Unveiling Cognitive Processes in Mice through Reinforcement Learning Models in the FST and TST

Zhihan Li[1], Tianyu Lu[1], Jiaozhao Yan[2], Xiang Zhang[1] & Yun-Feng Li[1,3]

[1]Beijing Institute of Basic Medical Sciences, Beijing, China
[2]School of Life Science and Technology, ShanghaiTech University, Shanghai, China
[3]Beijing Institute of Pharmacology and Toxicology, Beijing Key Laboratory of Neuropsychopharmacology, Beijing, China

Abstract

The forced swim test (FST) and tail suspension test (TST) are widely used to assess depression-like and stress-related behaviors in rodents. However, traditional approaches for analyzing these tests often oversimplify behavioral data, limiting insights into underlying cognitive processes. In this study, we employed reinforcement learning models to analyze behavioral trajectories, providing new perspectives on the cognitive functions of mice. Our findings suggest that behavior in the FST and TST can be conceptualized as reinforcement learning processes involving both learning and decision-making. Notably, we identified differences in the cognitive mechanisms driving behavior in the two tests, challenging the assumption that they are analogous and interchangeable for cross-validation. Using chronic restraint stress (CRS) exposure and medial prefrontal cortex (mPFC) pyramidal neuron inhibition as established paradigms, we demonstrate that these manipulations influence cognitive processes similarly in FST. Regression analyses of model parameters reveal distinct phases of behavior: early stages are dominated by learning rate-related components, while outcome sensitivity- and vigor-related components become more influential later. This dynamic suggests that traditional metrics, which often emphasize behavior during the final minutes of testing, may underestimate the significance of learning processes and overemphasize outcome- and vigor- driven responses. These findings underscore the value of computational approaches in uncovering the complex cognitive mechanisms underlying animal behavior.

Introduction

Recent advances in rodent research have significantly advanced our understanding of the neural circuits and molecular mechanisms underlying mental disorders, establishing itself as a cornerstone of neuroscience and pharmacology[1–5]. However, the inability of rodents to verbally communicate subjective experiences or engage in complex cognitive tasks limits their ability to fully capture the cognitive distortions and information-processing abnormalities seen in conditions like depression[6–8]. To overcome this limitation, simplified behavioral models such as the forced swim test (FST) and tail suspension test (TST) have been widely adopted[9–13]. As shown in Fig. 1C and D, research employing these methods has surged over the past two decades, and these tests remain some of the most popular behavioral assays today. These tests are typically interpreted by measuring immobility duration, which is commonly seen

Correspondence and requests for materials should be addressed to Z.L. (email: Li_Zhihan@outlook.com)

as an indicator of depression-like behavior. However, this narrow focus on immobility overlooks the cognitive mechanisms that underlie these actions, raising concerns about the true nature of what these tests measure. One leading explanation, the learning and memory hypothesis, suggests that the gradual increase in immobility is part of an adaptive learning process[9,14–19], Despite its prominence, empirical evidence supporting this hypothesis remains sparse, highlighting the need for further investigation into the cognitive processes that drive behavior in these tests.

A major obstacle to understanding these cognitive processes is the limited scope of current behavioral metrics. Most studies reduce complex behaviors, typically measured over a 5-6 minute period, to a single metric—total immobility time in the final minutes of the test—thus obscuring detailed behavioral patterns that could reveal deeper cognitive insights. As illustrated in Fig. 1G, animals with identical total immobility times may exhibit very different behavioral patterns, driven by distinct cognitive processes. While some pioneering studies have attempted to analyze high-resolution behavioral sequences[10], these often rely on labor-intensive manual annotations, which limits scalability and broader application. Consequently, there is a growing need for automated tools that can efficiently track behavior and generate fine-grained data to advance research.

In addition to improving behavioral data collection, another challenge is interpreting this data to uncover the cognitive mechanisms behind observed behaviors. In human psychiatric research, computational models have been employed to analyze behavioral sequences in tasks such as probabilistic reversal learning[20–22], Iowa gambling[23], fear conditioning[24] and two-stage decision[25,26]. These models have successfully identified cognitive differences between individuals, particularly in how they process rewards and punishments[27–32]. For example, studies show that patients with depression exhibit altered reward sensitivity[33], prediction errors[34–36], and learning rates[37,38] compared to healthy individuals. Such findings have refined cognitive models of depression and provided a framework for understanding the disorder and developing effective treatments. Inspired by these successes in human research, there is a growing trend in animal studies to apply similar computational models, such as reinforcement learning models, to analyze rodent behavior[39–43]. This paper explores the cognitive mechanisms underlying rodent behavior in the FST and TST using reinforcement learning models, with the broader goal of developing new methods for investigating implicit cognitive processes in rodent research.

To this end, we developed an automated tool designed to detect animal action states in the FST and TST. The tool utilizes passive motion offset technology in combination with a frame-difference algorithm to accurately recognize and track behavioral trajectories. Using the resulting data, we applied several reinforcement learning models to simulate the animals' learning and decision-making processes. Through model comparison, we identified those models that best characterize the cognitive mechanisms underlying behavior. Our findings demonstrate that reinforcement learning models effectively capture the evolution of mouse behavior, supporting the hypothesis that increased immobility in the FST and TST reflects underlying learning and memory processes. Notably, we found that chronic restraint stress (CRS) and inhibition of pyramidal neuron activity in the medial prefrontal cortex (mPFC) induced significant cognitive changes, leading to altered behavioral responses. Although the FST and TST are often considered analogous, our results reveal different cognitive mechanisms associated with each test. Additionally, linear regression analysis indicated that during the early stages of the experiments, mouse behavior was primarily influenced by the learning rate, while consequences sensitivity and vigor became more prominent in the later stages. This suggests that traditional analyses, which focus solely on immobility during the final minutes, may misrepresent cognitive processes by overemphasizing consequence sensitivity and vigor, while neglecting crucial aspects of learning.

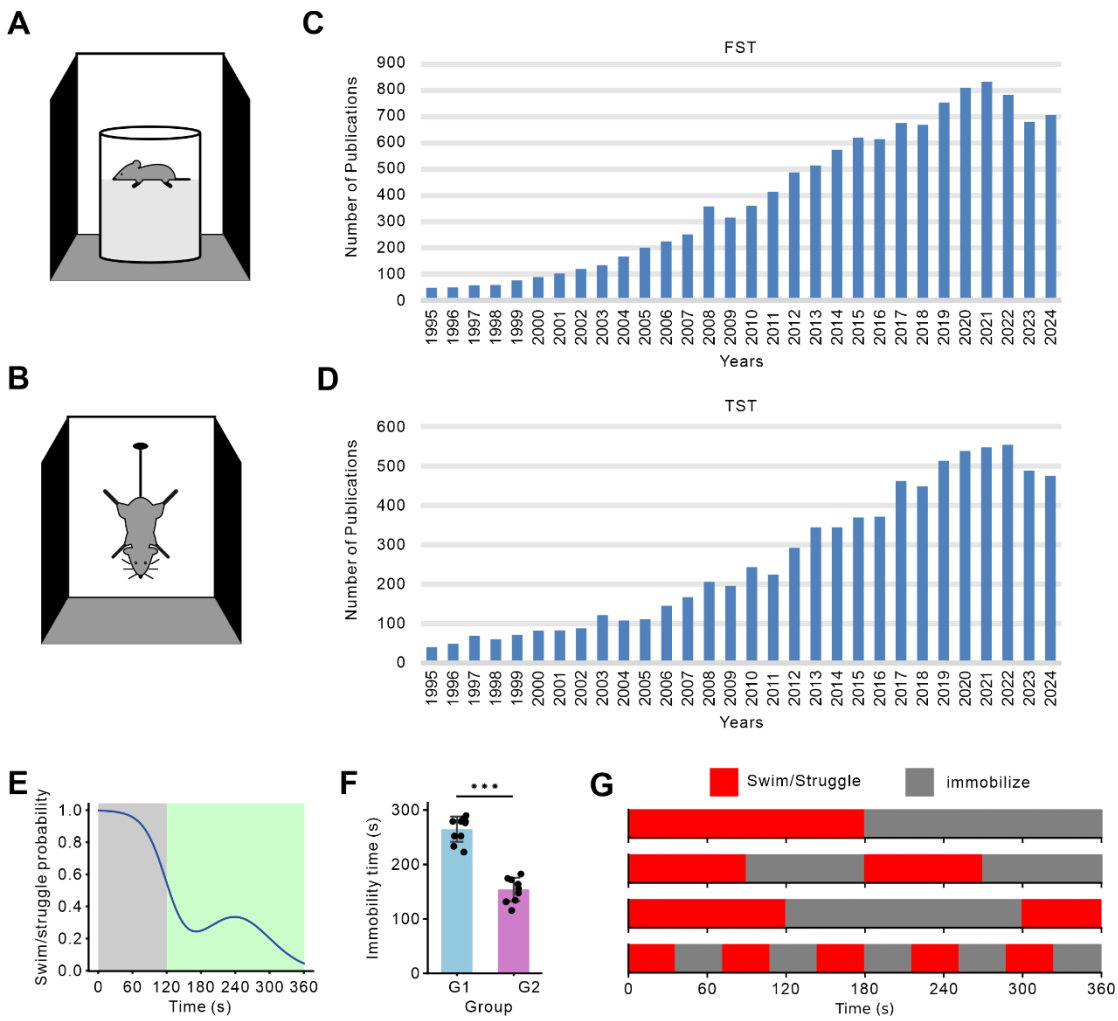

Fig. 1 | Behavioral Metrics in the FST and TST. **A, B.** Experimental setup for the FST and TST. In the FST (A), mice are placed in a container of water, while in the TST (B), they are suspended upside down. Both tests subject the mice to inescapable conditions for 5 to 6 minutes. Behavioral classifications in the FST include swimming and immobility, while in the TST, behaviors are categorized as struggling or immobility. **C, D.** Trends in research using the FST (C) and TST (D) based on the number of publications indexed in Web of Science over the past two decades, showing a steady increase in studies utilizing these paradigms. **E.** Illustration of an averaged behavioral trajectory, highlighting a general decline in swimming or struggling behavior over time with intermittent fluctuations in the later stages of the test. Traditional analyses often emphasize behavior during the final minutes (in green background) while neglecting data from the earlier phases (in grey background). **F.** Conventional methods summarize total immobility or swimming/struggling times by presenting group mean immobility times, standard error of the mean (SEM), individual data points, and statistical results. **G.** Cartoons illustrating potential behavioral trajectories for individual animals highlight the complexity of their actions. Even if a mouse remains immobile for half of the 6-minute test, distinct patterns of behavior can still emerge. Asterisks indicate significance differences between groups, as determined by the Mann–Whitney U test; ***$p < 0.001$.

Result

Tool for automatic behavior recognition

Current behavior metrics offer limited insights, highlighting the need for more advanced analyses based on behavioral trajectories. Manual annotation of these trajectories is time-consuming and prone to variability, which can affect both accuracy and consistency. To address these challenges, we developed an automated tool to annotate animal behavior in the FST and TST. As shown in Fig. 2A-D, the tool's core workflow includes several key modules: frame calibration and region of interest (ROI) selection, passive motion elimination, active motion detection, and behavioral trajectory smoothing and binning.

The process begins with video recordings of the mice, which are processed through the calibration and ROI selection module. This module corrects for image distortion due to camera angles and allows users to define the region where mouse movement is expected, termed the ROI. Next, the passive motion elimination module filters out irrelevant movements, such as floating or swaying, which could otherwise result in false detections. This step enhances the accuracy of detecting active behaviors like swimming or struggling. The active motion detection module then identifies the mouse's state—whether it is swimming, struggling, or immobile—on a frame-by-frame basis. Finally, the time-window statistical analysis module smooths the data and bins it into 2-second intervals, producing a clean and reliable behavioral trajectory for further analysis and modeling.

We validated the tool's performance by comparing its results to human expert annotations. A strong correlation was found in total immobility times (Fig. 2E: FST: $r = 0.411$, $p = 0.002$; TST: $r = 0.823$, $p < 10^{-5}$, Pearson correlation) as well as in the ranking of immobility times (Fig. 2E, G: FST: $r = 0.419$, $p = 0.001$; TST: $r = 0.720$, $p < 10^{-5}$, Spearman correlation). These results demonstrate that the automated tool produces outcomes closely aligned with manual annotations, and all subsequent analyses in this study will be based on the automated results.

Behavior of mice in FST and TST

We analyzed mouse behavior in two experimental paradigms: chronic restraint stress (CRS) exposure and inhibition of mPFC pyramidal neurons using chemical genetics. In the CRS paradigm, 29 mice were housed under identical conditions and divided into two groups: a CRS group ($n = 13$), which underwent daily restraint stress for 6-8 hours over 21 days, and a control group ($n = 16$), which experienced no stress. On days 22 and 24, both groups underwent the FST and TST, respectively (Fig. 3A). Behavioral trajectories and immobility times are shown in Fig. 3B-G. The CRS group demonstrated significantly increased immobility time across both tests compared to the control group (FST: 218.00 ± 31.87 vs. 147.25 ± 83.15s, mean ± std, $p = 0.009$; TST: 205.08 ± 65.76 vs. 145.13 ± 55.19, $p = 0.016$), and also showed higher immobility in the final 4 minutes of the FST (182.00 ± 23.42 vs. 119.50 ± 71.20, $p = 0.007$).

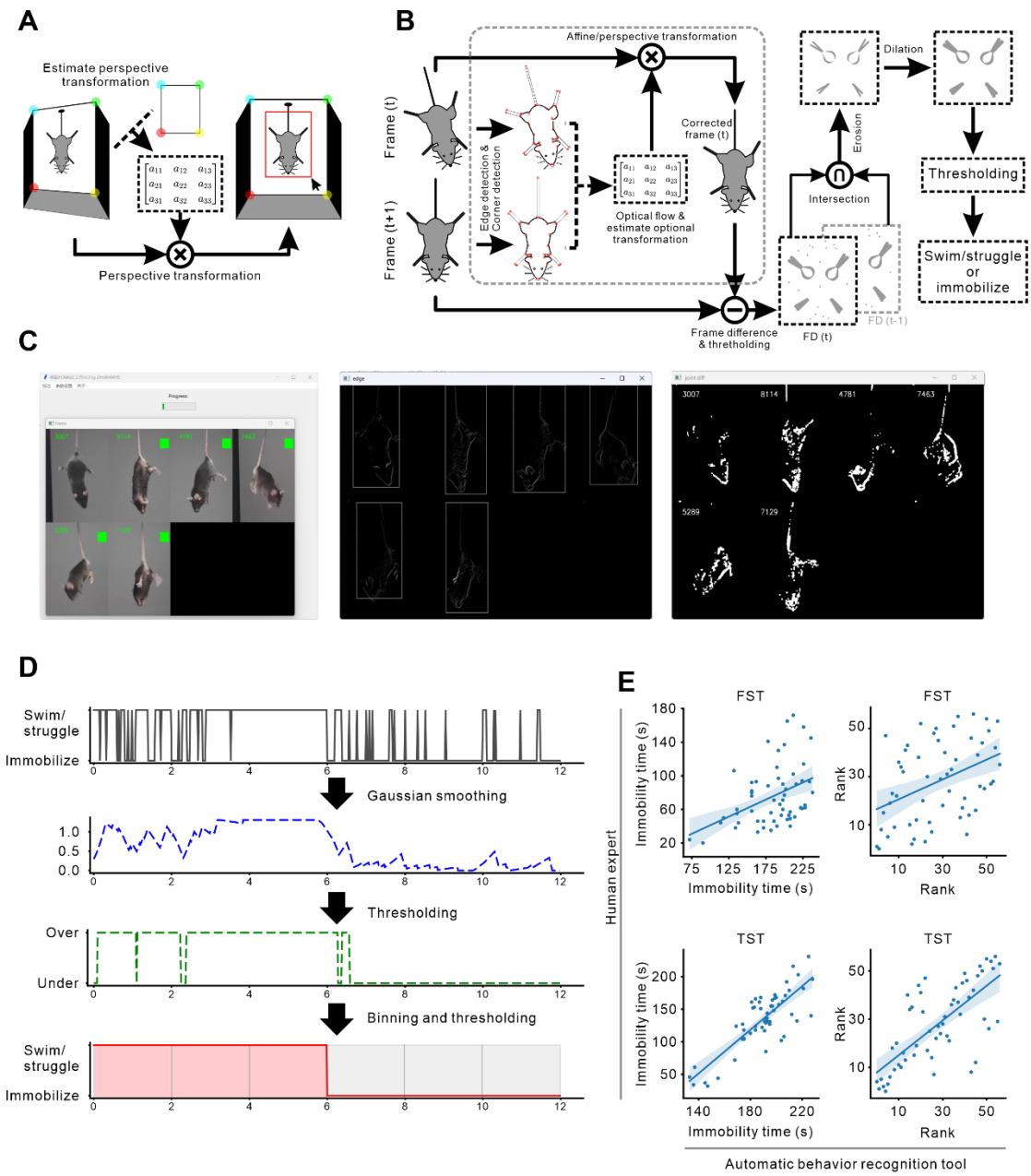

Fig. 2 | Automated Tool for Mouse Behavior Recognition in FST and TST. **A.** Frame calibration and ROI selection: To improve behavior recognition accuracy and account for camera angles, perspective transformation is applied to raw video frames. Users manually select the four vertices of the experimental box in a selected video frame, which are then compared to the pre-measured dimensions of the box. This process calculates a perspective transformation matrix to correct the frames. After correction, the region of interest (ROI) is defined as the maximum possible range of mouse movement. **B.** Workflow of the automatic recognition: The tool determines behavioral states by calculating inter-frame differences across consecutive video frames. Noise reduction is achieved through morphological operations. Behavioral trajectories are refined using smoothing, binning, and thresholding techniques. An optional module for eliminating passive motion (e.g., floating or swinging) uses geometric transformations to further enhance detection accuracy. **C.** Interface of the tool: The user-friendly interface enables streamlined interaction with the tool, providing visualization of intermediate outputs from key processing steps. **D.** Smoothing and binning of behavioral trajectories: Raw behavioral trajectories are often noisy, as seen in cases where

a 6-second struggle followed by 6 seconds of immobility produces erratic data. Applying Gaussian smoothing assigns a behavioral value to each frame, reducing noise. Thresholding transforms the smoothed trajectory into a binary profile, which is then binned into 2-second intervals to create a clear, finalized trajectory. **E.** Validation through comparison with expert annotations: Automated immobility times for both the FST and TST show strong correlations with expert manual annotations. Additionally, the ranking of immobility times produced by the tool aligns closely with expert evaluations, confirming the tool's reliability across both experimental paradigms.

Given previous studies linking alterations in the excitation-inhibition (E-I) balance of the prefrontal cortex to depression-like behaviors[44], we next examined how modulation of mPFC neuronal activity affects behavior in the FST and TST. A total of 24 mice were equally divided into two groups: a control group and a hM4D group, in which mPFC pyramidal neurons were inhibited. The hM4D groups received injections of rAAV-CaMKIIα-hM4D(Gi)-mCherry virus, while the control group received injections of rAAV-CaMKIIα-mCherry virus from days 1 to 7. After 28 days of viral expression, the mice underwent the FST and TST on days 36 and 38, following CNO injections to inhibit mPFC neuronal activity in the hM4D group (Fig. 3H). Behavioral trajectories and immobility times for this paradigm are shown in fig. 3I-N. While the hM4D group showed a trend toward increased immobility compared to controls in the FST, the differences were not statistically significant (overall: 199.17 ± 69.73 vs. 161.00 ± 60.36, $p = 0.184$; final 4 minutes: 162.17 ± 59.10 vs. 127.33 ± 59.25, $p = 0.181$). However, excluding one outlier mouse from the hM4D group—characterized by almost continuous struggling—revealed statistically significant differences in immobility times (Fig. S1, overall: 216.73 ± 40.05, $p = 0.060$; final 4 minutes: 176.73 ± 35.59, $p = 0.049$).

These findings are consistent with existing literature but offer a more detailed analysis compared to traditional methods, which typically yield a single scalar value per test. Using our automated recognition tool, we obtained behavioral trajectories that provide a rich dataset for more nuanced analysis. Notably, the probability of swimming and struggling typically decreases after the start of the test but shows a rebound and increased variability during the later stages. This dynamic pattern highlights the importance of considering the entire behavioral trajectory, as it provides insights into the cognitive processes underlying behavior—insights often overlooked in conventional analyses focused solely on immobility.

Computational models for analyzing behavior in FST and TST

Reinforcement learning (RL) models have proven valuable in understanding the cognitive processes underlying behavioral trajectories in psychological and cognitive neuroscience research with human subjects. These models also hold significant promise for analyzing mouse behavior in both the FST and the TST. In this context, we conceptualize the mice as agents making decisions at each time step, aiming to optimize outcomes—either by maximizing rewards or minimizing punishments—through interactions with the experimental environment.

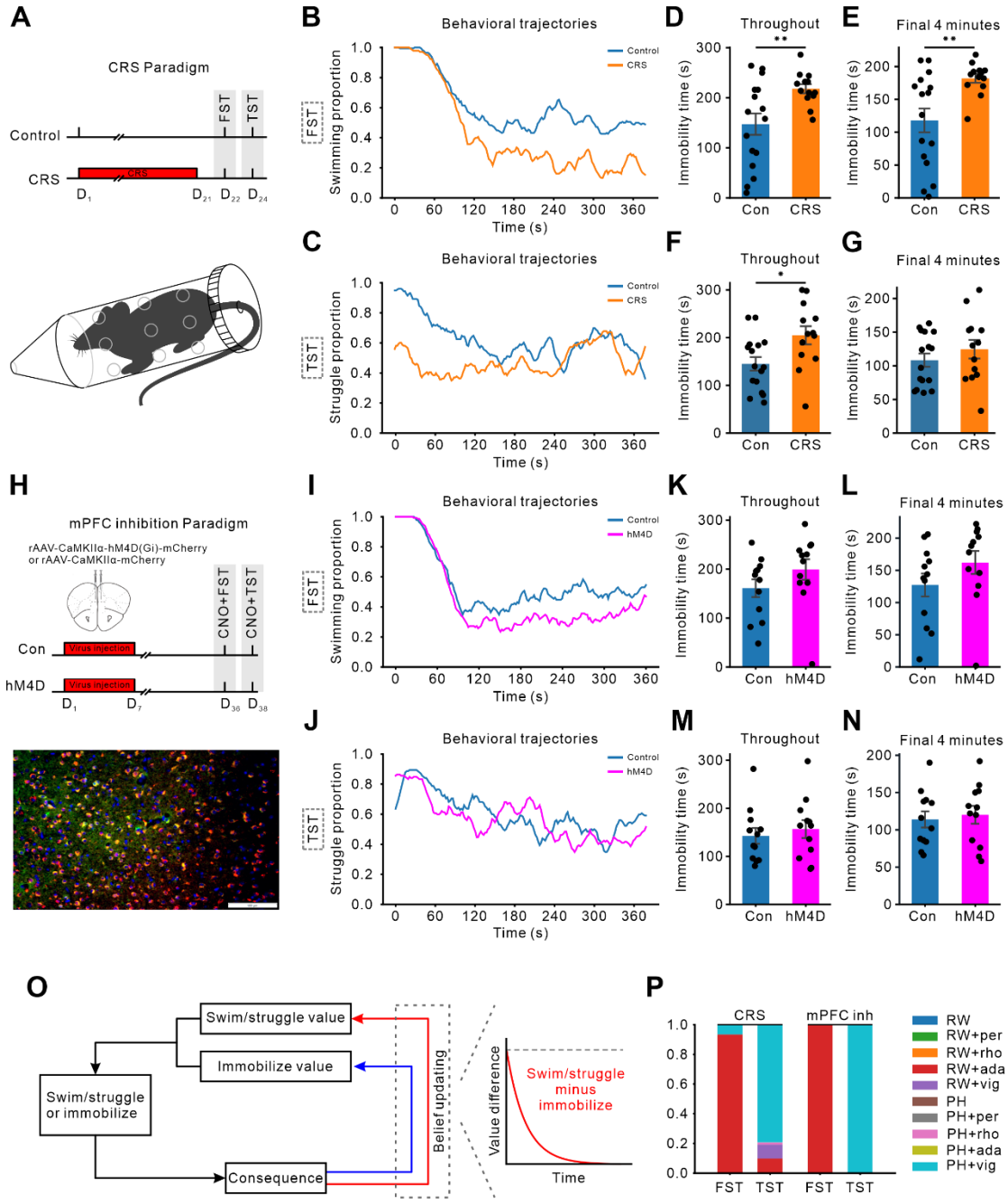

Fig. 3 | Behavior analysis of mice in FST and TST under two experimental paradigms. **A.** The chronic restraint stress (CRS) paradigm. Following a 7-day acclimation period, 29 mice were divided into a CRS group (n=13) and a control group (n=16). The CRS group underwent daily restraint stress from days 1 to 21, while the control group was unstressed. The FST and TST were conducted on days 22 and 24, respectively, for both groups. **B, C.** Averaged behavioral trajectories of mice during the FST are depicted, with blue representing the control group and yellow indicating the CRS group. **D-G.** Immobility duration during FST and TST. The immobility times for mice in both the FST (D, E) and TST (F, G) are shown for the entire testing period (D, F) and the final 4 minutes (E, G). Blue bars indicate the control group, yellow bars represent the CRS group, and black dots denote individual animal data. Error bars indicate the standard error of the mean (SEM). **H.** The mPFC pyramidal neuron inhibition paradigm. Twenty-four mice were evenly assigned to control and hM4D groups after a 7-day acclimation period. From days 1 to 7, hM4D mice received bilateral injections of rAAV-CaMKIIα-hM4D(Gi)-mCherry into the mPFC, while

control mice were injected with rAAV-CaMKIIα-mCherry. Viral expression was allowed to mature over 28 days. The FST and TST were conducted on days 36 and 38, respectively. Ninety minutes before testing, both groups received clozapine-N-oxide (CNO) injections, which specifically inhibited pyramidal neuron activity in the hM4D group. Immunofluorescence confirmed the expression of hM4D(Gi)-mCherry in mPFC pyramidal neurons in the hM4D group (red=hM4D(Gi), green=CaMKIIα, blue=DAPI; scale bar = 100 μm). Viral expression colocalized with pyramidal neurons at a rate of 89.27% (n=3 mice). **I, J.** Averaged FST behavioral trajectories in mPFC inhibition paradigm. As in panels B and C, behavioral trajectories are shown with blue representing the control group and magenta indicating the hM4D group. **K-N.** Immobility duration during FST and TST in mPFC inhibition paradigm. The immobility times for the FST (K, L) and TST (M, N) are presented, following the same format as panels D-G. Blue bars represent the control group, and magenta bars represent the hM4D group **O.** Reinforcement learning model schematic. The reinforcement learning model illustrates how mice assign expected values to two potential actions—swimming versus immobility (FST) or struggling versus immobility (TST). Action selection is guided by the difference in expected values, which are continuously updated based on environmental feedback. Initially, active behaviors like swimming or struggling are assigned higher values as escape strategies, but repeated failures decrease these values, leading to reduced active responses and increased immobility over time. **P.** Model weights across tests and experimental paradigms. Bar colors represent distinct models, while bar lengths indicate their respective weights. The winning model, defined as the model with the highest weight, is RW+ada for the FST and PH+vig for the TST in both paradigms. Asterisks indicate significance differences compared to the control groups, as determined by the Mann–Whitney U test; $^{**}p < 0.01$, $^{*}p < 0.05$.

We hypothesize that, at each moment, mice decide between active coping behaviors, such as swimming or struggling, and passive coping behaviors, like immobility, based on their expectations of future consequences. As the experiment progresses, these decisions are continually re-evaluated and adjusted according to the consequences of previous actions, leading to shifts in behavioral preferences. For instance, early in the experiment, mice may perceive swimming or struggling as more valuable, believing that these actions could help them escape. However, with repeated failures and mounting exhaustion, the perceived value of these active behaviors declines, making immobility a more likely choice. This dynamic process of action re-evaluation, driven by updates from past experiences, is illustrated in Fig. 3O.

To model this process, we fitted ten RL models (described in detail in the Methods section), including the Rescorla-Wagner (RW) model and the Pearce-Hall (PH) model, along with variants that incorporate sensitivity to consequences, sensitivity adaptation, behavioral perseveration, and vigor. Bayesian model selection was used to identify the model that best captures the underlying cognitive processes. Models were compared using Bayesian bootstrap and model averaging, with weights assigned based on log-likelihood and model complexity. The model with the highest weight was selected as the "winning model."

Fig. 3P illustrates the winning models for both the FST and TST. An intriguing finding was that the winning models differed between the two tests. In the FST, the RW+ada model, which updates beliefs with a fixed learning rate while dynamically adjusting sensitivity to the consequences of actions, emerged as the winning model. In contrast, the PH+vig model, which adjusts attention allocation based on discrepancies between perceived and expected consequences, dynamically alters the learning rate, and incorporates a vigor factor, was the winning model in the TST. Notably, all winning models successfully

recovered model parameters (Fig. S2), demonstrating their robustness in capturing the behavioral data.

Model validations

To assess the efficacy of these computational models, we conducted posterior predictive checks to evaluate how well the models predicted mouse behavior. As shown in Fig. 4A-D and K-N, the winning models accurately recovered the observed immobility times across all groups throughout the task and during the final 4 minutes. Additionally, nearly all individual immobility times fell within the simulated 95% highest density interval (HDI) (Fig. 4E-H, and O-R), with strong correlations between the predicted and observed immobility times across all experimental conditions and time windows ($r>0.95$, $p<10^{-5}$). Fig. 4I, J, S, T further show that the models made relatively accurate predictions of each mouse's latency to immobility. These results underscore the predictive validity of the models for individual behaviors. Moreover, the winning models successfully replicated the behavioral trajectories of the mice (Fig. 4U-X), capturing the dynamic patterns of behavior over time. Overall, these computational models provide a robust simulation of mouse behavior.

These findings indicate that the evolution of mouse behavior in the FST and TST can be effectively described by reinforcement learning models, representing an ongoing process of learning and decision-making. Furthermore, the differences between the winning models for these two tasks suggest that the cognitive processes driving behavior in the FST and TST are different, challenging the common assumption that these two tests are interchangeable for cross-validation.

Impact of CRS and mPFC pyramidal neuron inhibition on model parameters

The model parameters for individuals in each experimental group are presented in Fig. 5A-D. In the FST, both the CRS and hM4D groups showed significant reductions in the learning rate (α) and the lower bound of consequence sensitivity (lb), while the adaptation rate of consequence sensitivity (β) was significantly increased compared to the control group. Additionally, the hM4D group exhibited a higher inverse temperature (τ), which reflects the confidence in the agent's decisions. In the TST, both the CRS and hM4D groups exhibited an increase in the normalization parameter for the learning rate (κ) and a decrease in the τ relative to the control group. Furthermore, the CRS group displayed an elevated attention updating rate (η), while showing decreased sensitivity to consequences (ρ) and vigor bound (vb). In contrast, the hM4D group showed an increase in ρ.

We further analyzed changes in intermediate variables such as the value differences between swimming/struggling and immobility, adapted consequence sensitivity, and attention allocation over the course of the test (Fig. 5E-L). Overall, the value differences showed a trend of gradually decreasing to a stable low level. In the FST, both the CRS and hM4D groups exhibited lower value differences at later stages compared to the control group. In the TST, the CRS group had lower value difference than the control group in the early stages, while the hM4D group did not showed significant differences throughout the test. Additionally, both the CRS and hM4D groups displayed reduced sensitivity to consequences in the early stages of the FST, while the CRS group exhibited higher attention allocation to struggling behavior in the early stages of the TST.

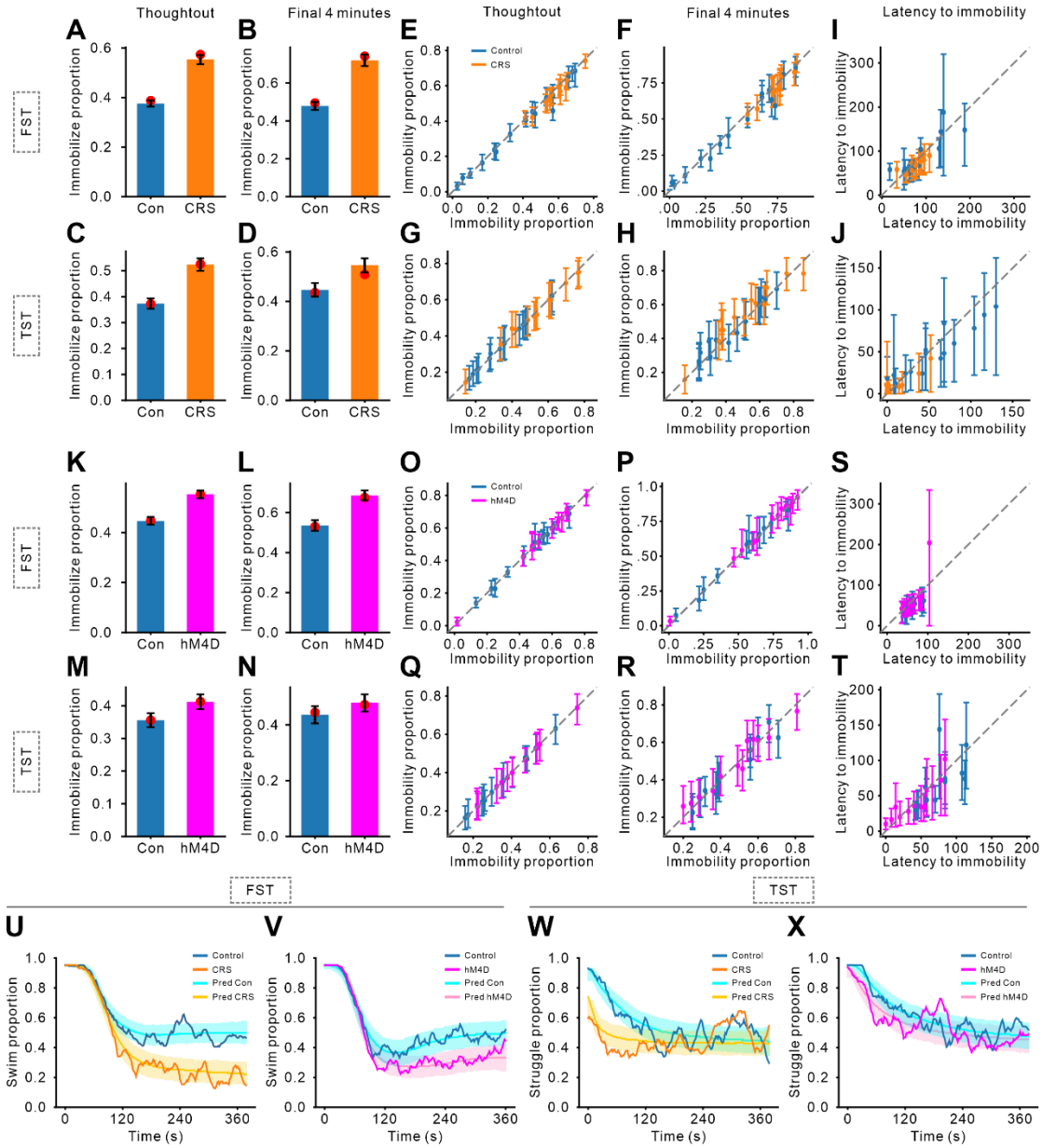

Fig. 4 | Posterior predictive checks for the winning models. **A-D.** Predicted immobility proportions for the CRS paradigm during the full testing period (A, C) and the final 4 minutes (B, D) for the FST (A, B) and TST (C, D). Colored bars indicate the median predictions from the winning models, with error bars showing the 95% highest density interval (HDI). Red dots represent observed experimental averages, demonstrating alignment between predictions and data. **E-H.** Predicted versus observed immobility times for individual mice during the full experiment (E, G) and the final 4 minutes (F, H) in the FST (E, F) and TST (G, H). The horizontal axis reflects observed values, and the vertical axis shows median model predictions based on 2000 simulations. Error bars indicate the 95% HDI, with distinct colors denoting experimental groups. **I, J.** Comparison of predicted versus observed latencies to immobility in the FST (I) and TST (J), respectively. The data are presented in the same format as panels E-H. **K-N.** Results analogous to panels A-D, focusing on mPFC pyramidal neuron inhibition. **O-T.** Predictions versus observations for mPFC pyramidal neuron inhibition, corresponding to panels E-J. **U-X.** Behavioral trajectory predictions over the testing period. Panels U, V show swimming behavior predictions for the FST, while W, X depict struggling behavior for the TST across groups. Cyan, yellow, and pink curves

represent median predictions, with shaded areas indicating the 95% HDI based on 2000 simulations. Blue, orange, and magenta lines correspond to observed group averages. The models effectively capture behavioral dynamics, as most observed values fall within the 95% HDI.

The decrease in lb indicating that mice subjected to CRS and mPFC pyramidal neuron inhibition in the FST may experience heightened feelings of despair following failed escape attempts. Simultaneously, the increase in β likely accelerated this sense of despair, as indicated by the yellow and magenta curves in Fig. 5F and H, which consistently fall below the blue curve, showing a steeper decline in the earlier stages. This shift led to a reduction in the relative value of active coping behaviors (Fig. 5E and G, where the yellow and magenta curves are lower than the blue curve), ultimately leading to a higher incidence of immobility. In the TST, for the CRS group, the increase in η suggests that these mice more rapidly attended to the discrepancy between expected and actual outcomes of struggling (Fig. 5J, where the yellow line is above the blue line in the early stages). Coupled with a higher κ, this led to a quicker learning of the negative outcomes of struggling (Fig. 5I, where the yellow line declines more rapidly to a stable level), which subsequently resulted in faster onset of immobility.

Contribution of model parameters to behavior

To examine the influence of model parameters on behavior, we conducted simulations by varying two parameters at a time while holding others constant at their median values, which were derived from the posterior distribution of the winning models (the result of the CRS and mPFC inhibition paradigms were shown in Fig. 6B-F and Fig. S3, respectively). The results for the FST, shown in Fig. 6B and C, reveal that higher values of both the upper bound (ub) and lower bound (lb) of consequence sensitivity were associated with longer swimming durations. The relationship between α and β was more complex: very low α values prolonged swimming, while moderate β values, when combined with higher α, promoted swimming behavior. Fig. 6D-F presents the TST results, showing that longer struggling durations were linked to higher values of vb and ρ, along with lower values of η, κ, and τ.

To explore the impact of model parameters on behavior in greater detail, we analyzed their relationship with behavioral performance at each time step. Linear regression models were used to predict cumulative swimming or struggling times based on parameters sampled from the posterior distribution of the winning models. The contribution of each parameter was assessed by comparing their explanatory power (r^2). The analysis revealed that different parameters exerted varying degrees of influence at different stages of the experiment. In the early stages of the FST, α from the RW+ada model was the strongest predictor of cumulative swimming time. However, as the experiment progressed, the influence of α diminished, while the bounds of consequence sensitivity adaptation (lb and ub) became more significant (Fig. 6H, I). In the TST, η from the PH+vig model was the dominant factor early in the test, while vb and ρ gained importance in the later stages (Fig. 6J, K).

These findings suggest that traditional analytical methods, which often overlook early-phase data, may underestimate the role of learning rate-related factors and overemphasize consequence sensitivity and vigor-related factors when interpreting animal behavior.

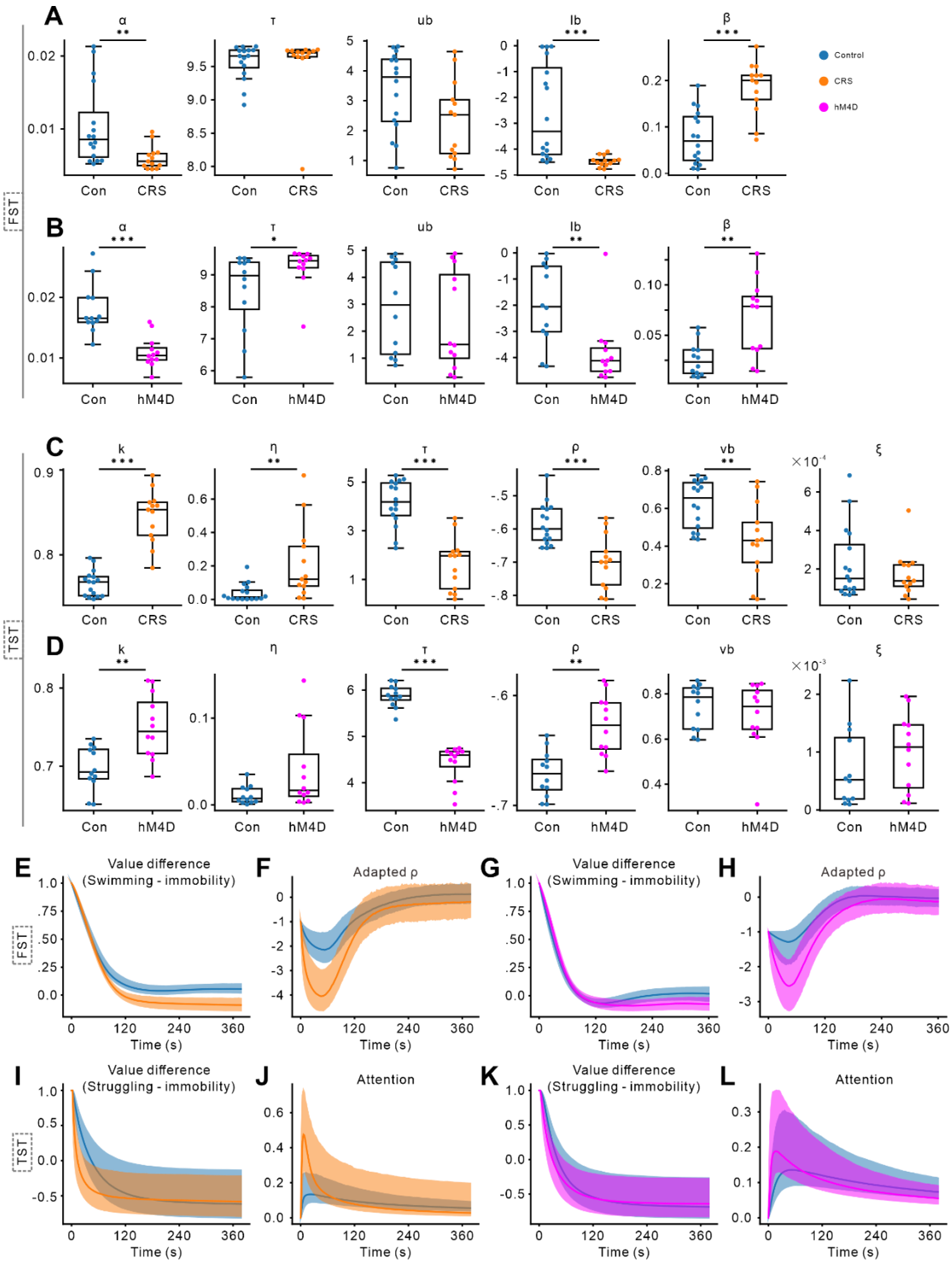

Fig. 5 | Comparison of model parameters. **A, B.** Parameter distributions for the winning model (RW+ada) in the FST. Each panel represents a parameter, with each point showing the mean of 2,000 samples drawn from the posterior distribution of an individual mouse. Boxplots display the group-level medians and interquartile ranges. Blue, orange, and magenta represent the control, CRS, and hM4D groups, respectively. **C, D.** Parameter distributions for the winning model (PH+vig) in the TST, presented similarly to panels A and B. **E-H.** Predicted dynamics of value expectation differences (swimming – immobility, E, G) and adapted sensitivity to consequences (F, H) over the course of the FST, based on the winning models. Panels E and F show results for the CRS paradigm, while panels G and H display results

for the mPFC pyramidal neuron inhibition paradigm. Solid lines represent group-level medians, with shaded areas indicating the 95% HDI based on 2,000 posterior samples. **I-L.** Predicted dynamics of value expectation differences (struggling – immobility, I, K) and attentional allocation (J, L) over the course of the TST. Panels I and J show results for the CRS paradigm, and panels K and L present results for the mPFC inhibition paradigm. As in the FST panels, solid lines indicate group-level medians, with shaded areas representing the 95% HDI. Asterisks indicate significance differences compared to the control groups, as determined by the Mann–Whitney U test; ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

Discussion

In this study, we introduced an automated tool for recognizing animal behavior, enabling the efficient collection of detailed behavioral trajectories in the FST and TST. Our computational modeling analysis revealed that the evolution of mouse behavior in these tasks is driven by reinforcement learning processes involving both learning and decision-making. Notably, the cognitive processes underlying behavior in the FST and TST were not entirely identical, challenging the common assumption that these tasks are equivalent or mutually validating. Through analyses of model parameters representing distinct cognitive components, we found that in the FST, CRS and mPFC inhibition primarily increased immobility by altering the adaptation process of consequence sensitivity. In contrast, in the TST, CRS promoted immobility by modulating attention-based learning rates, consequence sensitivity, and vigor. Furthermore, we observed that different cognitive components shaped behavior at various stages of the experiment. This suggests that traditional methods, which often exclude early-phase data, may underestimate the influence of learning rate-related factors while overemphasizing the roles of consequence sensitivity and vigor. Overall, our findings demonstrate that detailed behavioral trajectories provide deeper insights than conventional metrics, highlighting the potential of computational modeling to uncover the complex cognitive processes driving behavior.

The necessity of using fine-grained behavioral trajectories

The FST and TST are widely used as preliminary methods for studying depression in rodents. However, as interest in understanding animal behavior has grown, attention has shifted to the cognitive insights these tests can provide. Traditional metrics, which typically report immobility time during the final minutes as a single value, fail to capture the complexity and evolution of behavior throughout the test. This limitation has restricted our ability to fully understand the cognitive processes assessed by the FST and TST. Consequently, there is a growing need for fine-grained behavioral trajectory data. Currently, researchers often rely on labor-intensive manual annotation to identify behaviors in the FST and TST. This process, which involves second-by-second or frame-by-frame analysis, is both time-consuming and labor-intensive, creating significant barriers to the widespread adoption of behavioral trajectory analysis.

To overcome these challenges, we developed an easy-to-use tool specifically designed for automatically generating behavioral trajectories in the FST and TST. This Python-based tool enables researchers to identify animal action states with frame-level temporal resolution. By leveraging the detailed behavioral trajectories provided by this tool, researchers can detect subtle changes in behavior patterns, apply computational modeling, and gain deeper insights into the cognitive processes underlying animal behavior.

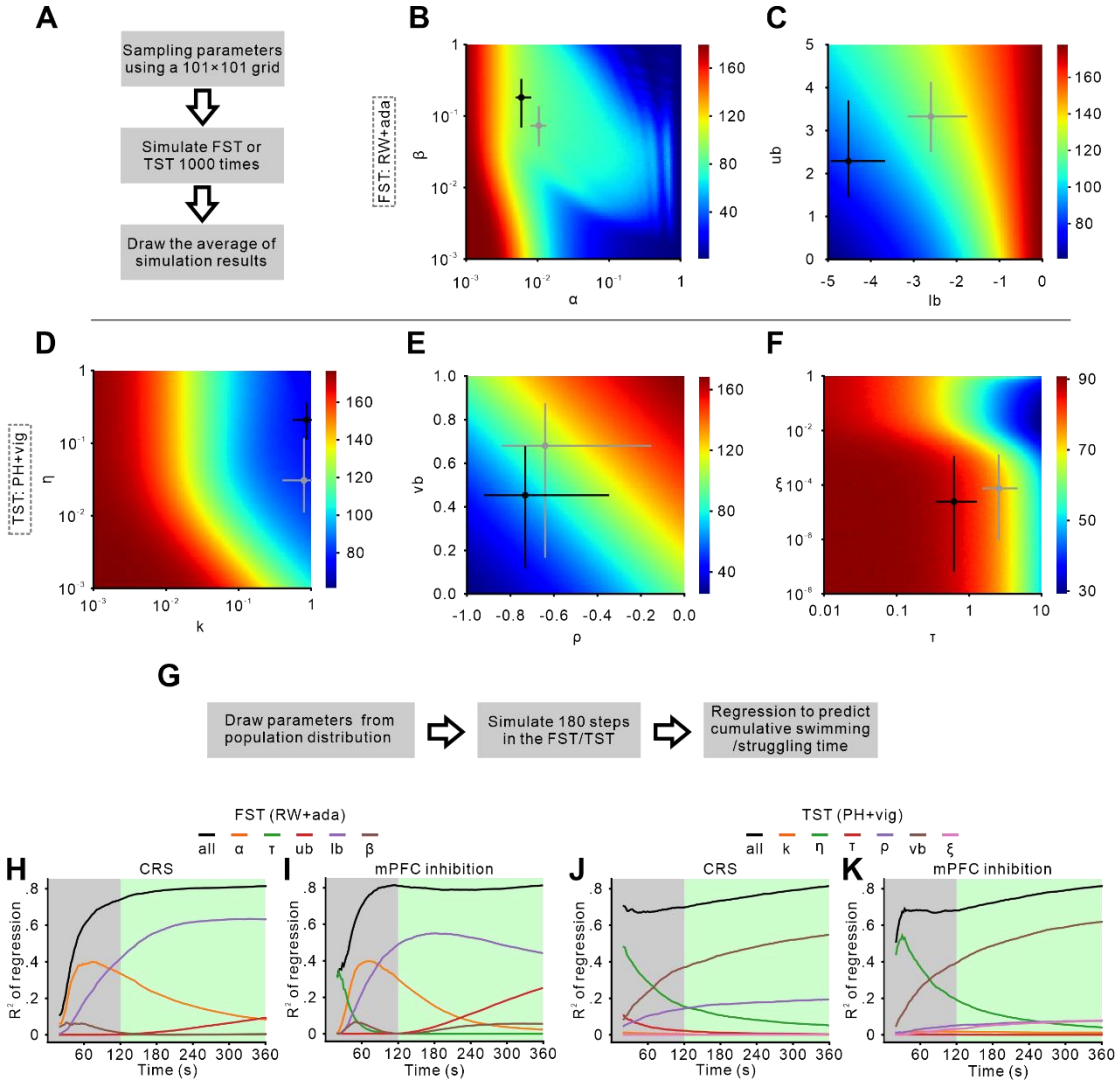

Fig. 6 | Linking model parameters to behavioral patterns. **A.** Diagram outlining the simulation workflow. Parameters were sampled from a predefined grid, simulations were run for the FST or TST, and the resulting average behavioral metrics were computed. **B-F.** Heatmaps showing simulated total swimming and struggling times across the parameter space of the winning models. Each heatmap explores two selected parameters while holding others constant at their posterior medians. The x- and y-axes represent the sampled parameter values, with brighter colors indicating longer swimming or struggling durations. Circles mark the posterior median estimates, and horizontal/vertical lines show the 95% HDI. Black markers represent the CRS or hM4D groups, while gray markers indicate the control group. Panels B and C correspond to parameters from the RW+ada model (winning model for FST), while panels D-F display parameters from the PH+vig model (winning model for TST). **G.** Workflow for deriving relationships between model parameters and behavior. Parameters were drawn from population-level posterior distributions, simulated across 180 steps, and used to compute explanatory regressions for behavioral metrics. **H-K.** Explanatory power (r^2) of individual and combined model parameters from linear regression analyses for cumulative swimming and struggling times in FST (H, I) and TST (J, K). Black curves represent r^2 values for models including all parameters, while colored curves reflect models using only one parameter. Panels H and J display results for the CRS paradigm, while panels I and K illustrate findings for the mPFC inhibition paradigm. Parameter contributions evolved over time: learning rate-related parameters (α, η) were most influential in the early phases, whereas parameters tied to

consequence sensitivity (ub, lb, ρ) and vigor (vb) gained importance in later stages.

Using computational models to analyze FST and TST data

This study represents a pioneering effort to analyze mouse behavior in the FST and TST using reinforcement learning models, providing a deeper understanding of the cognitive processes driving behavioral evolution. We framed the continuous decisions between swimming/struggling and immobility as a reinforcement learning task, applying the classic Rescorla-Wagner (RW) and Pearce-Hall (PH) frameworks along with their variants to model behavior. While the RW and PH models predict that the value expectations of actions gradually converge with perceived consequences over time, our observations in the FST and TST show that mouse behavior does not stabilize in the later stages. Instead, after extended periods of immobility, swimming and struggling behaviors exhibit rebounds and fluctuations. To enhance the realism of our simulations, we incorporated two additional factors: consequence sensitivity adaptation and vigor.

Consequence sensitivity adaptation refers to adjustments in how mice perceive outcomes following repeated failed escape attempts through swimming or struggling or after prolonged reliance on immobility as a passive coping strategy. These adjustments influence value learning over time. Vigor reflects the motivational drive to engage in active coping behaviors, which decreases due to cumulative fatigue from repeated swimming or struggling but recovers during periods of rest. Vigor influences behavioral decisions jointly with value expectations. Models incorporating these factors emerged as the winning models for the FST and TST, accurately predicting immobility durations and the dynamic evolution of behavior, demonstrating effective simulation of both tasks.

Computational modeling is a powerful tool for uncovering the cognitive mechanisms underlying behavior and exploring related neural processes. For example, a recent study using sucrose preference tests showed that stress exposure and mPFC inhibition affect behavior primarily by altering reward sensitivity and learning rates, respectively[43]. Our findings build on this understanding by showing that CRS and mPFC inhibition modulate behavior in the FST primarily through changes in consequence sensitivity adaptation. In the TST, CRS influences behavior via alterations in attention-based dynamic learning rates and vigor. These computational modeling studies enhance our understanding of the cognitive dimensions assessed in behavioral experiments and provide valuable insights into the mechanisms underlying mental disorders.

Behavior in the FST and TST as reinforcement learning processes

The cognitive processes assessed by the FST and TST remain poorly understood. Although the learning and memory hypothesis is a prevailing perspective, it has not been rigorously supported with quantitative evidence. In this study, reinforcement learning models successfully predicted various behavioral metrics at both the individual and group levels, as well as the behavioral trajectories of mice in the FST and TST. These results suggest that behavior in these tests can be interpreted as reinforcement learning processes involving both learning and decision-making.

When initially placed in an uncomfortable environment, mice perceive swimming and struggling as potential escape strategies, assigning higher expected value to these active coping behaviors. This leads to a preference for these actions in the early stages of the experiment. However, repeated failed escape attempts generate prediction errors—discrepancies between the expected value of these actions and their

perceived consequences. These prediction errors drive a continuous reassessment of the expected values of each behavior. Over time, the expected value of swimming and struggling declines, perhaps falling below that of immobility eventually. This shift accounts for the observed increase in immobility during the later phases of the experiment.

FST and TST exam different cognitive component

The FST and TST are widely used in neuroscience and pharmacology to evaluate behaviors associated with depression and to assess the effects of antidepressant drugs. Although these tests share similar principles and analytical methods, our findings reveal that the winning models differ between them, suggesting that mice engage somewhat different cognitive processes to drive behavior in each task. The winning model for the FST is RW+ada, which indicates that mice update the value of actions at a fixed rate while dynamically adapting their sensitivity to consequences based on their history of action outcomes. In contrast, the winning model for the TST is PH+vig, which suggests that mice update value expectations at dynamic rates influenced by attention and consider both value and vigor during decision-making.

These findings challenge the common practice of using the FST and TST interchangeably for cross-validation, highlighting meaningful differences in the cognitive processes underlying these tasks. The FST appears to emphasize how animals dynamically perceive and adapt to the consequences of actions, whereas the TST may better capture how attention adjust the learning of value and motivation to adopt active coping behaviors such as swimming or struggling. In practice, inconsistent results often arise when using the FST and TST to evaluate depression-like behaviors. Our findings may help explain this phenomenon by demonstrating that these tasks assess different aspects of cognition. We hope that future research will further elucidate the cognitive processes measured by behavioral experiments and propose more appropriate methods for cross-validation.

Traditional methods overlook learning rates and overemphasize outcome sensitivity and vigor

Conventional analyses of FST and TST data often focus exclusively on the final minutes of the experiment. However, our findings suggest that this approach may lead to incomplete or misleading interpretations. Using computational models, we observed that in the FST, the learning rate (α) was the most influential factor in explaining simulated behaviors during the early stages. As the experiment progressed, the influence of α diminished, and the upper (ub) and lower (lb) bounds of consequence sensitivity became increasingly dominant. A similar progression was observed in the TST. Early in the experiment, the attention updating rate (η) exerted the strongest influence on behavior. However, its impact declined over time, and in later stages, behavior was increasingly driven by vigor bound (vb) and consequence sensitivity (ρ). These shifts highlight the critical role of learning rate in shaping behavior at the beginning of the test, with consequence sensitivity and vigor gaining prominence in the later phases.

Disregarding the first two minutes of the experiment—a common practice—risks underestimating the importance of learning rates, while overemphasizing the roles of outcome sensitivity and vigor. Such an oversight can limit our understanding of depressive-like behaviors and impede the development of effective antidepressant treatments. To address this, we recommend that future studies analyze full

behavioral trajectories, using computational models to gain deeper insights into the cognitive processes underlying behavior throughout the entire duration of the experiment.

Method

Subjects

Male C57Bl/6 mice (weighing 20–30g, 8 weeks old at the start of experiments) were obtained from Beijing SPF Animal Technology Co. for all experiments. The mice were housed under controlled conditions: a 12:12-hour light-dark cycle (lights on at 7:00 AM), a temperature of 25 ± 2°C, and humidity levels of 50 ± 10%. Behavioral experiments commenced after one week of acclimatization and were conducted.

Chronic restraint stress (CRS) model

The CRS procedure followed a widely accepted paradigm with slight modifications. In brief, mice were restrained in tubes with limited movement for 8 hours daily (10:00 AM–6:00 PM) over 21 consecutive days. Control mice were kept in their home cages without exposure to stress. After each restraint session, the mice were released and returned to their home cages. All experimental procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of the Beijing Institute of Basic Medical Sciences (Approval number: IACUC-06-2019-002).

Stereotaxic surgeries and chemical genetics

Stereotaxic surgeries were performed under general anesthesia with 1% sodium pentobarbital using a stereotaxic apparatus (RWD Life Science Co., Ltd, Shenzhen, China). To inhibit pyramidal neurons in the mPFC, 500nL of rAAV-CaMKIIα-hM4D(Gi)-mCherry (BrainVTA Co., Ltd, Wuhan, China) was bilaterally injected into the mPFC (250nL for each side) of C57Bl/6 mice at the coordinates: AP +2.34 mm, ML ±0.3 mm, and DV −1.8 mm from Bregma. The control group received the same injection protocol but with rAAV-CaMKIIα-mCherry (BrainVTA Co., Ltd, Wuhan, China). The infusion was performed using a glass pipette connected to a Hamilton syringe fitted with a 33-gauge needle. The syringe was mounted on a syringe pump (RWD Life Science Co.) to ensure precise control over the infusion process. The infusion rate was set to 50 nl/min, allowing for consistent and accurate delivery of the solution. The injection pipette was slowly withdrawn 10 minutes after the infusion. Behavioral experiments were conducted 4 weeks post-surgery.

On the day of the testing, Clozapine N-oxide (CNO) was administered intraperitoneally to inhibit pyramidal neurons(5mg/kg), and behavioral assessments began 90 minutes later. After the behavioral tests, the mice were perfused, and whole brains were collected for immunofluorescent staining to verify viral transduction. The results confirmed successful infection of 89.27% of pyramidal neurons, ensuring effective viral expression.

Forced swim test (FST) and tail suspension test (TST)

The FST was conducted in a well-lit room where mice were individually placed in water-filled cylinders (35 cm height × 10 cm diameter) with water maintained at 24–25°C. The water depth was sufficient to prevent the mice from touching the bottom with their limbs. Mice were allowed to swim freely for about 6 minutes while their movements were recorded on video.

For the TST, mice were gently removed from their cages, and their tails were secured with adhesive tape approximately 1 cm from the tip. The mice were then suspended by their tails, with the distance between the tip of the tail and the ground set at approximately 30 cm, placing the mice in an inverted position. The test lasted for approximately 6 minutes, and video recordings of the mice's behavior were obtained during the experiment.

Tool for automated behavioral recognition

In brief, the automated behavior recognition tool consists of four modules: the Calibration and Region of Interest (ROI) Selection module, the Passive Motion Compensation module, the Active Motion Detection module, and the Smooth and Thresholding Analysis module. We briefly describe the function of each module below.

1. **Calibration and ROI Selection Module:** As shown in Fig. 2A, we mitigate the impact of camera angles on recognition accuracy by applying a perspective transformation to the original images. Prior to detection, the user manually selects the coordinates of the experimental box's vertices in the video frame. By comparing these coordinates with pre-measured box dimensions, we calculate and apply a perspective transformation matrix to correct the original video frames. On the corrected frames, the largest possible movement area of the mouse is defined as the ROI. The tool offers two options for selecting the ROI: manually adjusting the size by dragging the mouse or choosing from preset sizes. The selected ROI is resized to 360x480 pixels for further analysis.

2. **Passive Motion Compensation Module (Optional):** As illustrated in the dashed box in Fig. 2B, this module focuses on analyzing the animal's contour map extracted from the region of interest (ROI), along with key points required for estimating optical flow. Optical flow techniques are used to track these key points in successive frames. In cases where the animal is passively moving—such as during inertial drift or pendulum motion—its posture remains static. Consequently, the relative positions of key points between adjacent frames remain constant, indicating only an overall shift. The Pyramid Lucas-Kanade optical flow algorithm is applied to detect corresponding key points in the following frame. By minimizing the geometric distance between predicted and actual key point positions, the algorithm determines the geometric transformation that best represents the animal's overall movement. This fitted transformation models the passive motion of the animal. Using this module, the current frame is adjusted to generate a passive motion-compensated prediction, enabling direct comparison with the subsequent frame for further analysis.

3. **Active Motion Detection Module:** As shown in Fig. 2B, active movement by the animal results in significant differences between adjacent frames, even after passive motion compensation. We use a three-frame differencing method to calculate inter-frame differences based on the ROI images from frames t and t+1. These differences are then applied a threshold to generate a binary difference image at frame t, which is combined with a cached binary difference image from frame t-1 using intersection

operations. Morphological operations (erosion and dilation) are adopted to remove noise, and the processed image is then summed and thresholded to determine the movement state (struggling or immobilize). Fig. 2C displays the current video frame, contour detection results (if passive motion compensation is enabled), and combined inter-frame differences to illustrate the tool's effectiveness.

4. **Smooth and Thresholding Analysis Module:** As illustrated in Figure 2D, the original behavioral trajectory for each frame often contains significant noise. For example, in a video where a mouse struggles for the first 6 seconds and remains immobile for the following 6 seconds, initial recognition may produce noisy results, as shown in the top row. To address this, Gaussian smoothing is applied to reduce noise and produce a clearer trajectory, as depicted in the second row. Next, thresholding is used to convert the smoothed data into a binary spectrum, shown in the second row from the bottom. Finally, the data is processed through binning and an additional thresholding step, resulting in a behavior trajectory with 2-second bins. This final trajectory, shown in the bottom row, is both more accurate and robust for subsequent analysis.

Reinforcement learning models

We employed reinforcement learning models to fit the behavioral trajectories observed in the FST and TST. In this framework, mice are treated as agents with the goal of optimizing outcomes, which includes maximizing rewards and minimizing punishments. Specifically, the mice make decisions based on their value expectations about actions, and update the value expectations by calculating prediction errors between value expectations and actual outcomes. Future action decisions are made based on these updated values. We implemented 10 reinforcement learning models to explore different value updating mechanisms and cognitive processes underlying behaviors in FST and TST.

1. Rescorla-Wagner model (RW): The Rescorla-Wagner model is a classic model of error-driven predictive learning. It assumes that mice update their beliefs by incorporating a fixed proportion of prediction error. Given that swimming and struggling deplete the mice's physical energy, making the subsequent failing to escape experiences more unpleasant than remaining immobile, we encoded the outcome of swimming and struggling as -1, while immobility was encoded as 0. The predictive error at timestep t was calculated as follows:

$$\delta_t = \begin{cases} -1 - Q_{a,t} & swim/strggle \\ 0 - Q_{a,t} & immobilize \end{cases}$$

The value expectation was updated as:

$$Q_{a,t+1} = Q_{a,t} + \alpha\delta_t$$

Here, δ_t represents the prediction error at timestep t, while $Q_{a,t}$ and $Q_{a,t-1}$ represent the value expectations for the actions taken by the mice at timesteps t and t-1, respectively. The parameter α is the learning rate, which governs the extent to which the prediction error is integrated into the updated value expectation.

A sigmoid function was applied to compute the probability of swimming/struggling based on the values expectation of actions. At each trial t, the probability of swimming/struggling $p_{s,t}$ was given by:

$$p_{s,t} = \frac{1}{1 + e^{-\tau(Q_{s,t} - Q_{i,t})}}$$

In this equation, $Q_{s,t}$ and $Q_{i,t}$ represent the value expectations of swimming/struggling and immobility, respectively. The probability of immobility is $1-p_{s,t}$. This decision function was adopted by RW, RW+ρ, RW+ada, PH, PH+ρ, and PH+ada models.

2. RW model with perseveration (RW+per): Building on the RW model, this model incorporates the influence of behavioral perseveration on the decision-making process. Behavioral perseveration is modeled based on the frequency of unchosen actions $uct_{a,t}$, which is updated according to the following rule:

$$uct_{a,t} = \begin{cases} 0 & if\ the\ action\ was\ not\ taken\ at\ t-1 \\ uct_{a,t-1} + 1 & if\ the\ action\ was\ taken\ at\ t-1 \end{cases}$$

The $uct_{a,t}$ are considered in decision. At each trial t, the probability of swimming/struggling $p_{s,t}$ was given by:

$$p_{s,t} = \frac{1}{1 + e^{-(\tau(Q_{s,t} - Q_{i,t}) + \phi(uct_{s,t} - uct_{i,t}))}}$$

Where $uct_{s,t}$ and $uct_{i,t}$ refers to the unchosen time for swimming/struggling and immobile, respectively. ϕ refers to the importance of perseveration.

3. RW model with consequence sensitivity (RW+ρ): Building on the RW learning rule, this model incorporates the sensitivity to the consequence of swimming and struggling. The sensitivity to the consequence is represented by the parameter ρ. In this model, after swimming or struggling, the predictive error is computed as follows:

$$\delta_t = \rho - Q_{a,t}$$

This adjustment reflects the idea that the mice perceive the consequences of their actions (e.g., swimming and struggling) with varying degrees of sensitivity, as captured by ρ .

4. RW model with consequence sensitivity adaptation (RW+ada): This model further extends the RW learning rule by accounting for the adaptation of the mice's sensitivity to the consequence of actions. We hypothesize that repeated failed escape attempts through swimming or struggling or after prolonged reliance on immobility lead to changes in perceiving consequence of actions. The adaptation process is described as:

$$\rho_t^{ada} = \begin{cases} \beta(lb - \rho_{t-1}) & swim/strggle \\ \beta(ub - \rho_{t-1}) & immobilize \end{cases}$$

Here, the adapted consequence sensitivity ρ_t^{ada} replaces the fixed parameter ρ from the previous model, with β representing the rate of adaptation. The lower and upper bounds of sensitivity adaptation are represented by parameters lb and ub, respectively. The prediction error is then updated as:

$$\delta_t = \begin{cases} \rho_t^{ada} - Q_{a,t} & swim/strggle \\ 0 - Q_{a,t} & immobilize \end{cases}$$

5. RW model with vigor (RW+vig): Vigor reflects the motivational drive to engage in active coping behaviors, which decreases due to cumulative fatigue from repeated swimming or struggling but recovers during periods of immobile.

$$vig_t = \begin{cases} vig_{t-1} + \xi(-vb - vig_{t-1}) & swim/strggle \\ vig_{t-1} + \xi(vb - vig_{t-1}) & immobilize \end{cases}$$

Where ξ refers to the rate of change in vigor. Vigor influences behavioral decisions jointly with value expectations:

$$p_{s,t} = \frac{1}{1 + e^{-\tau(Q_{s,t} - Q_{i,t} + vig_t)}}$$

6. Pearce-Hall (PH) Model: Unlike the Rescorla-Wagner rule, which assumes a constant belief updating rate, the Pearce-Hall model dynamically adjusts the belief updating rate based on the level of surprise or unpredictability. The value expectation is updated as follows:

$$Q_{a,t} = Q_{a,t-1} + \kappa a_{a,t} \delta$$

Here, $a_{a,t}$ represents the attention allocated to the action at time t, κ is a normalization parameter for learning rate. The attention $\alpha_{a,t}$ is updated based on the prediction error δ_t as follows:

$$a_{a,t+1} = (1-\eta)a_{a,t} + \eta|\delta_t|$$

Where η refers to the learning rate for attention.

7. PH model with perseveration (PH+per): Combines the PH learning rule with perseveration, as defined in the RW+per model.

8. PH model with consequence sensitivity (PH+ρ): Combines the PH learning rule with consequence sensitivity, as defined in the RW+ρ model.

9. PH Model with consequence sensitivity adaptation (PH+ada): Combines the PH learning rule with consequence sensitivity adaptation, as defined in the RW+ada model.

10. PH Model with vigor (PH+vig): Combines the PH learning rule with vigor, as defined in the RW+vig model.

Hierarchical bayesian model

We used a hierarchical Bayesian framework to model individual behavioral trajectories. This approach allowed us to estimate participant-specific parameters while simultaneously leveraging data from the entire experimental group to improve estimation precision. Specifically, individual parameters were drawn from group-level distributions as follows:

$$param_{individual} = \mu_{group} + \sigma_{group} * var_{individual}$$

Here, $param_{individual}$represents the parameter for each individual mouse, μ_{group} and σ_{group} are the group-level mean and standard deviation of the parameter, respectively, and $var_{individual}$ refers to the individual-level variance. This hierarchical structure effectively captures both individual differences and shared underlying processes, offering a robust and nuanced representation of the observed behavioral dynamics.

Parameter estimation and model selection/validation

Parameter estimation was conducted using hierarchical Bayesian analysis (HBA) implemented in the Stan language[45] via Python (cmdstanpy[46]). Posterior inference was performed through Markov chain Monte Carlo (MCMC) sampling. The models were fitted separately for each and each test. Model comparison was used to evaluate how well the models fit the data, accounting for model complexity. Model comparison was assessed using Bayesian bootstrap and model averaging, wherein the log-likelihoods of each model were evaluated based on the posterior simulations, and a weight was assigned to each model. The model weights included a penalization term for complexity and a normalization term relative to the number of models being compared. For each group, the model weights summed to 1. The model with the highest weight was defined as the "winning model," indicating that it best fit the observed behavioral data of the mice.

To perform a posterior predictive check, we simulated behavioral trajectories for each animal in each experimental paradigms and test, based on the parameters of the winning model. Specifically, the model generated random draws from each animal's joint posterior distribution to predict actions. This process was repeated 2,000 times per animal for each test. We compared each animal's synthetic data to their experimental observations, focusing on the immobility times both throughout the test and in the

final 4 minutes, the first immobile time, and the behavioral trajectories of mice.

Histology and microscopy

Following the final day of the behavioral experiments (Figure 5), mice were injected with 0.2 mL of a urethane solution at a concentration of 0.3 mg/mL to achieve anesthesia. Subsequently, they underwent transcardial perfusion with 4% paraformaldehyde in PBS for fixation. The brains were then post-fixed overnight, and coronal brain sections (20 μm) were prepared. Cell samples were imaged using a fluorescent microscope, and images were analyzed using ImageJ Pro Plus Version 6.0 (Bethesda, Maryland, USA).

Analysis of trends in research publications

To examine the research trends involving the FST and TST, the annual number of publications from 1995 to 2024 was collected from the Web of Science database. The analysis followed a two-step query process to ensure specificity to studies related to affective and stress-related disorders as well as the FST and TST.

The initial query used the following terms to identify publications related to affective and stress-related disorders:

TS = ((depression) OR (depressed) OR (depressive) OR (dysthymia) OR ("affective disorder") OR ("mood disorder") OR (anhedonia) OR (despair) OR (helplessness) OR (anxiety) OR (stress) OR ("social defeat")).

This search was refined by adding terms specific to the FST or TST. For the FST, the query was:

TS = ((FST) OR ("Forced Swimming Test") OR ("Forced Swim Test")).

For the TST, the query was:

TS = ((TST) OR ("tail suspension test") OR ("suspension test") OR (tail-hanging test)).

The results from each refined search were categorized by year of publication to analyze annual trends. The data were visualized as a bar graph to illustrate the temporal progression of research activity in the field.

Statistics

Comparative statistical tests, correlation analyses, and linear regressions were conducted using Python libraries, including SciPy[47] and scikit-learn[48]. Mann-Whitney U tests were conducted to compare differences in immobility times between groups (Figs. 3B-G, I-N) and differences in model parameters across groups (Fig. 5A-D). Pearson correlation analysis was used to examine the relationships between the posterior predictive mean of immobility time and the latency to immobility for each individual with their corresponding experimental observations (Figs. 4E-J, O-T). Additionally, Pearson correlation was applied to evaluate the agreement between immobility times identified by the automated recognition tool and those identified by human experts. Spearman correlation analysis was employed to assess the ranking consistency of immobility times identified by the automated tool and human experts. The linear regression models shown in Figs. 6G-J used model parameters as inputs to predict cumulative immobility times at each time step.

Reference

1. J, G., Lg, B. & M, G. Rodent models for Alzheimer disease. *Nature reviews. Neuroscience* **19**, (2018).

2. Powell, C. M. & Miyakawa, T. Schizophrenia-Relevant Behavioral Testing in Rodent Models: A Uniquely Human Disorder? *Biological Psychiatry* **59**, 1198–1207 (2006).

3. Ellenbroek, B. & Youn, J. Rodent models in neuroscience research: is it a rat race? *Disease Models & Mechanisms* **9**, 1079–1087 (2016).

4. Baker, M., Hong, S.-I., Kang, S. & Choi, D.-S. Rodent models for psychiatric disorders: problems and promises. *Laboratory Animal Research* **36**, 9 (2020).

5. Fenton, A. A. Excitation-Inhibition Discoordination in Rodent Models of Mental Disorders. *Biological Psychiatry* **77**, 1079–1088 (2015).

6. Makowska, I. J. & Weary, D. M. Assessing the emotions of laboratory rats. *Applied Animal Behaviour Science* **148**, 1–12 (2013).

7. Wong, A. H. C. & Josselyn, S. A. Caution When Diagnosing Your Mouse With Schizophrenia: The Use and Misuse of Model Animals for Understanding Psychiatric Disorders. *Biological Psychiatry* **79**, 32–38 (2016).

8. Shemesh, Y. & Chen, A. A paradigm shift in translational psychiatry through rodent neuroethology. *Mol Psychiatry* **28**, 993–1003 (2023).

9. Molendijk, M. L. & Kloet, E. R. de. Forced swim stressor: Trends in usage and mechanistic consideration. *European Journal of Neuroscience* **55**, 2813–2831 (2022).

10. Chen, L., Faas, G. C., Ferando, I. & Mody, I. Novel insights into the behavioral analysis of mice subjected to the forced-swim test. *Transl Psychiatry* **5**, e551–e551 (2015).

11. Reardon, S. Depression researchers rethink popular mouse swim tests. *Nature* **571**, 456–457 (2019).

12. Can, A. *et al.* The Tail Suspension Test. *J Vis Exp* 3769 (2012) doi:10.3791/3769.

13. Stukalin, Y., Lan, A. & Einat, H. Revisiting the validity of the mouse tail suspension test: Systematic review and meta-analysis of the effects of prototypic antidepressants. *Neuroscience & Biobehavioral Reviews* **112**, 39–47 (2020).

14. De Pablo, J. M., Parra, A., Segovia, S. & Guillamón, A. Learned immobility explains the behavior of rats in the forced swimming test. *Physiology & Behavior* **46**, 229–237 (1989).

15. Molendijk, M. L. & De Kloet, E. R. Immobility in the forced swim test is adaptive and does not reflect depression. *Psychoneuroendocrinology* **62**, 389–391 (2015).

16. Kloet, E. R. de & Molendijk, M. L. Coping with the Forced Swim Stressor: Towards Understanding an Adaptive Mechanism. *Neural Plasticity* **2016**, 6503162 (2016).

17. Molendijk, M. L. & de Kloet, E. R. Coping with the forced swim stressor: Current state-of-the-art. *Behavioural Brain Research* **364**, 1–10 (2019).

18. Anyan, J. & Amir, S. Too Depressed to Swim or Too Afraid to Stop? A Reinterpretation of the Forced Swim Test as a Measure of Anxiety-Like Behavior. *Neuropsychopharmacol.* **43**, 931–933 (2018).

19. Armario, A. The forced swim test: Historical, conceptual and methodological considerations and its relationship with individual behavioral traits. *Neuroscience & Biobehavioral Reviews* **128**, 74–86 (2021).

20. Crawley, D. *et al.* Modeling flexible behavior in childhood to adulthood shows age-dependent learning mechanisms and less optimal learning in autism in each age group. *PLOS Biology*

18, e3000908 (2020).

21. Brown, V. M. *et al.* Reinforcement Learning Disruptions in Individuals with Depression and Sensitivity to Symptom Change following Cognitive Behavioral Therapy. *JAMA Psychiatry* **78**, 1113–1122 (2021).

22. Dombrovski, A. Y., Szanto, K., Clark, L., Reynolds, C. F. & Siegle, G. J. Reward Signals, Attempted Suicide, and Impulsivity in Late-Life Depression. *JAMA Psychiatry* **70**, 1020–1030 (2013).

23. Halim, J. *et al.* Computational mechanisms underpinning greater exploratory behaviour in excess weight relative to healthy weight adolescents. *Appetite* **183**, 106484 (2023).

24. Homan, P. *et al.* Neural computations of threat in the aftermath of combat trauma. *Nature Neuroscience* **22**, 470–476 (2019).

25. Langley, C. *et al.* Chronic escitalopram in healthy volunteers has specific effects on reinforcement sensitivity: a double-blind, placebo-controlled semi-randomised study. *Neuropsychopharmacology 2023 48:4* **48**, 664–670 (2023).

26. Feher da Silva, C. & Hare, T. A. Humans primarily use model-based inference in the two-stage task. *Nat Hum Behav* **4**, 1053–1066 (2020).

27. Chen, C., Takahashi, T., Nakagawa, S., Inoue, T. & Kusumi, I. Reinforcement learning in depression: A review of computational research. *Neuroscience & Biobehavioral Reviews* **55**, 247–267 (2015).

28. Huys, Q. J. M., Maia, T. V. & Frank, M. J. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience* **19**, 404–413 (2016).

29. Lan, D. C. L. & Browning, M. What Can Reinforcement Learning Models of Dopamine and Serotonin Tell Us about the Action of Antidepressants? *Computational Psychiatry* **6**, 166–188

(2022).

30. Palminteri, S. & Lebreton, M. The computational roots of positivity and confirmation biases in reinforcement learning. *Trends in Cognitive Sciences* **26**, 607–621 (2022).

31. Kube, T. & Rozenkrantz, L. When Beliefs Face Reality: An Integrative Review of Belief Updating in Mental Health and Illness. *Perspect Psychol Sci* **16**, 247–274 (2021).

32. Katthagen, T., Fromm, S., Wieland, L. & Schlagenhauf, F. Models of Dynamic Belief Updating in Psychosis—A Review Across Different Computational Approaches. *Frontiers in Psychiatry* **13**, (2022).

33. Huys, Q. J., Pizzagalli, D. A., Bogdan, R. & Dayan, P. Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biology of Mood & Anxiety Disorders* **3**, 12 (2013).

34. Gradin, V. B. *et al.* Expected value and prediction error abnormalities in depression and schizophrenia. *Brain* **134**, 1751–1764 (2011).

35. Kumar, P. *et al.* Abnormal temporal difference reward-learning signals in major depression. *Brain* **131**, 2084–2093 (2008).

36. Kumar, P. *et al.* Impaired reward prediction error encoding and striatal-midbrain connectivity in depression. *Neuropsychopharmacology 2018 43:7* **43**, 1581–1588 (2018).

37. Pike, A. C. & Robinson, O. J. Reinforcement Learning in Patients With Mood and Anxiety Disorders vs Control Individuals: A Systematic Review and Meta-analysis. *JAMA Psychiatry* **79**, 313–322 (2022).

38. Bottemanne, H. *et al.* Evaluation of Early Ketamine Effects on Belief-Updating Biases in Patients With Treatment-Resistant Depression. *JAMA Psychiatry* **79**, 1124–1132 (2022).

39. Liao, C. & Kwan, A. C. Applying Reinforcement Learning to Rodent Stress Research. *Chronic Stress* **5**, (2021).

40. Kim, H. R. *et al.* A Unified Framework for Dopamine Signals across Timescales. *Cell* **183**, 1600-1616.e25 (2020).

41. Dezfouli, A., Morris, R., Ramos, F., Dayan, P. & Balleine, B. W. Integrated accounts of behavioral and neuroimaging data using flexible recurrent neural network models. *Advances in Neural Information Processing Systems* **2018-Decem**, 4228–4237 (2018).

42. Dabney, W. *et al.* A distributional code for value in dopamine-based reinforcement learning. *Nature* **577**, 671–675 (2020).

43. Verharen, J. P. H., de Jong, J. W., Zhu, Y. & Lammel, S. A computational analysis of mouse behavior in the sucrose preference test. *Nat Commun* **14**, 2419 (2023).

44. Sohal, V. S. & Rubenstein, J. L. R. Excitation-inhibition balance as a framework for investigating mechanisms in neuropsychiatric disorders. *Molecular Psychiatry* **24**, 1248–1257 (2019).

45. Documentation. *Stan* //mc-stan.org/users/documentation/.

46. cmdstanpy – Python interface to CmdStan — CmdStanPy 1.2.4 documentation. https://mc-stan.org/cmdstanpy/.

47. SciPy 1.0: fundamental algorithms for scientific computing in Python | Nature Methods. https://www.nature.com/articles/s41592-019-0686-2.

48. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 32300853) and the STI2030-Major Projects (No. 2022ZD0208200 and No. 2021ZD0200900).

Author contributions

Z.L., and Y.L. conceived the project. Z.L., T.L., J.Y., and X.Z. performed the original experiments. Z.L. developed the tool for automated behavioral recognition and computational models, designed and performed analysis. Z.L., T.L., and Y.L. discussed the results and wrote the manuscript.

Competing interests

The authors declare no competing interests.