A low dimensional manifold of human exploration and exploitation behavior reveals behavioral states of apathy and anxiety

Xinyuan Yan [1], R. Becket Ebitz [3], Nicola Grissom [2], David P. Darrow [4], Alexander B. Herman [1]*

[1]Department of Psychiatry and Behavioral Sciences, University of Minnesota, Minneapolis, MN 55455, USA

[2]Department of Psychology, University of Minnesota, 75 E River Rd, Minneapolis, MN 55455, USA

[3]Department of Neuroscience, Universite de Montreal, 2900 Edouard Montpetit Blvd, Montreal, Quebec H3T 1J4, Canada

[4]Department of Neurosurgery, University of Minnesota, Minneapolis, MN 55455, USA

* Corresponding author: Alexander B. Herman

Email: herma686@umn.edu

Author Contributions: Conceptualization, A.B.H, R.B.E, N.G, D.P.D; Methodology, X.Y, R.B.E, A.B.H; Formal Analysis, X.Y; Writing – Original Draft, X.Y, D.P.D, A.B.H; Funding Acquisition, A.B.H and N.G; Supervision, A.B.H, D.P.D.

Competing Interest Statement: The authors declare no competing interests.

Classification: SOCIAL SCIENCES, BIOLOGICAL SCIENCES

Keywords: explore/exploit, manifold, apathy, anxiety, computational psychiatry

This PDF file includes:

Main Text
Figures 1 to 4
Tables 0

Abstract

The question of whether to explore for new rewards or to exploit one's knowledge to gain known rewards is a fundamental problem for all animals, including

humans. This balance is modulated by subjective states of anxiety and apathy, which, although they would seem to be incompatible with each other, often co-occur psychiatrically. Here, we report a latent structure underlying sequential exploration and exploitation strategies on a decision task, which reveals separate decision patterns aligned to anxiety and apathy. 1001 participants in a gender-balanced sample completed a three-armed restless bandit task along with psychiatric symptom surveys. Using advances in dimensionality reduction methods, we found that the decision sequences reduced to a low-dimensional manifold. The axes of this manifold explained individual differences in the balance between states of exploration and exploitation and the stability of those states, as determined by a statistical mechanics model of explore/exploit dynamics. Position along the balance axis was correlated with opposing symptoms of behavioral apathy and anxiety, while position along the stability axis was correlated with the level of emotional apathy. This work provides a basis for using behavioral manifolds to reveal relationships between behavioral dynamics and affective states, with important implications for behavioral measurement approaches to neuropsychiatric conditions.

Introduction

Humans make decisions based on their history of past success and their anticipation of future rewards. For instance, individuals may be likely to repeat a choice when rewarded but look for better options when not. However, in a changing and uncertain environment, simply exploiting when rewarded and exploring when not may be suboptimal (1). Furthermore, human beings exhibit a range of behaviors independent of optimality, owing to individual differences in the psychological determinants of decision-making (2, 3). Shifts in the balance of exploration and exploitation have been observed in several neuropsychiatric disorders (4, 5) such as schizophrenia (6), addictive behavior (7) and anxiety (8, 9).

As the number of possible options and reward history a person considers and the number of sequential decisions they make about those options increases, the potential complexity of their choice history increases exponentially, implying a "curse of dimensionality." Although computational models of decision-making theory postulate that these choices can be described by a low-dimensional representation (1, 10), the shape of that low-dimensional manifold and its relation to individual differences remains an open empirical question.

We employed advances in dimensionality reduction methods to test whether human explore-exploit decision-making occupies a low-dimensional shape and whether such a shape can explain individual differences in psychiatric measures important for decision-making. Dimensionality reduction on personality traits has demonstrated utility in identifying the latent structures of complex psychological processes. For example, factor analysis in personality psychology research has revealed patterns of covariation among latent constructs over hundreds of questionnaire items, such as those composing the big-five personality inventory (11, 12). These methods have helped us understand fundamental dimensions of human personalities (13, 14). In neurophysiological research, dimensionality reduction has also proved extremely powerful in describing the latent structure underlying the activity of thousands of neurons (15).

In the current study, 1001 participants completed a three-armed restless bandit explore-exploit task (Figure 1A) and psychiatric measurements of mood, motivation, and anxiety. Based on a growing literature, we hypothesized that these particular domains would play a key role in shaping the latent structure of explore-exploit decision-making (9, 16–18). Extreme exploratory behaviors may reflect the uncertainty aversion characteristic of anxiety, while over-exploitation may result from a decreased motivation to explore new possibilities. Our results show that exploration and exploitation behaviors indeed live on a low-dimensional manifold, and the shape of that manifold can, in part, be explained by the opposing forces of apathy and anxiety, which have been studied separately but rarely together.

Results

Gender-matched adults (n=1001, age: range 18-54, mean ± SD = 28.446 ± 10.354 years; gender, 493 female) performed a restless 3-armed bandit (Figure 1A) (19, 20), and completed symptom surveys on mood, anxiety, and motivation (See methods and SI Section1, Table S1). In the bandit task, participants were presented with images of three playing cards and moved their cursor over their choice. The probability of reward associated with each card deck drifted randomly over time, encouraging participants to balance periods of exploitation, when they've found a good option, with periods of exploration, when they search for a better one.

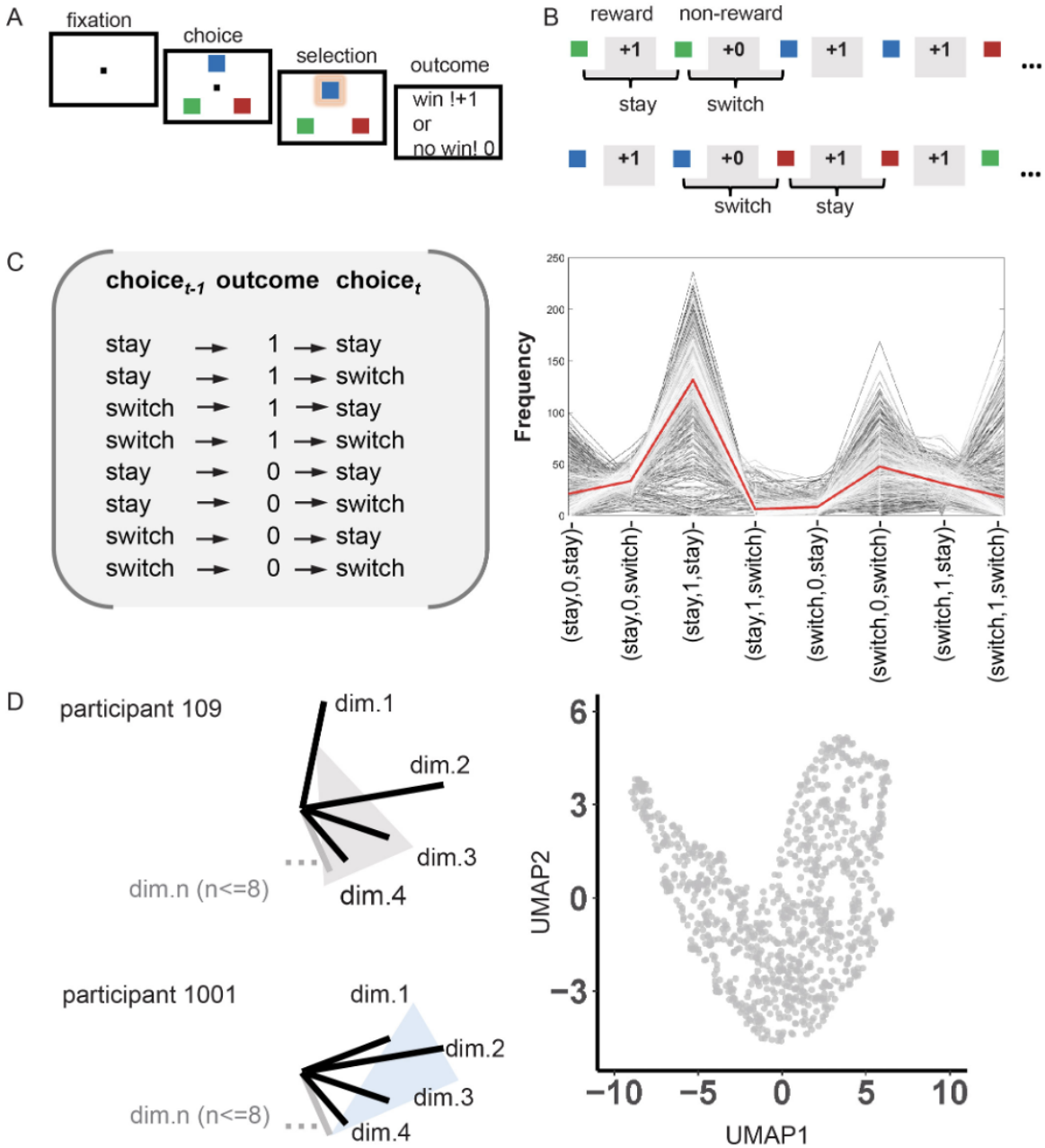

Figure 1. Dimensionality reduction of behavior on a three-armed restless bandit explore/exploit task (A) three-armed restless bandit task. Participants chose one option from among the three targets to receive reward or non-reward feedback. Each target is associated with a hidden reward probability that randomly and independently changed throughout the task. (B) example choice and reward sequence and the definition of *stay* and *switch*. Specifically, *stay* was defined as choosing the same target as in the previous trial, while *switch* was defined as choosing a different target. "+1" denotes reward feedback, "+0" denotes non-reward feedback. (C) all possible choice-reward sequences (left) and the distribution (right) of individual differences in frequency for all decision patterns. The red line represents the mean value for each decision pattern across participants. (D) schematic high dimensional space of participants' decision-making pattern (left) and the manifold-projected behavior (right).

Dimensionality reduction of behavior on a three-armed restless bandit explore/exploit task

To capture the full richness of behavior in the bandit task, we formatted each participant's trial-by-trial task data into sequences of choices to stay (repeat the choice on the last trial) or switch (choose a different option) and reward outcome for two consecutive trials ({$choice_{t-1}$, $outcome_{t-1}$, $choice_t$}, Figure 1B and 1C). The behavioral data for each participant was then transformed into counts for each of these eight unique sequences. To determine whether behavior on this high-dimensional task can be reduced to a low-dimensional representation, we applied Uniform Manifold Approximation and Projection (UMAP) (21), a computationally efficient algorithm that can preserve both the local and global distances between data points in high-dimensional space, to learn the two dimensional manifold underlying the eight-dimensional behavioral data, see Methods for more algorithm details). Including additional reward history did not affect the dimensionality reduction results (SI, Section 2, Figure S1).

A low dimensional manifold of individual differences in explore/exploit decision-making

We next asked how the manifold explains the underlying variation in explore/exploit decision-making in our sample. To reveal latent states of exploitation and exploration and the dynamics of transitions between those states, we fit a Hidden Markov Model (HMM) to the sequences of decisions, following previous work (20, 22). We mapped model-free (Figure 2A) and model-based (Figure 2B and 2C) behavioral measures onto the UMAP (all model-free and model-based measures, see SI Section 3, Table S2). We found that patterns of exploration and exploitation were indeed well-described by the UMAP projection, yielding clearly interpretable results. We then conducted correlation analyses between all indices and the axes of the manifold (i.e., UMAP1, the horizontal axis, and UMAP2, the vertical axis) to quantify the relationships.

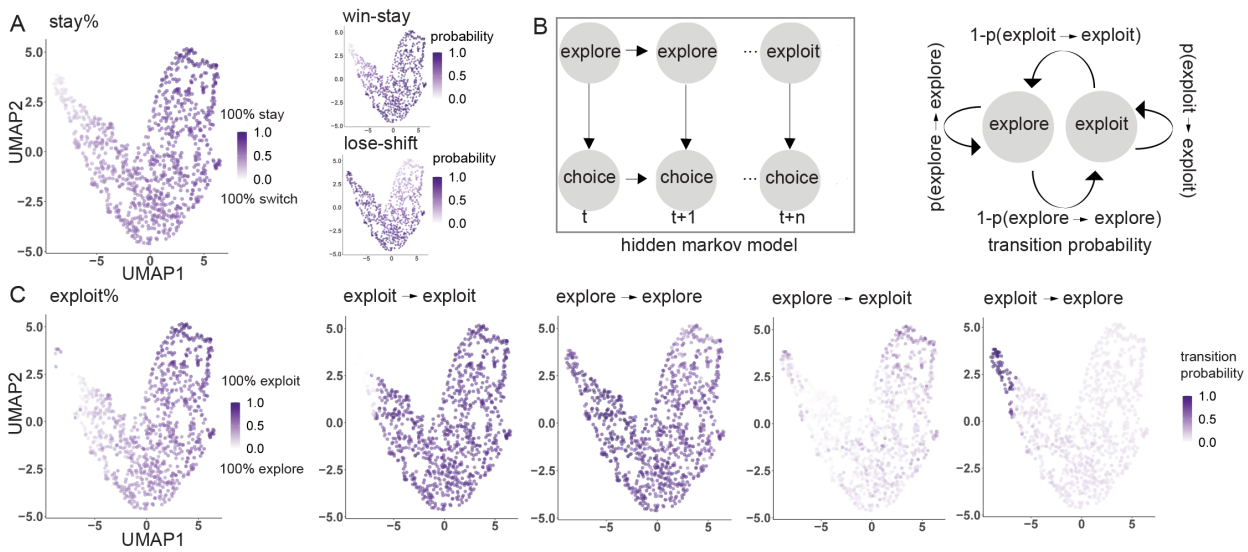

Figure 2. A low dimensional manifold of individual differences in explore/exploit decision-making (A) individual differences in model-free measures of sequential behavior projected onto the manifold (percentage of stay, win-stay, and lose-shift decisions). (B) schematic of Hidden Markov Model (HMM) (left) and transition probabilities within and between states (right). (C) individual differences in HMM model-based measures of decision states and transition probabilities projected onto the manifold (from left to right, percentage of exploit states; probability of transition from: exploit to exploit, explore to explore, explore to exploit, and exploit to explore).

The horizontal-axis dimension of the manifold explained most of the variance in both model-free stay-switch and model-based explore/exploit behavior. Moving from left to right along the horizontal axis (UMAP1) of the manifold, the probability of switching decreased, as did the probability of exploration, while the probability of staying and the probability of exploiting increased, with a high correlation coefficient (Figure 2A and 2C, SI Section 4, Table S3 & Table S4, all $p < 0.001$, corrected by FDR $p < 0.05$). When we decomposed the behavior into the degree with which individuals stay after a reward ("win-stay') and switch after reward omission ("lose-shift"), we found that the horizontal axis largely explained the win-stay behavior in the positive direction and the lose-shift behavior in the opposite direction.

The energy landscape explaining manifold dimensions

To understand how the dynamics of explore/exploit decision-making map onto the manifold, we examined the probabilities of the different state transitions (20, 22).

We found that the axes of the manifold explained different features of the explore/exploit transition dynamics. The horizontal axis (UMAP1) was aligned

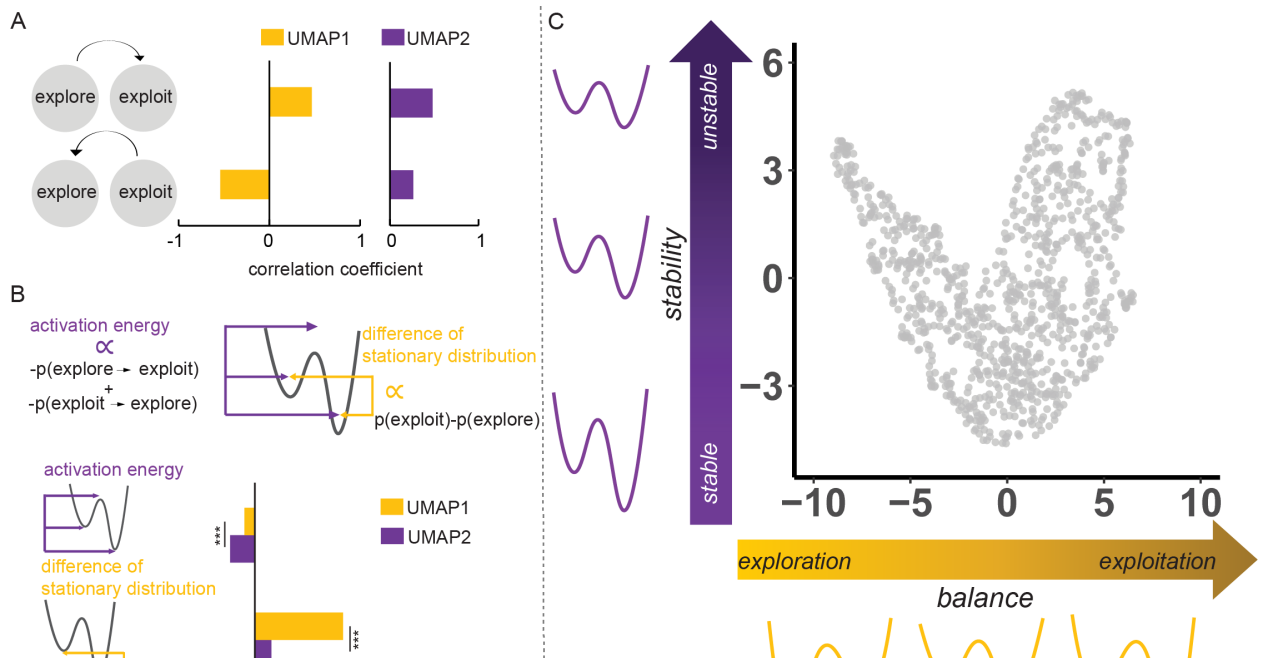

Figure 3. The energy landscape of explore/exploit decision-making explaining manifold dimensions (A) correlation coefficients between the state transitions and the manifold horizontal axis (UMAP1, yellow) and the vertical axis (UMAP2, purple), demonstrating that UMAP1 most closely tracked the explore/exploit balance, while UMAP2 was associated with state stability. (B) schematic (top) and correlation coefficients (bottom) showing that the UMAP1-balance association is explained by the difference in energy well depths (stationary distribution probabilities) of each state, while the UMAP2-stability association relates to the activation energy of state transitions. (C) schematic with manifold-projected data summarizing how the manifold is shaped by a horizontal dimension of explore/exploit state balance and a vertical dimension of state stability.

with the overall probability of exploitation. It was positively associated with the probability of transitioning from explore to exploit, positively associated with exploit to exploit, negatively associated with exploit to explore, and negatively associated with explore to explore (all $p < 0.001$, corrected by FDR $p < 0.05$; see Table S4). In contrast, the vertical axis (UMAP2) was negatively associated with the probability of remaining in either an exploit or an explore state and positively associated with transitioning between states (Figure 3A, all $p < 0.001$, corrected by FDR $p < 0.05$; see Table S4). We further analyzed the dynamics with the tools of thermodynamics (22). The stationary probability of occupying a given state is proportional to the depth of that state's potential energy well. The probability of transitioning between states, in turn, is related to the height of the activation energy barrier separating them. A higher transition probability indicates that less activation energy is needed to cross the barrier, leading to a more unstable system. To further confirm our explanations for UMAP1 and UMAP2, we did correlation

analyses between the UMAP score and the difference of the stationary distribution probabilities (i.e., the difference in energy that is associated with exploration and exploitation), as well as the activation energy, which captures the system's stability. Results showed that UMAP1 more strongly tracked the difference of potential energy ($r = 0.812$, $p<0.001$) for each state compared to UMAP2 (z-score = 21.971, $p<0.001$). While UMAP2 more strongly tracked the activation energy ($r = -0.228$, $p<0.001$) compared to UMAP1 (z-score = 3.125, $p = 0.001$) (Figure 3B). Thus, the horizontal axis (UMAP1) was most closely associated with the relative depths of the energy wells corresponding to explore and exploit states, which we call the *balance,* whereas the vertical axis was best characterized by the height of the activation energy barrier between them (UMAP2), which we refer to as the *stability* (Figure 3C).

Manifold projection of individual differences in anxiety and apathy subdomains

We next examined whether the latent manifold explained features of affect important for decision-making. We conducted Spearman correlation analyses to reveal associations between the UMAP scores and psychiatric symptoms.

Our results showed that both the horizontal axis (UMAP1) of the manifold (rho = 0.117, $p < 0.001$) and vertical axis (UMAP2) (rho = 0.101, $p = 0.001$) correlated positively with apathy, as well as social apathy (horizontal axis, rho = 0.117, $p < 0.001$; vertical axis, rho = 0.058, $p = 0.065$). Behavioral apathy correlated positively with the horizontal axis (rho = 0.082, $p = 0.009$) but not the vertical (rho = 0.061, $p = 0.055$), while emotional apathy correlated positively with the vertical axis (rho = 0.097, $p = 0.002$) but not horizontal axis (rho = 0.040, $p = 0.212$). Moreover, GAD7 (i.e., generalized anxiety) correlated negatively with the horizontal manifold (rho = -0.071, $p = 0.025$) but not the vertical one (rho = -0.061, $p = 0.054$). No significant results were found for SHPS or PHQ-9 (all $p > 0.06$). All significant results reported here survive FDR correction for multiple comparisons ($p < 0.05$).

We considered whether position on the UMAP axes could be explained by combinations of symptoms, which might yield insight into their underlying meaning. We constructed a linear regression model to map several significant psychiatric symptoms on UMAP1 (horizontal) and UMAP2 (vertical), respectively. A linear regression model composed of UMAP1 as the dependent variable and anxiety, social apathy, and behavioral apathy as independent variables showed a significant contribution of anxiety ($t = -3.419$, $p < 0.001$), as well as the two

subscales from Apathy (social apathy, $t = 2.854$, $p = 0.004$; behavioral apathy, $t = 2.156$, $p=0.031$). In contrast, a linear regression model with UMAP2 as the dependent variable and emotional apathy and social apathy as independent variables showed a significant contribution of emotional sensitivity ($t = 2.529$, $p =0.011$) and a marginally significant contribution from social motivation ($t = 1.784$, $p=0.074$) (Figure 4). We used variance inflation factor (VIF) to assess multicollinearity (23), and the results showed the VIF for both linear models is low ($VIF<5$).

Traversing the manifold from the upper left, the propensity to explore gradually decreases, behavioral and social apathy increase, and anxiety decreases. At the same time, the overall stability of states of both exploration and exploitation increases, and emotional apathy decreases. Moving through the trough of the manifold and up the right arm, while the probability of exploration continues to decrease, the decision states start to destabilize, and emotional apathy increases. Approaching the upper right, we find the highest level of exploitation, overall state instability, behavioral and social apathy, and the greatest emotional apathy.

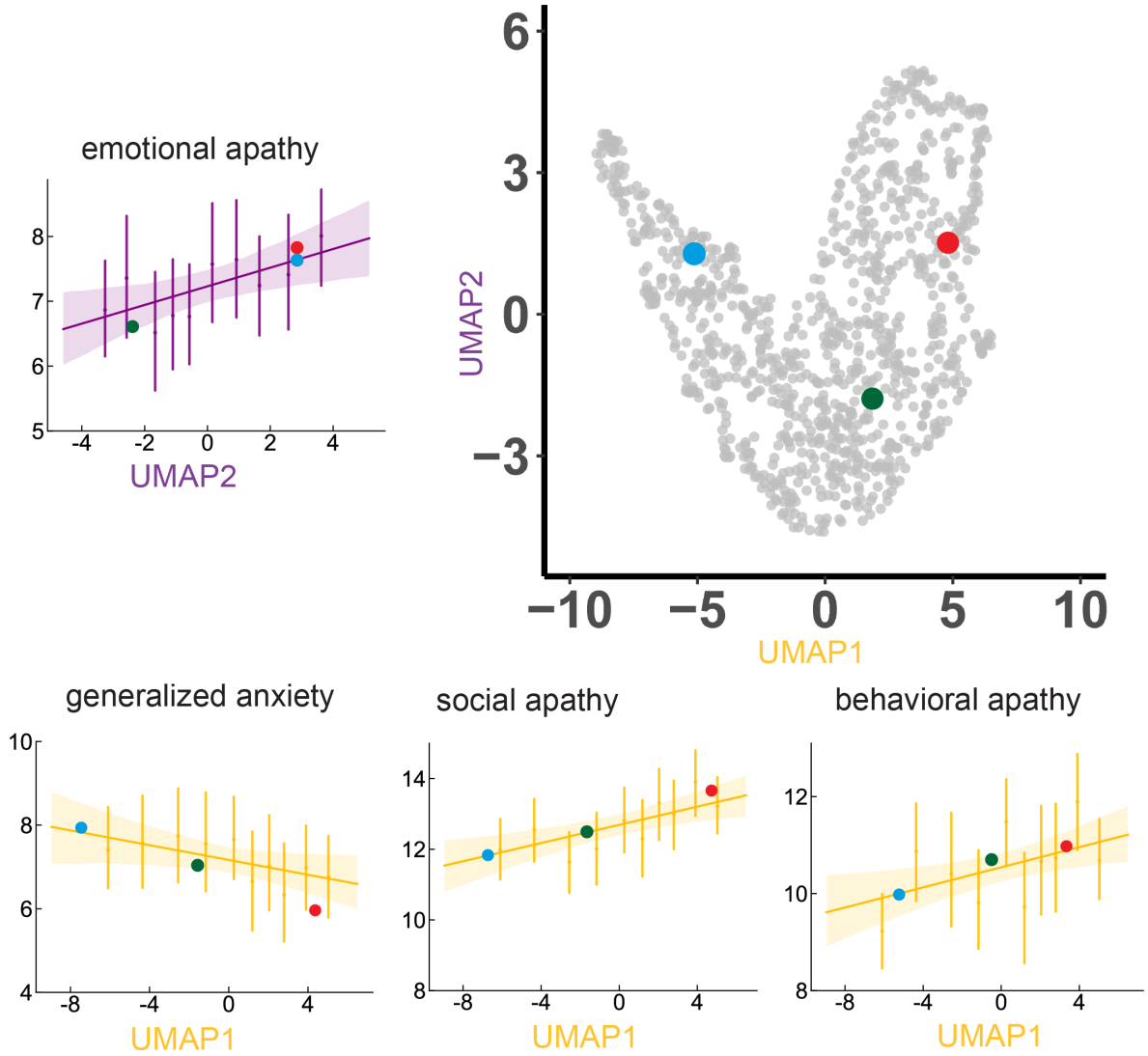

Figure 4. Manifold projection of individual differences in anxiety and apathy subdomains. Linear regression results demonstrate that levels of generalized anxiety, and social and behavioral apathy vary in opposite directions along the horizontal dimension UMAP1, while emotional apathy explains position along the vertical dimension UMAP2. The red, green, and blue dots are examples of different participants, showing how an individual's position on the manifold is associated with their symptoms.

Discussion

Sequential decision-making is often assumed to follow simple patterns, despite a wide range of possible complexity in individual differences. We examined the complexity of human exploration and exploitation strategies in a large-scale online experiment. Using dimensionality reduction and computational modeling, we found that the distribution of individual differences in this behavior collapsed to a low-dimensional manifold and that variation along the axes of the manifold explained individual differences in the dynamics of decision-making. Importantly, position on the manifold also predicted mental health symptoms that impact decision-making, revealing distinct roles for the transdiagnostic constructs of apathy and anxiety. These findings were consistent across linear and non-linear methods of dimensionality reduction (SI Section 5, Figure S2, Table S5&S6), robust to cross-validation (SI Section 6, Table S7&S8), and gender-independent (SI, Section 7 & Figure S3).

Anxiety and apathy are often viewed as affective expressions of two opposing factors in the motivational organization of emotion: aversive and appetitive, respectively (24). While we might predict, then, that apathy and anxiety would be anti-correlated, in our large sample of 1001 people, they were instead positively correlated ($r = 0.190$, $p<0.001$). However, when projected onto the behavioral manifold, we found that they did indeed oppose each other along the dimension controlling the exploration/exploitation balance. These results are consistent with significant prior work that anxious individuals are more exploratory (25), while apathetic individuals exhibit decision inertia and decreased exploration (26, 27). The counterbalancing of apathy and anxiety in exploration resembles the way that appetitive and aversive signals determine the balance between approach and avoidance. While approach and exploratory behavior may be responses to different kinds of environmental contexts and thus may provide differential advantages in those environments, they share common dopaminergic signaling pathways (28, 29), potentially explaining why they are related to similar affective states.

We found that the manifold axis corresponding to decision-state stability also explained individual differences in emotional apathy. Specifically, movement along this axis was associated with both a decrease in emotional *responsiveness* and a decrease in the persistence of decision-making states. One possible explanation is that the decreased reward sensitivity characteristic of emotional apathy also leads to a lowering of the activation energy barrier and the destabilization of decision-states. In line with this interpretation, a change in reward sensitivity is one of the most well-studied mechanisms by which emotions can affect decisions (30–32). While the decrease in both the win-stay and lose-shift behavior we

observed along this same axis supports this conclusion, we did not observe a direct correlation between emotional apathy and exploratory behavior, suggesting that further work would be required to substantiate it.

Our findings bear on the ongoing quest to characterize psychiatric pathology through behavior, particularly when that pathology affects the kinds of flexible decision-making captured by the dynamics of exploration and exploitation. While recent work has challenged the notion that symptoms and behavior correspond in meaningful ways (33), we found that symptoms and behavior can be complementary in explaining the variation in biological markers of psychiatric state, with symptoms shaping how exploratory behavior relates to physiology(19). Consistent with this result, the present findings demonstrate that because variation in exploratory behavior follows a parabolically shaped manifold, co-variation in symptoms related to exploratory behavior, in turn, becomes shaped by this manifold, leading, for instance, to the opposing directions of apathy and anxiety, despite their positive correlation. This finding resolves an apparent contradiction between the models of apathy and anxiety in the psychological literature and the clinical neuropsychiatric literature (concordant with clinical experience) documenting their high degree of comorbidity (34, 35).

In summary, our results suggest that a generalizable latent structure links the balance and stability of exploration and exploitation decision strategies to apathy and anxiety. Although we examined normative behavior in an online population, one key implication of our findings is that in disorders characterized by cognitive rigidity, such as Parkinson's Disease (36) or Schizophrenia (37), the relationship between behavior and affective symptoms may be similarly constrained. Locating an individual on the behavioral manifold might then help to predict how behavior would change with symptoms in response to treatment. For instance, the parabolic shape of the manifold with greater spread around its vertex suggests a potential "sweet spot" of balanced mild symptoms and flexible behavior that can be approached from different directions, depending on the starting point. To address this hypothesis, several important questions need to be answered: Do changes within an individual follow a trajectory on this same manifold? Do clinical populations project to the same manifold, or might they, in fact, be off manifold, projecting to the large regions of unoccupied space? The answers may help us to implement dimensional approaches to diagnosis and individualized neuropsychiatric care (38).

Materials and Methods

Ethics approval

The experimental procedures of all experiments were in line with the standards set by the Declaration of Helsinki and were approved by the local Research Ethics Committee of the University of Minnesota, Twin Cities. Participants provided written informed consent after the experimental procedure had been fully explained and were reminded of their right to withdraw at any time during the study.

Participants

We recruited a sample of 1512 participants via Amazon Mechanical Turk (MTurk) as well as Prolific (Prolific.co), exclusion criteria included current or history of neurological, and psychiatric disorders. And 1001 participants completed all questionnaires and the bandit task (age: range 18-54, mean ± SD = 28.446 ± 10.354 years; gender, 493 female). All participants were compensated for their time in line with minimum wage.

Questionnaire measurement

Participants' depressive symptoms, anhedonia, general anxiety, and apathy symptoms were measured using the Patient Health Questionnaire (PHQ-9)(39), the Snaith-Hamilton Pleasure Scale (SHPS) (40), the General Anxiety Disorder Screener (GAD-7) (41), and the Apathy-Motivation Index (42), respectively. More specifically, the PHQ-9, a 9-item multiple-choice inventory, was employed to measure depressive symptoms. The SHPS, a 4-point Likert scale, measured symptoms of anhedonia. Participants' general anxiety was measured using the GAD-7, which contains 7 items for assessing anxiety severity in the last two weeks. All items were rated on a 4-point scale, with higher scores indicating greater anxiety. Participants' apathy level was measured using the 18-item Apathy-Motivation Index (AMI), which was designed to identify and measure the general apathy, as well as subtypes of apathy in behavioral, social, and emotional domains. Higher scores on AMI represent greater apathy. For all questionnaire scores, see SI, Section 1 & Table S1.

Three-armed restless bandit task

We assessed explore/exploit behavioral dynamics using a 300-trial three-armed restless bandit task (20). Participants were free to choose between three targets for the potential to earn a reward of 1 point. Each target is associated with a hidden reward probability that randomly and independently changes throughout the task.

We seeded each participant's reward probability walks randomly to prevent biases due to particular kinds of environments. We assessed performance by comparing the total number of rewarded trials to that expected by chance. Out of the 1001 participants, 985 accrued more rewarded trials than would be expected by chance.

Dimensionality reduction method

Popular and valid dimensionality reduction techniques to reveal manifolds include t-distributed stochastic neighborhood embedding (t-SNE) (43), uniform manifold approximation and projection (UMAP)(21), and Principal component analysis (PCA) (44). However, t-SNE suffers from limitations, including slow computation time, and loss of global data structure, and it is not a deterministic algorithm (45). The main drawback of PCA is that it is highly affected by outliers in the dataset (44). In contrast, UMAP is a deterministic and efficient algorithm, it also preserves both local and global structure of original high-dimensional data. Uniform Manifold Approximation and Projection (UMAP)

UMAP was implemented in the R language. The eight-dimensional datasets from all participants were passed into the R package *umap*, version 0.2.8.0 (21), available at https://cran.r-project.org/web/packages/umap/) with default parameter setting as n_component = 2, n_neighbors = 15, min_dist = 0.3, metric = 'Euclidean'. For reproducibility reasons, we fixed the random_state in this algorithm. The hyperparameter *n_neighbors* decide the radius of the search region. Larger values will include more neighbors, thus forcing this algorithm to consider more global structure of original n-dimension data. Another important hyper-parameter, *min_dist* determines the allowed minimum distance apart for cases in lower-dimensional space. *metric* defines the way that UMAP is used to measure distances along the manifold.

Model-free analyses

We adopted some widely used model-free measures, including win-stay and lose-shift (46, 47) as the direct measurement for this learning task.

Win-stay. Win-stay is defined as the percentage of times that the choice in trial *t*-1 was repeated on trial *t* following a reward.

Lose-shift. In contrast, lose-shift equals the percentage of trials that the choice was shifted or changed when the outcome of trial t-1 was non-reward.

Computational Modeling

Hidden Markov Model

We fit a Hidden Markov Model (HMM) to the behavior, following previous studies(20) to decode the hidden state of each trial for each participant. Fundamentally, the HMM has two layers, the hidden layer (i.e., state) and observable layer. The hidden dimension should satisfy the Markov property. That is, the current hidden state only depends on the previous state but not any past model history. The observable dimension entirely depends on the current hidden states and is independent with other observations. Parameters of Hidden Markov Model can be represented as $\Omega \rightarrow (T, O, c)$. Specifically, T is the transition probabilities matrix, O is the observation probabilities matrix, or emissions matrix, where the c refers to a vector with initial probabilities for each hidden state. Here, we have two hidden states, "exploration" state and the "exploitation" state.

The transition probability between exploration and exploitation states can be represented by:

$$T = \begin{pmatrix} t_{11} & \cdots & t_{1m} & \vdots & \ddots & \vdots & t_{m1} & \cdots & t_{mm} \end{pmatrix}$$

$$t_{explore,\, exploit} = P(q_k = s_{exploit} | q_{k-1} = s_{explore})$$

(equation1)

Where $t_{explore,\, exploit}$ refers to the transition probability from hidden state $s_{exploit}$ to another hidden state $s_{explore}$

k = time instant, m = state sequence length

Then matrix O represent the transition probabilities between hidden and observable states.

$$O = \begin{pmatrix} o_{11} & \cdots & o_{1n} & \vdots & \ddots & \vdots & o_{m1} & \cdots & o_{mn} \end{pmatrix}$$

$$o_{explore,\, exploit} = P(\lambda_k = r_{exploit} | q_{k-1} = s_{explore})$$

(equation2)

k = time instant, n = observation sequence length

Both matrix O and T satisfy the principle that the sum along the rows must be equal to 1.
c is a m dimensional row vector which refers to the initial probability distribution. In our current study, the initial probability was fixed and equal to $\frac{1}{N}$, N equals the available choices.
We fit HMM via expectation-maximization using the Baum Welch algorithm and decode hidden state from observed choice sequences by the Viterbi algorithm (48).

Analyzing model dynamics.

When we successfully decode the state sequence from HMM, the next question would be, what is the long-term probability that the system will be in either state of exploration or exploitation? To answer this question, we further extracted the steady state analysis from the hidden Markov process from the transition matrix. Let π⇒ steady-state probabilities, mathematically,

$$T * \pi = \pi$$

$$T * \vec{v} = \theta \vec{v}$$

(equation3)

Where $\vec{v}$ is the eigenvector and θ is the eigenvalue, T is the transition matrix again. The steady-state vector is the eigenvector with eigenvalue 1.

To develop an intuition for how psychiatric symptoms covary with the dynamics of explore/exploit decision-making, we describe the energetic landscape of behavior as previous work did(22). Each state has an energy associated with it. Like a valley between two high peaks, a low-energy state is very stable and deep, while a high-energy condition is less stable. The steady-state probability is related to the energy of that state via Boltzmann distribution:

$$p_i = \frac{1}{Z} e^{\frac{-E_i}{k_B T}} \text{ (equation 4)}$$

where Z is the partition function of the system, k_B is the Boltzmann constant, and T is the temperature of the system (a constant). To characterize the difference in energy associated with *exploration* and *exploitation*, we have:

$$\frac{p_i}{p_j} = e^{\frac{-(E_i - E_j)}{k_B T}} \text{ (equation 5)}$$

after rearranging, we will have:

$$ln(\frac{p_i}{p_j})k_BT = E_j - E_i \text{ (equation 6)}$$

Next, we are going to quantify the activation energy required to transition between exploration and exploitation. We adopted the Arrhenius equation from chemical kinetics that relates the transitional probabilities between different states to the energy required to alert these transitions. The activation energy required to escape state (E_a):

$$k = Ae^{\frac{E_a}{k_BT}} \text{ (equation 7)}$$

where the k is the transition probability from one state to another state (e.g., from exploitation to exploration), E_a represents the activation energy required to escape that state (e.g., exploitation)

where A is a constant and we set this to one for convenience.

after rearranging, we will have:

$$E_a = -ln(\frac{k}{A})k_BT \text{ (equation 8)}$$

Acknowledgements

This work was supported by NIMH under award #R21MH127607, NIDA under award #K23DA050909, and the University of Minnesota's MnDRIVE (Minnesota's Discovery, Research and Innovation Economy) initiative.

We thank David Redish, Cathy Chen, Seth D Koenig, and Brian Sweiss for helpful comments on the manuscript.

Code and data availability.

The data, and source code used to do data analysis, as well as generate main and supplementary figures, is published at https://github.com/hermandarrowlab/umap-bandit

References

1. R. S. Sutton, A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd Ed. (Bradford Books, 2018).

2. C. Summerfield, K. Tsetsos, Do humans make good decisions? *Trends Cogn. Sci.* **19**, 27–34 (2015).

3. J. Hewig, *et al.*, Why humans deviate from rational choice. *Psychophysiology* **48**, 507–514 (2011).

4. M. A. Addicott, J. M. Pearson, M. M. Sweitzer, D. L. Barack, M. L. Platt, A Primer on Foraging and the Explore/Exploit Trade-Off for Psychiatry Research. *Neuropsychopharmacology* **42**, 1931–1939 (2017).

5. M. Dubois, T. U. Hauser, Value-free random exploration is linked to impulsivity. *Nat. Commun.* **13**, 4542 (2022).

6. G. P. Strauss, *et al.*, Deficits in Positive Reinforcement Learning and Uncertainty-Driven Exploration Are Associated with Distinct Aspects of Negative Symptoms in Schizophrenia. *Biol. Psychiatry* **69**, 424–431 (03/2011).

7. M. A. Addicott, J. M. Pearson, J. Wilson, M. L. Platt, F. J. McClernon, Smoking and the bandit: A preliminary study of smoker and nonsmoker differences in exploratory behavior measured with a multiarmed bandit task. *Exp. Clin. Psychopharmacol.* **21**, 66–73 (2013).

8. C. Gagne, O. Zika, P. Dayan, S. J. Bishop, Impaired adaptation of learning to contingency volatility in internalizing psychopathology. *Elife* **9** (2020).

9. H. Fan, S. J. Gershman, E. A. Phelps, Trait somatic anxiety is associated with reduced directed exploration and underestimation of uncertainty. *Nat Hum Behav*, 1–12 (2022).

10. Z. C. Ashwood, *et al.*, Mice alternate between discrete strategies during perceptual decision-making. *Nat. Neurosci.* **25**, 201–212 (2022).

11. M. Gurven, C. von Rueden, M. Massenkoff, H. Kaplan, M. Lero Vie, How universal is the Big Five? Testing the five-factor model of personality variation among forager-farmers in the Bolivian Amazon. *J. Pers. Soc. Psychol.* **104**, 354–370 (2013).

12. T. J. Bouchard Jr, J. C. Loehlin, Genes, evolution, and personality. *Behav. Genet.* **31**, 243–273 (2001).

13. D. I. Tamir, M. A. Thornton, J. M. Contreras, J. P. Mitchell, Neural evidence that three dimensions organize mental state representation: Rationality, social impact, and valence. *Proceedings of the National Academy of Sciences* **113**, 194–199 (2016).

14. S. T. Fiske, A. J. C. Cuddy, P. Glick, J. Xu, A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *J. Pers. Soc. Psychol.* **82**, 878–902 (2002).

15. S. Saxena, J. P. Cunningham, Towards the neural population doctrine. *Curr. Opin. Neurobiol.* **55**, 103–111 (2019).

16. R. Smith, *et al.*, Lower Levels of Directed Exploration and Reflective Thinking Are Associated With Greater Anxiety and Depression. *Front. Psychiatry* **12** (2022).

17. R. Nusslock, L. B. Alloy, Reward processing and mood-related symptoms: An RDoC and translational neuroscience perspective. *J. Affect. Disord.* **216**, 3–16 (2017).

18. M. Kirschner, A. Rabinowitz, N. Singer, A. Dagher, From apathy to addiction: Insights from neurology and psychiatry. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **101**, 109926 (2020).

19. E. A. Kaske, *et al.*, Prolonged Physiological Stress Is Associated With a Lower Rate of Exploratory Learning That Is Compounded by Depression. *Biol Psychiatry Cogn Neurosci Neuroimaging* (2022) https:/doi.org/10.1016/j.bpsc.2022.12.004.

20. R. B. Ebitz, E. Albarran, T. Moore, Exploration Disrupts Choice-Predictive Signals and Alters Dynamics in Prefrontal Cortex. *Neuron* **97**, 450-461.e9 (01/2018).

21. L. McInnes, J. Healy, N. Saul, L. Großberger, UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software* **3**, 861 (2018).

22. C. S. Chen, E. Knep, A. Han, R. B. Ebitz, N. M. Grissom, Sex differences in learning from exploration. *Elife* **10**, e69748 (2021).

23. E. R. Ziegel, J. Neter, M. Kutner, C. Nachtsheim, W. Wasserman, Applied linear statistical models. *Technometrics* **39**, 342 (1997).

24. P. J. Lang, M. M. Bradley, B. N. Cuthbert, Emotion, motivation, and anxiety: brain mechanisms and psychophysiology. *Biol. Psychiatry* **44**, 1248–1263 (1998).

25. K. C. Aberg, I. Toren, R. Paz, A neural and behavioral trade-off between value and uncertainty underlies exploratory decisions in normative anxiety. *Mol. Psychiatry* **27**, 1573–1587 (03/2022).

26. J. Scholl, H. A. Trier, M. F. S. Rushworth, N. Kolling, The effect of apathy and compulsivity on planning and stopping in sequential decision-making. *PLoS Biol.* **20**, e3001566 (2022).

27. B. Batrancourt, *et al.*, Exploration Deficits Under Ecological Conditions as a Marker of Apathy in Frontotemporal Dementia. *Front. Neurol.* **10**, 941 (2019).

28. C. G. Deyoung, The neuromodulator of exploration: A unifying theory of the role of dopamine in personality. *Front. Hum. Neurosci.* **7**, 762 (2013).

29. G. F. Antunes, *et al.*, Dopamine modulates individual differences in avoidance behavior: A pharmacological, immunohistochemical, neurochemical and volumetric investigation. *Neurobiol Stress* **12**, 100219 (2020).

30. N. Watanabe, M. Sakagami, M. Haruno, Reward prediction error signal enhanced by striatum-amygdala interaction explains the acceleration of probabilistic reward learning by emotion. *J. Neurosci.* **33**, 4487–4493 (2013).

31. H. van Steenbergen, G. P. H. Band, B. Hommel, In the mood for adaptation: how affect regulates conflict-driven control. *Psychol. Sci.* **21**, 1629–1634 (2010).

32. G. Dreisbach, T. Goschke, How positive affect modulates cognitive control: reduced perseveration at the cost of increased distractibility. *J. Exp. Psychol. Learn. Mem. Cogn.* **30**, 343–353 (2004).

33. A. Z. Enkavi, R. A. Poldrack, Implications of the Lacking Relationship Between Cognitive Task and Self-report Measures for Psychiatry. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **6**, 670–672 (07/2021).

34. B. M. Scott, *et al.*, Co-occurrence of apathy and impulse control disorders in Parkinson disease. *Neurology* **95**, e2769–e2780 (2020).

35. S. Nguyen, B. Krause-Sorio, P. Siddart, L. Ercoli, H. Lavretsky, Regional grey matter volume correlates with anxiety, apathy and resilience in geriatric depression. *Am. J. Geriatr. Psychiatry* **31**, S115 (2023).

36. R. Cools, R. A. Barker, B. J. Sahakian, T. W. Robbins, Mechanisms of cognitive set flexibility in Parkinson's disease. *Brain* **124**, 2503–2512 (2001).

37. S. Mohamed, J. S. Paulsen, D. O'Leary, S. Arndt, N. Andreasen, Generalized cognitive deficits in schizophrenia: a study of first-episode patients. *Arch. Gen. Psychiatry* **56**, 749–754 (1999).

38. B. N. Cuthbert, T. R. Insel, Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med.* **11**, 126 (2013).

39. K. Kroenke, R. L. Spitzer, J. B. W. Williams, The PHQ-9. *J. Gen. Intern. Med.* **16**, 606–613 (2001).

40. R. P. Snaith, *et al.*, A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. *Br. J. Psychiatry* **167**, 99–103 (1995).

41. B. Löwe, *et al.*, Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Med. Care* **46**, 266–274 (2008).

42. Y.-S. Ang, P. Lockwood, M. A. J. Apps, K. Muhammed, M. Husain, Distinct Subtypes of Apathy Revealed by the Apathy Motivation Index. *PLoS One* **12**, e0169938 (2017).

43. G. E. Hinton, S. Roweis, Stochastic neighbor embedding. *Adv. Neural Inf. Process. Syst.* **15** (2002).

44. I. T. Jolliffe, J. Cadima, Principal component analysis: a review and recent developments. *Philos. Trans. A Math. Phys. Eng. Sci.* **374**, 20150202 (2016).

45. H. Rhys, *Machine Learning with R, the tidyverse, and mlr* (Manning publications, 2020).

46. H. E. M. den Ouden, *et al.*, Dissociable effects of dopamine and serotonin on reversal learning. *Neuron* **80**, 1090–1100 (2013).

47. A. Bari, *et al.*, Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. *Neuropsychopharmacology* **35**, 1290–1301 (2010).

48. J. P. Coelho, T. M. Pinho, J. Boaventura-Cunha, *Hidden Markov Models: Theory and Implementation using Matlab* ®, 1st Ed. (CRC Press, 2019) (October 1, 2022).