiety. To examine the mechanistic link between these uncertainty signals and the rate of learning adjustment, we used a computational modeling approach. We tested and compared 3 Bayesian reinforcement learning models (see the Supplement). Our core model, dynamic Bayesian reinforcement learning (DBRL), was developed in our prior work (52) in which trial-level beliefs are adjusted through outcome history. As a Bayesian learner accumulates evidence that a partner is (un)trustworthy, it becomes more confident in that belief. Thus, when a partner changes their behavior, such a model will overly rely on the history of prior outcomes (56,57). To capture how the effect of prior outcomes on posterior beliefs should be adjusted in a nonstationary environment, our dynamic Bayesian model leverages changes in policy uncertainty to modulate decay (i.e., forgetting). By dynamically decaying prior beliefs as policy uncertainty increases, one can prioritize learning from more recent outcomes and quickly accumulate evidence, allowing new choice policies to form. Thus, rather than assuming a constant probability of change at a fixed rate, decay increases when the agent becomes more uncertain about what to do, thereby balancing the trade-off between stability and flexibility (6). In the model, policy uncertainty is calculated as the entropy, H, over the agent's choice probabilities, where p_1 and p_2 refer to the agent's probability of investing maximally ($10) or minimally ($1), respectively.

$$H_t = -[p_1 \times log_2(p_1) - p_2 \times log_2(p_2)] \tag{1}$$

Notably, the analogous formulation of H was used to calculate our gaze-derived estimate of policy uncertainty, H-gaze (see the Supplement). Gaze and model-derived H were positively correlated (t = 2.69, β = 0.036, p = .007), validating our assumption that H and H-gaze index uncertainty about one's current policy.

Consistent with our previous model for this task, decay was modeled separately for gains and losses (γ_{pos} and γ_{neg}, respectively). We further deconstructed γ into a constant γ_0 term (baseline beliefs about changeability) and a separate γ_1 term to allow decay to further increase or decrease as a function of the learner's change in policy uncertainty from trial to trial, quantified by ΔH. Note that, in our prior work, we constrained γ_1 to negative values, reflecting the assumption that as policy uncertainty increases, prior values are decayed. In our current model, DBRL-2, we relaxed this assumption and allowed changes in policy uncertainty to either decay prior reward history or exert the opposite effect of preserving prior knowledge (see the Supplement), allowing us to better capture the behavioral profile of participants with anxiety who show reduced learning when policy uncertainty increases.

$$\gamma_{pos} = \gamma_{0_{pos}} + \gamma_{1_{pos}} \times \Delta H \tag{2}$$

$$\gamma_{neg} = \gamma_{0_{neg}} + \gamma_{1_{neg}} \times \Delta H \tag{3}$$

We tested and compared a set of 4 nested Bayesian reinforcement learning models, 3 of which included additional γ_1 terms to dynamically adjust the decay rate (DBRL) (see the Supplement for details). Across both high- and low-anxiety groups, the DBRL-2 model best captured behavior (protected exceedance probability > 0.99) (Figure 4B; see the Supplement for model comparison details) and could reliably reproduce participant choices (see maximum likelihood estimation model simulation; Figure 4A). To validate that policy uncertainty estimates from the model (H) were generally

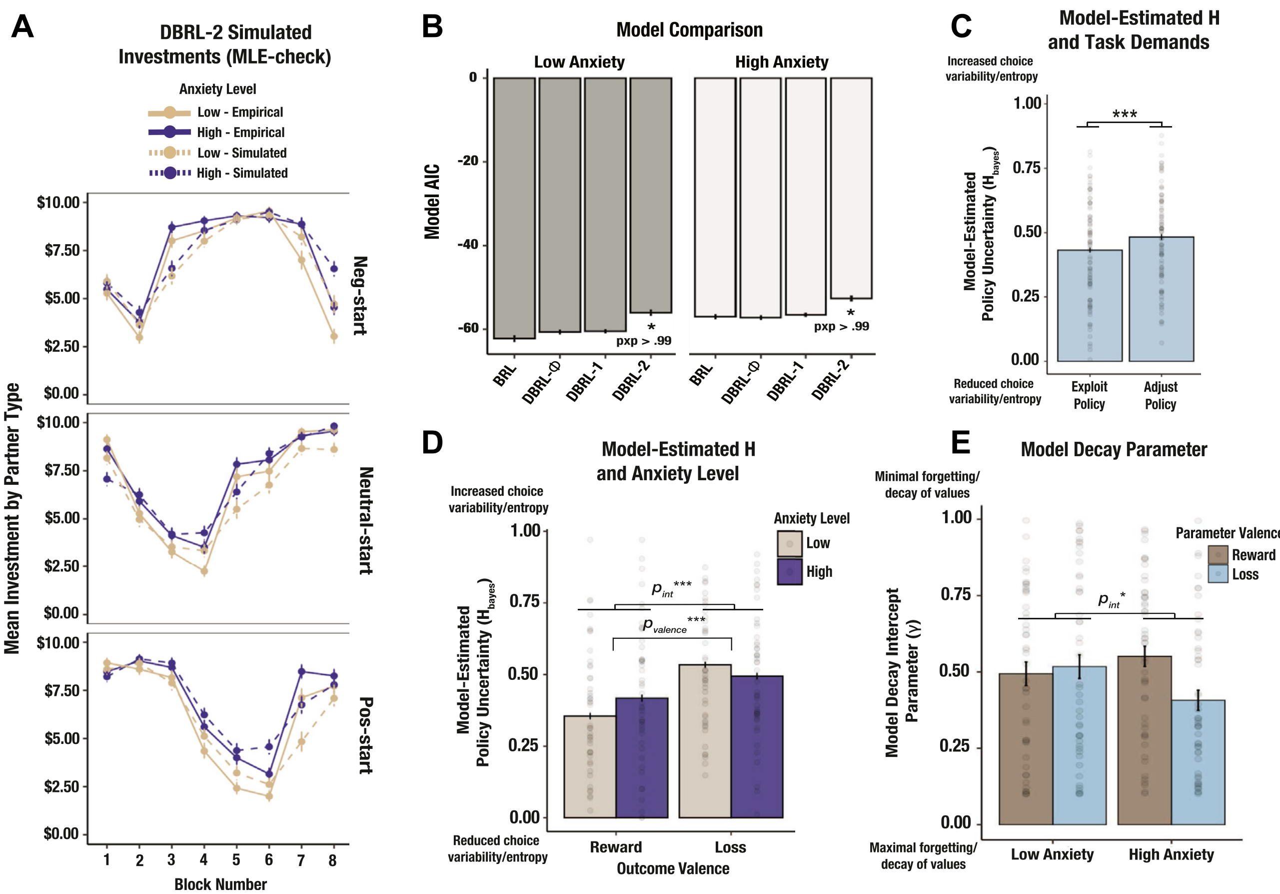

Figure 4. Dynamic Bayesian reinforcement learning (DBRL) model showed that individuals with anxiety were slower to adapt to untrustworthy partners. **(A)** DBRL-2 simulated investments from maximum likelihood estimation (MLE) check. Data were simulated using each participant's MLE-optimized (i.e., best-fitting) parameters. Model-simulated data recapitulated learning differences between groups. **(B)** Model comparison. Bayesian model comparison revealed that the best-fitting model to the data was adjusted DBRL-2. **(C)** Model-estimated H and task demands. Model estimates of policy uncertainty (H) increased when choice policies needed to be adjusted consistent with the pattern observed from our gaze-derived policy uncertainty estimates. **(D)** Model-estimated H and anxiety levels. Participants with high anxiety showed reduced, rather than increased, policy uncertainty (H) on loss trials, consistent with **(E)** the pattern we observed empirically. **(E)** Model decay parameter. The gamma intercept (γ_0) captured baseline beliefs about changeability. Highly anxious individuals exhibited a valence-specific learning asymmetry toward preserving previously learned rewarding values and having more uncertain beliefs (i.e., priors) when partners were untrustworthy, an asymmetry that produced a tendency to overinvest with untrustworthy partners. *$p < .05$, ***$p < .001$. Errors bars indicate the standard error of the mean. AIC, Akaike information criterion.

consistent with our gaze-based policy uncertainty measures (H-gaze), we examined the overall profile of trial-level H patterns from the model. Mirroring the physiological gaze-based analyses (Figure 2B), model-estimated H was greater during the adjust versus exploit policy trials when choice policies need to be revised ($t = 4.94$, $p < .001$) (Figure 4C). Furthermore, model-estimated H also recovered valence-dependent policy uncertainty differences across high- and low-anxiety groups (valence $\times$ anxiety group interaction: $t = 6.28$, $p < .001$) (Figure 4D). In particular, model-estimated policy uncertainty was lower for losses in the high-anxiety group, indicating less behavioral variability (i.e., perseveration) when partners were untrustworthy, recapitulating our gaze-based analysis (Figure 3E).

Comparing decay parameters from the winning DBRL-2 model, we observed a significant difference in baseline decay, γ_0, across groups in the loss domain (valence $\times$ anxiety group interaction: $t = -2.33$, $p = .0219$) (Figure 4E), revealing a learning asymmetry in the high-anxiety group for gains versus losses. In particular, decay parameters in the high-anxiety group revealed a tendency to disproportionately encode reward history while selectively forgetting the history of losses. This pattern of biased learning generated more precise and entrenched beliefs about a partner's trustworthiness, which would consequently produce a pattern of overinvesting.

DISCUSSION

Adaptively functioning in our social world requires integrating across multiple sources of uncertainty so we can expediently refine our beliefs and behaviors. We developed a novel eye-tracking procedure premised on information sampling theories, which granularly teased out distinct sources of uncertainty in real time, allowing us to evaluate how policy uncertainty and EU are differentially prioritized for learning. Our study reveals 2 key findings: people dynamically reoriented their attention toward each source of uncertainty as social interactions unfolded, and this attentional flexibility was critical for effectively adjusting one's behavior. In contrast, people with high anxiety showed reduced attentional switching between different sources of uncertainty and reduced expressions of policy uncertainty—a signal that is crucial for policy optimization.

Furthermore, while our findings dovetail with prior work showing that highly anxious people learn less from uncertain outcomes (42,50–52), we showed that anxiety altered learning through a biased information filter. In particular, by simultaneously measuring choice, predictions, and physiology, we found that individuals with anxiety disproportionately encoded reward histories, leading to more rigid, optimistic beliefs and inflexible choice policies, particularly in the loss domain. These findings rule out a social signaling explanation, instead suggesting that the locus of maladaptive behavior in our task was rigid beliefs and policies that become insensitive to feedback. Critically, our findings show altered social learning profiles emerging from a reward-encoding bias, which diverges from long-standing threat sensitivity accounts of generalized anxiety (58,59). While both perspectives converge on bias information processing and inflexible beliefs that ultimately result in maladaptive behavior, future work should investigate task environments and contexts that elicit altered reward versus threat processing. Along similar lines, a key element of our task design was evoking carefully controlled social rewards and threats through computerized agents, which were perceived as believable to varying degrees among participants (see the Supplement). Although posttask believability ratings did not significantly predict behavior, the use of social deception remains a fundamental limitation of the current design. Future work should aim to construct computerized agents that are indistinguishable from human partners or to eliminate the use of deception.

Our study sample was also characterized by a high prevalence of generalized anxiety symptoms. High and low anxiety levels were determined from scores on the GAD-7 scale and the STAI-Trait subscale. Within our final sample, approximately half of the participants ($n = 49$, ~52%) indicated symptoms above clinical significance thresholds. Although reported anxiety levels in our study are higher than those observed in a general population [estimated prevalence ~25% (60)], anxiety levels in the current study may be reflective of a unique combination of stressors. Recent high-profile reports have identified steeply increasing rates of mental illness among young adults and college-aged cohorts (61,62)—the primary demographic of the current study (ages 18–25, mean age: 20.4 years). One study (63) evaluating the prevalence of mental health disorders in undergraduates reported that most students were overwhelmed by their workload (~86%) and felt highly anxious day to day (~65%)—a pattern borne out across similar studies (64,65). Thus, the combined effects of an age-skewed sample and academic stress may explain higher reported anxiety levels in our study. Furthermore, we identified high anxiety levels using clinically recommended guidelines; however, it is worth noting that exceeding cutoffs on the GAD-7 or STAI-Trait scale is not necessarily diagnostic of pathological anxiety. This highlights the need for increased clinical translation work and methodological innovation to identify when self-reported symptoms and altered learning profiles are associated with maladaptive, real-world beliefs and behavior.

Finally, our eye-tracking procedure, which dissociated among different forms of uncertainty in real time, allowed us to evaluate whether altered learning might arise from individual variability in attending to distinct sources of uncertainty. While prior work speculates that reduced learning from losses in people with anxiety might arise from disrupted learning under uncertainty (42,50,51), here we explored the balance in learning from one specific type of uncertainty versus another. In particular, our findings leave open the possibility that the balance between prioritizing distinct uncertainty signals may be reconfigured in anxiety disorders such that one source of uncertainty can be overly prioritized in the system (e.g., epistemic knowledge), leaving other forms of uncertainty, such as policy uncertainty, unresolved. Future work should consider how the functional utility of adjusting one's attention toward epistemic knowledge or adjusting one's policy might be governed by prefrontal systems (i.e., a hypothetical meta-critic) (12), conveying the prioritization of uncertainty signals to solve a particular problem, and how biases in this system might alter learning. This account leaves open a new perspective in computational psychiatry approaches toward understanding anxiety-based disorders—one in which disrupted learning and decision making might reveal a divergent and heterogeneous set of goals and motivations of the learner.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was funded by the Centers of Biomedical Research Excellence (Grant No. P20GM103645 [to OFH]) and NARSAD Young Investigator Award from the Brain & Behavior Research Foundation (Grant No. 26210 [to OFH]).

AL, OFH, and MJF conceptualized the experiment. AL collected the data. AL performed data analysis under the supervision of OFH and MJF. AL, OFH, and MJF wrote the article.

We thank Matthew Nassar for helpful comments and feedback and Eric Ingram for assisting with data collection.

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Cognitive and Psychological Sciences, Brown University, Providence, Rhode Island (AL, MJF, OFH); Carney Institute of Brain Sciences, Brown University, Providence, Rhode Island (MJF, OFH); and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts (AL).

Address correspondence to Oriel FeldmanHall, Ph.D., at oriel.feldmanhall@brown.edu.

Received Mar 9, 2024; revised and accepted Jul 17, 2024.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsc.2024.07.015.

REFERENCES

1. Berg J, Dickhaut J, McCabe K (1995): Trust, reciprocity, and social history. Games Econ Behav 10:122–142.
2. Diaconescu AO, Mathys C, Weber LAE, Daunizeau J, Kasper L, Lomakina EI, *et al.* (2014): Inferring on the intentions of others by hierarchical Bayesian learning. PLOS Comput Biol 10:e1003810.
3. Fehr E, Fischbacher U, Kosfeld M (2005): Neuroeconomic foundations of trust and social preferences: Initial evidence. Am Econ Rev 95: 346–351.
4. FeldmanHall O, Nassar MR (2021): The computational challenge of social learning. Trends Cogn Sci 25:1045–1057.
5. Kramer RM, Wei J (2014): Social uncertainty and the problem of trust in social groups: The social self in doubt. Psychol Soc Self 145–168.
6. Franklin NT, Frank MJ (2015): A cholinergic feedback circuit to regulate striatal population uncertainty and optimize reinforcement learning. eLife 4:e12029.
7. Parr T, Friston KJ (2017): Uncertainty, epistemics and active inference. J R Soc Interface 14:20170376.
8. Payzan-LeNestour E, Bossaerts P (2011): Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. PLoS Comput Biol 7:e1001048.
9. Tomov MS, Truong VQ, Hundia RA, Gershman SJ (2020): Dissociable neural correlates of uncertainty underlie different exploration strategies. Nat Commun 11:2371.
10. Wilson RC, Bonawitz E, Costa VD, Ebitz RB (2021): Balancing exploration and exploitation with information and randomization. Curr Opin Behav Sci 38:49–56.
11. Friston KJ, Lin M, Frith CD, Pezzulo G, Hobson JA, Ondobaka S (2017): Active inference, curiosity and insight. Neural Comput 29:2633–2683.
12. Jaskir A, Frank MJ (2023): On the normative advantages of dopamine and striatal opponency for learning and choice. eLife 12:e85107.
13. Palminteri S, Khamassi M, Joffily M, Coricelli G (2015): Contextual modulation of value signals in reward and punishment learning. Nat Commun 6:8096.
14. Niv Y, Daw ND, Dayan P (2006): Choice values. Nat Neurosci 9:987–988.
15. Sutton RS, Barto AG (2018): Reinforcement Learning: An Introduction. Cambridge: MIT Press.
16. Bennett D, Niv Y, Langdon AJ (2021): Value-free reinforcement learning: Policy optimization as a minimal model of operant behavior. Curr Opin Behav Sci 41:114–121.
17. Collins AGE, Frank MJ (2014): Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol Rev 121:337–366.
18. Li J, Daw ND (2011): Signals in human striatum are appropriate for policy update rather than value prediction. J Neurosci 31:5504–5511.
19. Littman ML (1994): Memoryless Policies: Theoretical Limitations and Practical Results. Cambridge, MA: MIT Press, 238.
20. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, *et al.* (2015): Human-level control through deep reinforcement learning. Nature 518:529–533.
21. Wilson RC, Geana A, White JM, Ludvig EA, Cohen JD (2014): Humans use directed and random exploration to solve the explore–exploit dilemma. J Exp Psychol Gen 143:2074–2081.
22. Cohen JD, McClure SM, Yu AJ (2007): Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. Philos Trans R Soc B Biol Sci 362:933–942.
23. Schulz E, Gershman SJ (2019): The algorithmic architecture of exploration in the human brain. Curr Opin Neurobiol 55:7–14.
24. Vives M-L, FeldmanHall O (2018): Tolerance to ambiguous uncertainty predicts prosocial behavior. Nat Commun 9:2156.
25. Collins AGE, Shenhav A (2022): Advances in modeling learning and decision-making in neuroscience. Neuropsychopharmacology 47:104–118.
26. Niv Y, Langdon A (2016): Reinforcement learning with Marr. Curr Opin Behav Sci 11:67–73.
27. Nachum O, Norouzi M, Xu K, Schuurmans D (2017): Bridging the gap between value and policy based reinforcement learning. Adv Neural Inf Process Syst 30.
28. Gottlieb J, Hayhoe M, Hikosaka O, Rangel A (2014): Attention, reward, and information seeking. J Neurosci 34:15497–15504.
29. Fan H, Burke T, Sambrano DC, Dial E, Phelps EA, Gershman SJ (2023): Pupil size encodes uncertainty during exploration. J Cogn Neurosci 35:1508–1520.
30. Gottlieb J, Oudeyer P-Y, Lopes M, Baranes A (2013): Information-seeking, curiosity, and attention: Computational and neural mechanisms. Trends Cogn Sci 17:585–593.
31. Gottlieb J, Oudeyer P-Y (2018): Towards a neuroscience of active sampling and curiosity. Nat Rev Neurosci 19:758–770.
32. Bakst L, McGuire JT (2021): Eye movements reflect adaptive predictions and predictive precision. J Exp Psychol Gen 150:915–929.
33. Leckey S, Selmeczy D, Kazemi A, Johnson EG, Hembacher E, Ghetti S (2020): Response latencies and eye gaze provide insight on how toddlers gather evidence under uncertainty. Nat Hum Behav 4:928–936.
34. Harris D, Vine S, Wilson M, Arthur T (2024): The relationship between environmental statistics and predictive gaze behaviour during a manual interception task: Eye movements as active inference. Comput Brain Behav 7:225–241.
35. Wittek P, Liu Y-H, Darányi S, Gedeon T, Lim IS (2016): Risk and ambiguity in information seeking: Eye gaze patterns reveal contextual behavior in dealing with uncertainty. Front Psychol 7:1790.
36. Garland R (1991): The mid-point on a rating scale: Is it desirable. Mark Bull 2:66–70.
37. Subedi BP (2016): Using Likert type data in social science research: Confusion, issues and challenges. Int J Contemp Appl Sci 3:36–49.
38. Bechara A, Tranel D, Damasio H, Damasio AR (1996): Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. Cereb Cortex 6:215–225.
39. Critchley HD, Mathias CJ, Dolan RJ (2001): Neural activity in the human brain relating to uncertainty and arousal during anticipation. Neuron 29:537–545.
40. FeldmanHall O, Glimcher P, Baker AL, Phelps EA (2016): Emotion and decision-making under uncertainty: Physiological arousal predicts increased gambling during ambiguity but not risk. J Exp Psychol Gen 145:1255–1262.
41. Aston-Jones G, Cohen JD (2005): An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. Annu Rev Neurosci 28:403–450.
42. Browning M, Behrens TE, Jocham G, O'Reilly JX, Bishop SJ (2015): Anxious individuals have difficulty learning the causal statistics of aversive environments. Nat Neurosci 18:590–596.
43. Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI (2012): Rational regulation of learning dynamics by pupil-linked arousal systems. Nat Neurosci 15:1040–1046.

44. Urai AE, Braun A, Donner TH (2017): Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. Nat Commun 8:14637.
45. Bishop SJ (2007): Neurocognitive mechanisms of anxiety: An integrative account. Trends Cogn Sci 11:307–316.
46. Boelen PA, Reijntjes A (2009): Intolerance of uncertainty and social anxiety. J Anxiety Disord 23:130–135.
47. Buhr K, Dugas MJ (2009): The role of fear of anxiety and intolerance of uncertainty in worry: An experimental manipulation. Behav Res Ther 47:215–223.
48. Carleton RN, Mulvogue MK, Thibodeau MA, McCabe RE, Antony MM, Asmundson GJG (2012): Increasingly certain about uncertainty: Intolerance of uncertainty across anxiety and depression. J Anxiety Disord 26:468–479.
49. Koerner N, Dugas MJ (2006): A cognitive model of generalized anxiety disorder: The role of intolerance of uncertainty. In: Davey Graham CL, Wells Adrian, editors. Worry and its Psychological Disorders: Theory, Assessment and Treatment. Hoboken, NJ: Wiley, 201–216.
50. Aylward J, Valton V, Ahn W-Y, Bond RL, Dayan P, Roiser JP, Robinson OJ (2019): Altered learning under uncertainty in unmedicated mood and anxiety disorders. Nat Hum Behav 3:1116–1123.
51. Gagne C, Zika O, Dayan P, Bishop SJ (2020): Impaired adaptation of learning to contingency volatility in internalizing psychopathology. eLife 9:e61387.
52. Lamba A, Frank MJ, FeldmanHall O (2020): Anxiety impedes adaptive social learning under uncertainty. Psychol Sci 31:592–603.
53. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, Herzberg PY (2008): Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Med Care 46:266–274.
54. Spitzer RL, Kroenke K, Williams JBW, Löwe B (2006): A brief measure for assessing generalized anxiety disorder: The GAD-7. Arch Intern Med 166:1092–1097.
55. Spielberger CD (1983): State-Trait Anxiety Inventory for Adults. Available at: https://doi.org/10.1037/t06496-000. Accessed September 15, 2020.
56. Daw ND, Niv Y, Dayan P (2005): Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci 8:1704–1711.
57. Doll BB, Jacobs WJ, Sanfey AG, Frank MJ (2009): Instructional control of reinforcement learning: A behavioral and neurocomputational investigation. Brain Res 1299:74–94.
58. Davis M (1992): The role of the amygdala in fear and anxiety. Annu Rev Neurosci 15:353–375.
59. Shin LM, Liberzon I (2010): The neurocircuitry of fear, stress, and anxiety disorders. Neuropsychopharmacology 35:169–191.
60. Szuhany KL, Simon NM (2022): Anxiety disorders: A review. JAMA 328:2431–2445.
61. Goodwin RD, Weinberger AH, Kim JH, Wu M, Galea S (2020): Trends in anxiety among adults in the United States, 2008–2018: Rapid increases among young adults. J Psychiatr Res 130:441–446.
62. Blazer D (2020): Social isolation and loneliness in older adults-A mental health/public health challenge. JAMA Psychiatry 77:990–991.
63. Fried RR, Karmali S, Irwin JD (2022): Minding many minds: An assessment of mental health and resilience among undergraduate and graduate students; a mixed methods exploratory study. J Am Coll Health 70:898–910.
64. Kang HK, Rhodes C, Rivers E, Thornton CP, Rodney T (2021): Prevalence of mental health disorders among undergraduate university students in the United States: A review. J Psychosoc Nurs Ment Health Serv 59:17–24.
65. Sheldon E, Simmonds-Buckley M, Bone C, Mascarenhas T, Chan N, Wincott M, *et al.* (2021): Prevalence and risk factors for mental health problems in university undergraduate students: A systematic review with meta-analysis. J Affect Disord 287:282–292.