# Unveiling Cognitive Processes Underlying Depression-like Behaviors in the Force Swim and Tail Suspension Tests

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, with immobility time typically serving as a primary indicator of depression. However, this approach oversimplifies the complexity of the underlying cognitive processes, leaving the connection between increased immobility and cognitive biases unclear. In this study, we utilized a custom-developed tool to extract the entire behavioral trajectory of mice, followed by application of reinforcement learning models for systematic analysis. This approach yielded novel insights into the underlying cognitive mechanisms. Our findings suggest that behaviors observed in both tests can be interpreted as reinforcement learning processes, involving both learning and decision-making. However, the cognitive processes differ between the two tests, challenging the assumption that the FST and TST are interchangeable for cross-validation. We also investigate two common methods for inducing depression-like behavior: chronic restraint stress (CRS) and inhibition of the medial prefrontal cortex (mPFC). Both methods increase immobility time by altering learning and sensitivity to consequences. Regression analyses of model parameters reveal two distinct phases of behavior. Early in the tests, learning-related factors dominate, while later stages are more influenced by sensitivity to consequences. This dynamic suggests that traditional metrics, which often focus on the final minutes of testing, may underestimate the significance of learning and overemphasize the influence of consequences sensitivity. Overall, our findings highlight the importance of computational modeling in revealing 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, rodents' inability to verbally express subjective experiences or perform 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]. These methods, shown in Fig. 1C and D, have seen a significant increase in usage over the past two decades and continue to be some of the most popular behavioral

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

assays. The primary measure in these tests is immobility duration, which is often interpreted as an indicator of depression-like behavior. However, this narrow focus on immobility overlooks the cognitive mechanisms that underlie these actions, raising questions about the true nature of what these tests measure. One promising explanation, the behavioral adaptation 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 key challenge in understanding these cognitive processes lies in the limited scope of current behavioral metrics. Most studies reduce complex behaviors, typically measured over 5-6 minutes, to a single measure—total immobility time in the final minutes of the test. This simplification obscures the underlying behavioral patterns that could offer deeper insights into cognitive mechanisms. As illustrated in Fig. 1G, animals with identical immobility times may exhibit vastly different behavioral patterns driven by distinct cognitive processes. Although some pioneering studies have analyzed high-resolution behavioral trajectoires[10], these approaches often rely on labor-intensive manual annotations, limiting their scalability and broader application. Consequently, there is a growing need for automated tools that can efficiently track and generate fine-grained behavioral data, facilitating progress in research.

In addition to improving data collection methods, another challenge is interpreting this data to uncover the cognitive mechanisms behind observed behaviors. Computational modeling has emerged as a versatile methodological framework for formalizing latent cognitive computations underlying behavioral patterns, with growing applications in clinical investigations of psychiatric disorders. This approach enables quantitative decomposition of behavior sequences across multiple experimental paradigms, including probabilistic reversal learning (PRL)[20–22], Iowa gambling task (IGT)[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 better understanding the disorder and developing effective treatments. Currently, there is a growing trend in animal studies to apply computational methods, such as reinforcement learning models, to analyze rodent behavior[39–43]. However, previous studies have focused primarily on operant learning tasks, while FST and TST, the most commonly used tests in studies of depression and anxiety, remain underexplored. The cognitive mechanisms driving depressive-like behaviors in the FST and TST, as well as the cognitive biases associated with these behaviors, remain poorly understood.

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 to exam implicit cognitive processes in rodent research. We collected behavioral trajectories of mice using a custom-built tool, then applied several reinforcement learning models to interpret these behaviors. Our analysis demonstrates that the behavior observed in the FST and TST can be explained by different reinforcement learning processes. Furthermore, our analysis of model parameters suggests that behaviors at different stages of the tests are influenced by distinct cognitive components. This finding implies that traditional analytical methods, which often exclude the early stages of the test, may introduce biases in assessing cognitive processes.

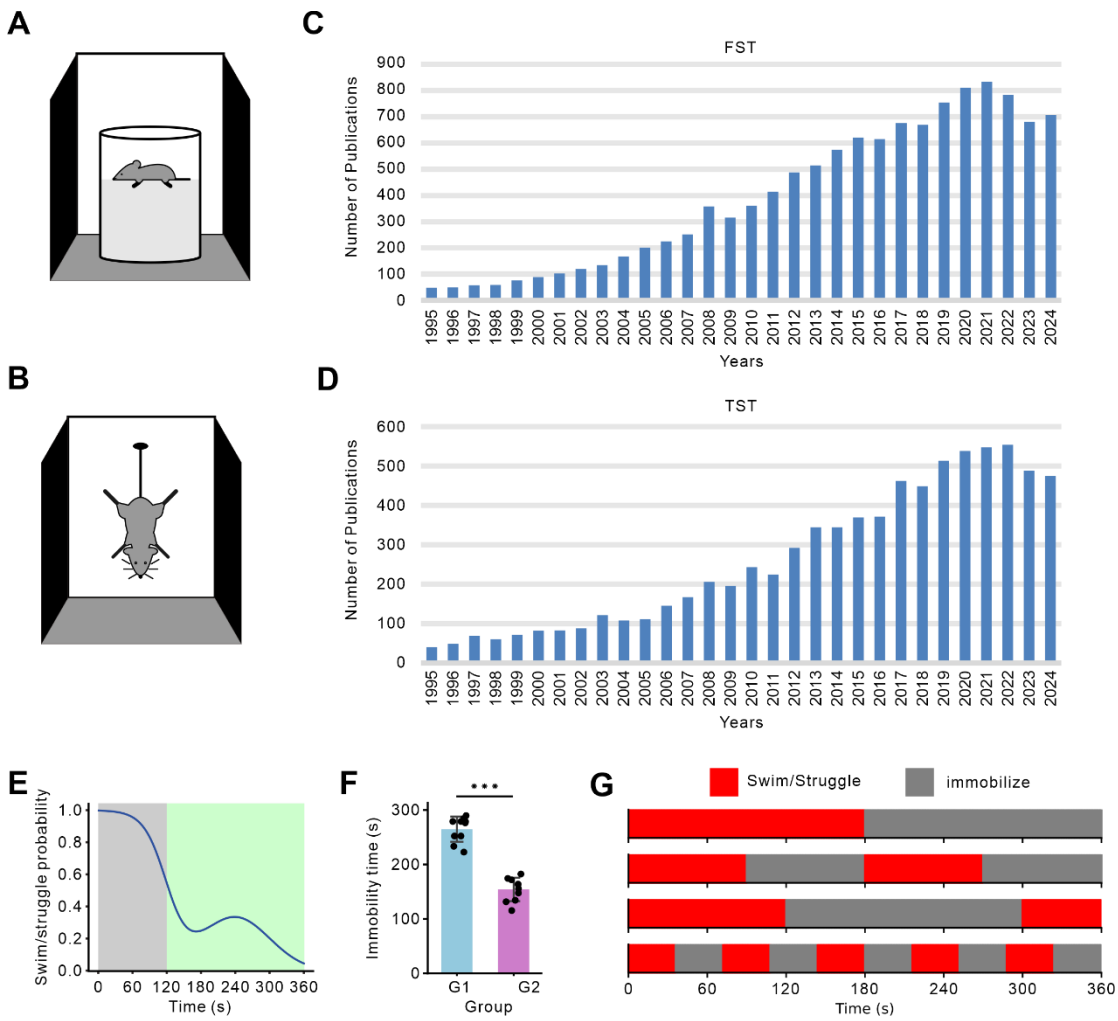


**Fig. 1** | The forced swim test (FST) and tail suspension test (TST). **A, B.** Experimental setup for the FST and TST. In the FST (A), mice are placed in a container of water, while in the TST (B), they are suspended upside down. Both tests subject the mice to inescapable conditions for 5 to 6 minutes. In the FST, behaviors are classified as swimming or immobility, whereas in the TST, behaviors are categorized as struggling or immobility. **C, D.** Trends in research using the FST and TST. Data from Web of Science over the past two decades shows a steady increase in the number of publications involving the FST (C) and TST (D), highlighting their growing use in behavioral research. **E.** Averaged behavioral trajectory. The illustration shows a typical decline in swimming or struggling behavior over time, with intermittent fluctuations observed in the later stages of the test. Traditional analyses often focus on the final minutes of the test (green background), neglecting the earlier phases (grey background). The data in this panel is artificially generated for demonstration purposes. **F.** Conventional methods for behavioral analysis. Traditional approaches summarize total immobility times by presenting group mean, standard error of the mean (SEM), individual data points, and statistical results. For demonstration, artificial data is used in this panel, with G1 and G2 representing two hypothetical experimental groups. **G.** Behavioral trajectories for individual mice. The cartoons highlight the complexity of behavior patterns, showing that even if a mouse remains immobile for half of the 6-minute test, distinct behavioral trajectories can 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. To analyze the underlying cognitive processes, we need detailed behavioral trajectory data. Manual annotation of these trajectories is time-consuming and prone to subjective criteria can affect both accuracy and consistency. To address these challenges, we developed a tool for automatic annotation of animal behavior in the FST and TST. As shown in Fig. 2A-D, the tool 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 calibration and ROI selection. 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 behavioral trajectory smoothing and binning 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

To induce depression-like behaviors, we employed two established experimental manipulations: chronic restraint stress (CRS) and chemogenetic inhibition of medial prefrontal cortex (mPFC) pyramidal neurons[44]. Comprehensive behavioral analyses were subsequently conducted to quantify immobility duration and struggle dynamics in the FST and TST under these controlled interventions. In the CRS experiment, 29 C57BL/6 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: 202.15 ± 30.00 vs. 137.13 ± 77.02s, mean ± std, $p = 0.024$; TST: 188.62 ± 59.89 vs. 130.50 ± 52.23, $p = 0.018$), and also showed higher immobility in the final 4 minutes of the FST (178.00 ± 20.96 vs. 118.75 ± 68.55, $p = 0.017$).

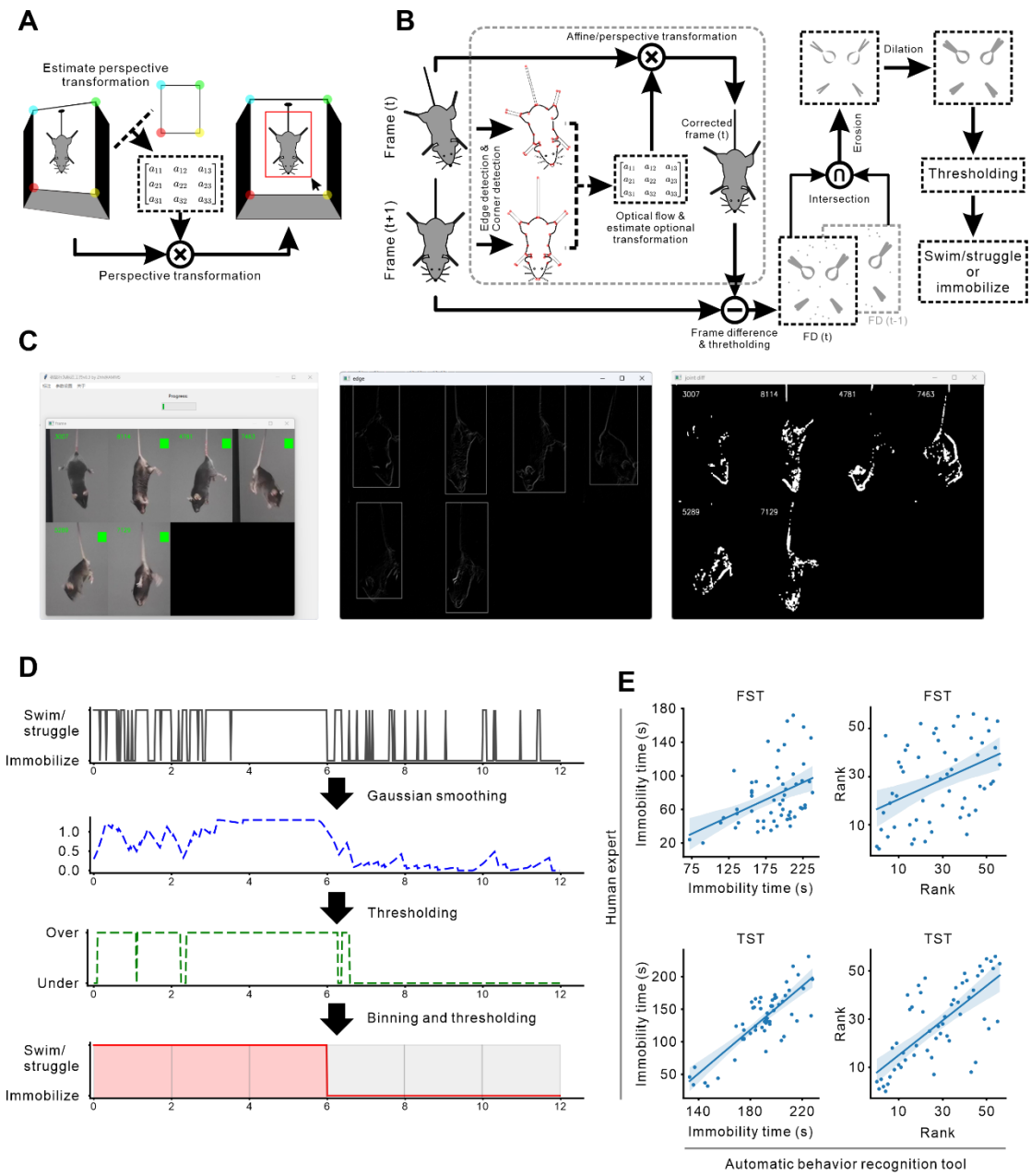


**Fig. 2** | Tool for automatic annotation of animal behavior in the FST and TST. **A.** Frame calibration and ROI selection. To improve behavior recognition accuracy, raw video frames undergo perspective transformation based on camera angles. The user manually selects the four vertices of the experimental box in a chosen video frame, which are then compared to the pre-measured box dimensions. This process generates a perspective transformation matrix that corrects the frames. The region of interest (ROI) is then defined as the maximum area within which the mouse can move. **B.** Workflow of the custom-developed tool. The tool identifies behavioral states by calculating the differences between consecutive video frames. An optional module removes passive motion (e.g., floating or swinging) using geometric transformations to enhance accuracy. Noise reduction is achieved through morphological operations, followed by trajectory refinement using smoothing, binning, and thresholding techniques. **C.** Tool interface. The user-friendly interface enables efficient interaction with the tool, providing real-time visualization of intermediate outputs from key processing steps. **D.** Smoothing and binning of behavioral trajectories. Raw behavioral trajectories are often noisy, as seen when a 6-second struggle is followed by

6 seconds of immobility, resulting in erratic data. Gaussian smoothing reduces this noise by assigns a behavioral value to each frame. Thresholding converts the smoothed trajectory into a binary profile, which is then binned into 2-second intervals to create a finalized, clear trajectory. **E.** Validation through expert comparison. Automated immobility times for both the FST and TST show strong correlations with expert manual annotations. Furthermore, the ranking of immobility times produced by the tool aligns closely with expert evaluations, confirming the tool's reliability across both tests.

We next examined how chemogenetic inhibition of mPFC pyramidal neurons affects behavior in the FST and TST. A total of 24 C57BL/6 mice were randomly assigned to either a control group (n=12, AAV-CaMKIIα-mCherry) or an experimental group (n=12, AAV-CaMKIIα-hM4D(Gi)-mCherry). Viral vectors were microinfused into the mPFC over 7 days (Fig. 3H), followed by a 4-week expression period. Behavioral assessments using the FST and TST were conducted on post-injection days 36 and 38, respectively, with clozapine-N-oxide (CNO) administration preceding each test to activate the inhibitory hM4D receptors. hM4D-mediated inhibition of mPFC neurons increased immobility duration compared to controls, particularly when excluding an outlier exhibiting persistent swimming behavior (Fig. 3I-L, throughout: 216.73 ± 40.05 vs. 161.00 ± 60.36, $p = 0.060$; final 4 minutes: 176.73 ± 35.59 vs. 127.33 ± 59.25, $p = 0.049$).

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. These dynamic patterns may reflect underlying cognitive processes in mice, establishing a data foundation for decoding cognitive mechanisms. Building upon this dataset, we propose to employ computational modeling approaches to elucidate the cognitive processes governing depression-like behavior.

## 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.

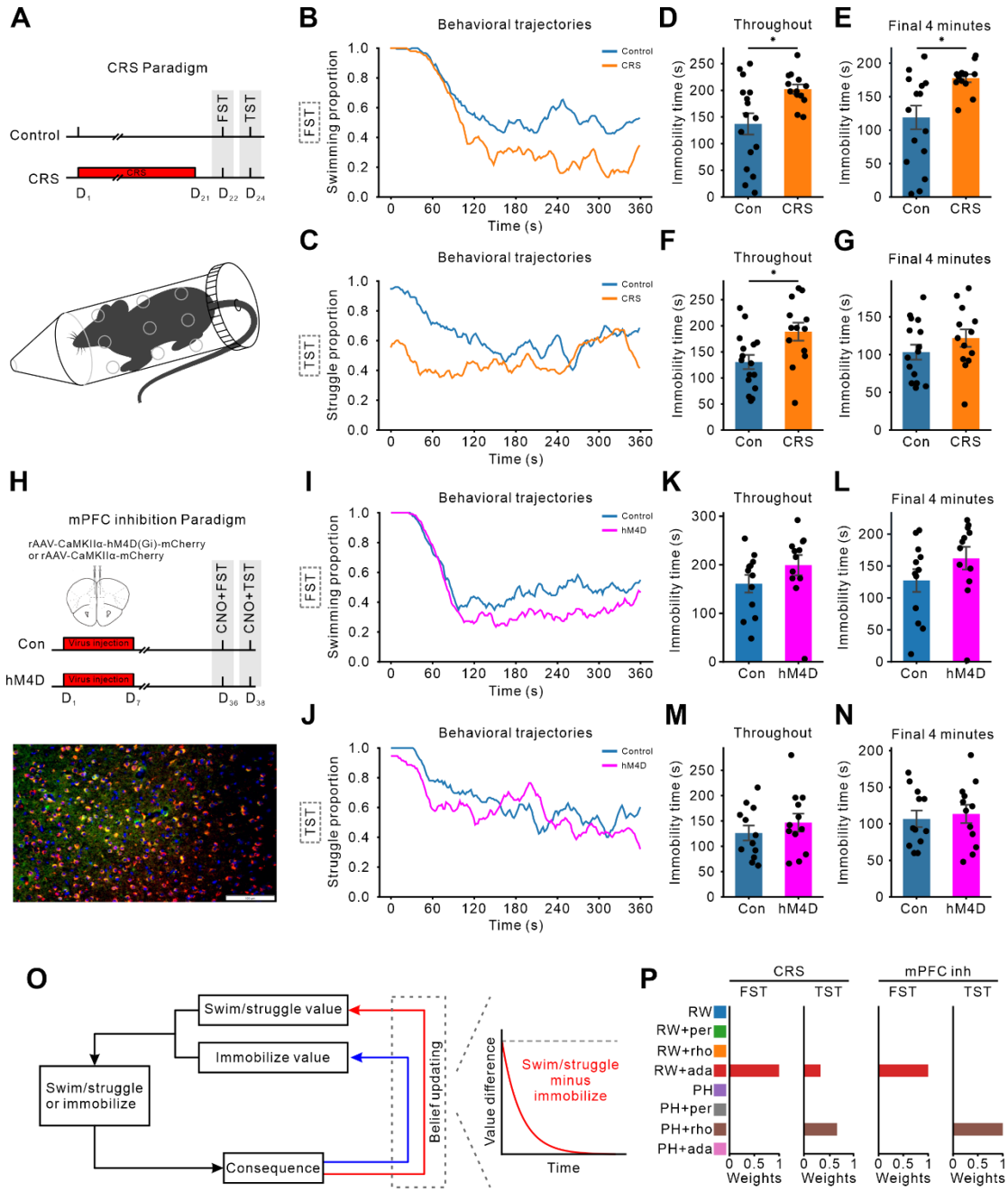


**Fig. 3** | Behavior analysis of mice in FST and TST under two experimental manipulations. **A.** The chronic restraint stress (CRS) protocol. Following a 7-day acclimation period, 29 mice were randomly assigned to either a CRS group (n=13) or a control group (n=16). The CRS group underwent daily restraint stress from days 1 to 21, while the control group remained unstressed. Behavioral testing were conducted on days 22 (FST) and 24 (TST) for both groups. **B, C.** Behavioral trajectories in the FST (CRS Experiment). Averaged behavioral trajectories for the FST are shown, with the control group represented in blue and the CRS group in yellow. **D-G.** Immobility duration in the FST and TST (CRS experiment). Immobility times for 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 indicate the CRS group, and black dots represent individual animal data. Error bars denote the standard error of the mean (SEM). **H.** Chemogenetic inhibition of mPFC pyramidal neuron. After a 7-day acclimation period, 24 mice were evenly divided into control and hM4D groups. From days 1 to 7, the hM4D group 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 for 28 days before behavioral testing, which was conducted on day 36 (FST) and day 38 (TST). Ninety minutes before testing, both groups received clozapine-N-oxide (CNO) injections, selectively inhibited pyramidal neuron activity in the hM4D group. Immunofluorescence confirmed hM4D(Gi)-mCherry expression in mPFC pyramidal neurons (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.** behavioral trajectories in the FST (mPFC inhibition experiment). Averaged behavioral trajectories for the FST are shown, with the control group in blue and the hM4D group in magenta. **K-N.** Immobility duration in the FST and TST (mPFC inhibition experiment). Immobility times for the FST (K, L) and TST (M, N) follow the same format as panels D-G. Blue bars indicate the control group, while 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 vs. immobility (FST) or struggling vs. immobility (TST). Action selection is driven by difference in expected values, which are continuously updated based on environmental feedback. Initially, active behaviors such as swimming or struggling are assigned higher values as escape strategies, but repeated failures reduce these values, leading to increased immobility over time. **P.** Model weights across tests and experimental conditions. The bar plot shows model weights for different conditions. With colors representing distinct models and bar lengths indicate their respective weights. The winning model, defined as the one with the highest weight, is RW+ada for the FSTs and PH+rho for the TSTs across both experimental manipulations. Asterisks denote significant differences from 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. These decisions will be continually re-evaluated and adjusted according to the consequences of previous actions, and may lead 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 eight 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, behavioral perseveration, and sensitivity adaptation. 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 weight of model indicates how likely the observed data is generated by this model. 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. The RW+ada model and PH+rho model emerged as the winning models in FSTs and TSTs, under both experimental manipulations, respectively. The RW+ada model updates beliefs with a fixed learning rate while dynamically adjusting sensitivity to the consequences of actions. In contrast, the PH+rho model, which dynamically alters the learning rate based on attention allocation determined by discrepancies between perceived and expected consequences, and incorporates consequence sensitivity, was the winning model in the TST. Notably, all winning models

can successfully recover their parameters (Fig. S2, r>0.52 p<0.0002, Pearson correlation), demonstrating their validity and reliability in capturing the behavioral data.

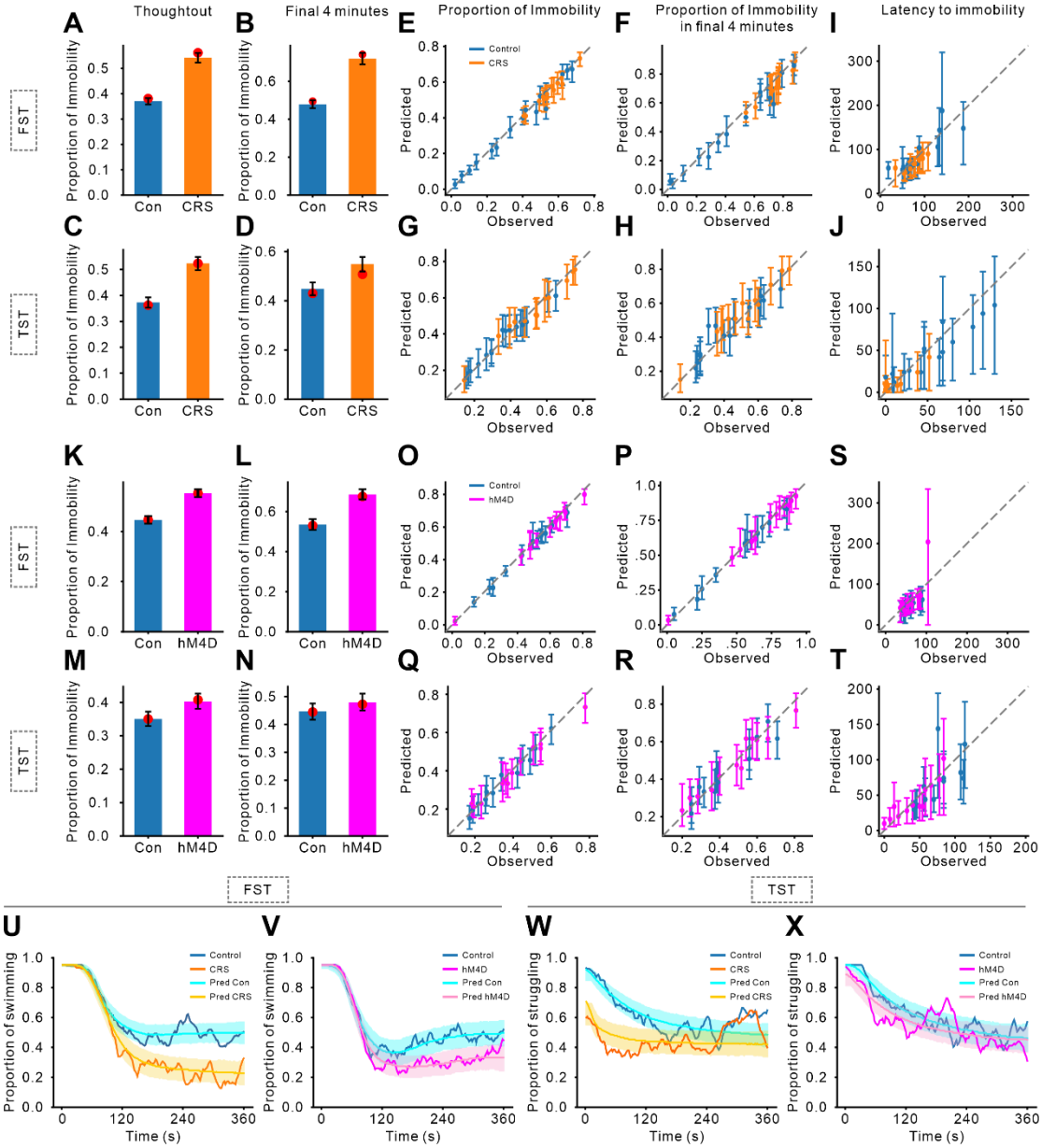


**Fig. 4** | Posterior predictive checks for the winning models. **A-D.** Model-predicted vs. observed immobility proportions (CRS experiment). Predicted immobility proportions for the CRS experiment are shown for the full testing period (A, C) and the final 4 minutes (B, D) in the FST (A, B) and TST (C, D). Colored bars indicate the median predictions from the winning models, with error bars showing the 95% highest density interval (HDI). Red dots mark the observed experimental averages, demonstrating strong alignment between predictions and actual data. **E-H.** Individual-level predicted vs. observed immobility proportions (CRS experiment). Comparison of predicted and observed immobility proportions for individual mice across the full test (E, G) and the final 4 minutes (F, H) in the FST (E, F) and TST (G, H). The x-axis represents observed values, while the y-axis shows median model predictions based on 2000 simulations. Error bars denote the 95% HDI, with colors distinguishing experimental groups. **I, J.** Predicted vs. observed latency to immobility (CRS experiment). Comparison of model-predicted and

observed latencies to immobility in the FST (I) and TST (J), formatted identically to panels E-H. **K-N.** Model-predicted vs. observed immobility proportions (mPFC inhibition experiment). Analogous to panels A-D, these panels present the results for the mPFC inhibition experiment. **O-R.** Individual-level predicted vs. observed immobility proportions. Analogous to panels E-H, showing the results for the mPFC inhibition experiment. **S, T.** Individual-level predicted vs. observed latency to immobility (mPFC inhibition experiment). Analogous to panels S and T, showing the results for the mPFC inhibition experiment. **U-X.** Behavioral trajectories predictions. Predicted behavioral trajectories over the full testing period. Panels U and V show swimming behavior predictions for the FST, while W and X display struggling behavior predictions for the TST across experimental groups. Cyan, yellow, and pink curves represent median predictions, with shaded areas indicating the 95% HDI based on 2000 simulations. Observed group averages are shown as blue, orange, and magenta lines. The models effectively capture behavioral dynamics, as most observed values fall within the 95% HDI.

## 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 inhibition on model parameters

The model parameters for each experimental group are presented in Fig. 5A-D. In the FST, both the CRS and hM4D groups demonstrated significant reductions in the learning rate ($\alpha$) and the lower bound of adaptation ($lb$) when compared to the control group. Additionally, the consequence sensitivity adaptation rate ($\beta$) was significantly increased in these groups. The hM4D group also showed a higher inverse temperature ($\tau$), indicating greater consistency in the agent's decisions. In the TST, both the CRS and hM4D groups had lower $\tau$ values than the control group. Furthermore, the CRS group exhibited higher learning rate normalization ($\kappa$) and attention updating rate ($\eta$), along with reduced sensitivity to consequences ($\rho$).

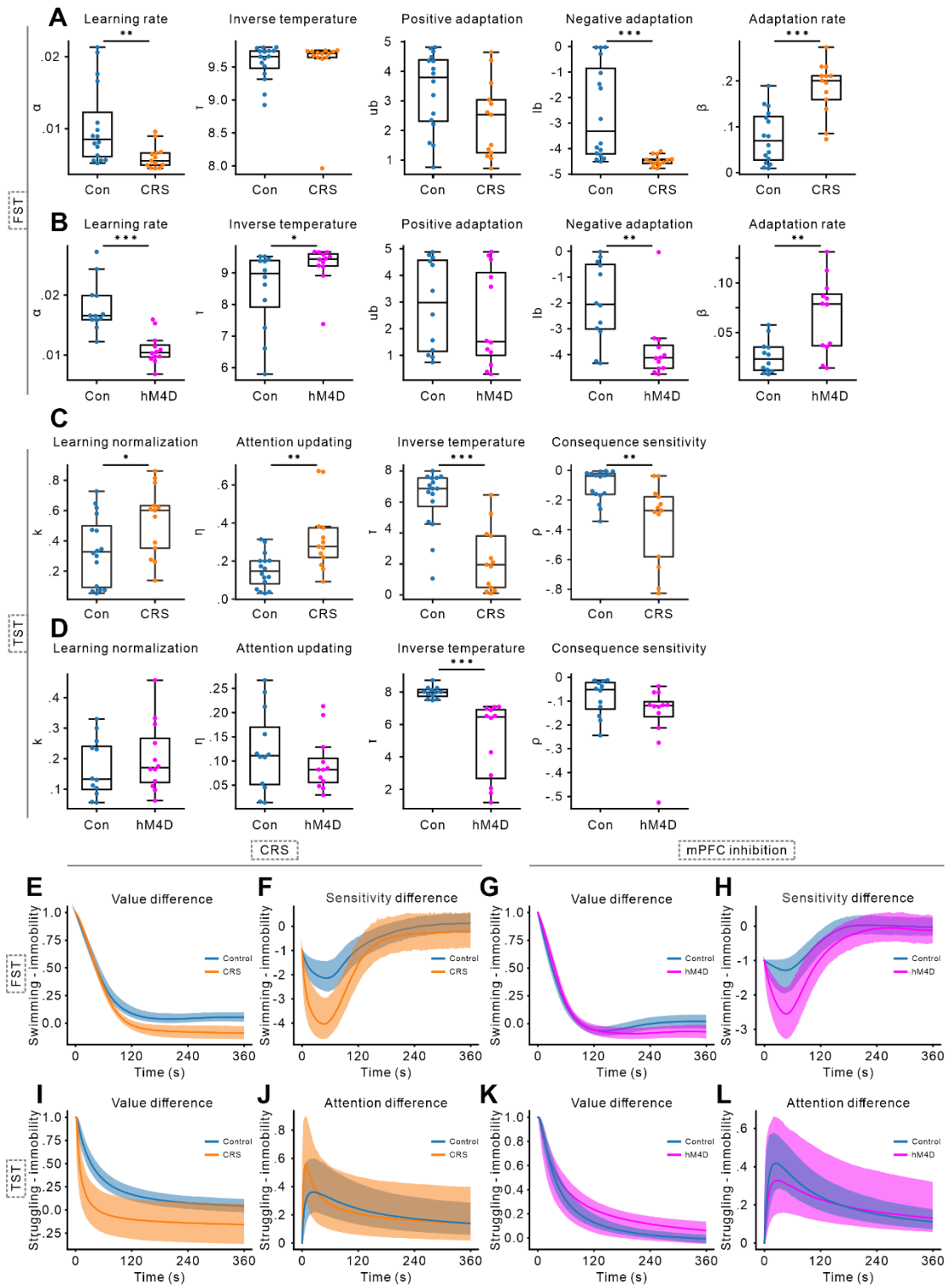


**Fig. 5** | Comparison of model parameters. **A, B.** Parameter distributions for the winning model (RW+ada) in the FST. The RW+ada model for the FST includes five parameters: the learning rate ($\alpha$), the inverse temperature parameter ($\tau$), which reflects how strongly decision-making is influenced by differences in expected values, and two adaptation bound parameters-$ub$, which accounts for positive adaptation by reducing the negative perception of escape failure after prolonged immobility, and $lb$, which represents negative adaptation by intensifying the negative experience following repeated escape failures during continuous swimming, the adaptation rate ($\beta$) governs the speed of these adjustments. Each panel displays

one parameter, with each point representing the mean of 2,000 samples from an individual mouse. Boxplots illustrate the group-level medians and interquartile ranges. With blue, orange, and magenta representing the control, CRS, and hM4D groups, respectively. **C, D.** Parameter distributions for the winning model (PH+rho) in the TST. Following the same format as panels A and B, these panels present the parameter distributions for the PH+rho model in the TST. This model consists of four parameters: $\kappa$, which normalizes the learning rate; $\eta$, which determines the rate of attention updates; the inverse temperature parameter ($\tau$), which functions similarly to its role in the RW-ada model; and $\rho$, which quantifies the intensity of negative experiences following an escape failure. **E-H.** Predicted dynamics of expected value and consequence sensitivity in the FST. Panels E and F show the predicted differences in expected value (E) and adapted sensitivity to consequences (F) between swimming and immobility throughout the FST in the CRS experiment. Panels G and H display the same predictions for the mPFC inhibition experiment. Solid lines represent group-level medians, while shaded areas indicate the 95% HDI based on 2,000 posterior samples. **I-L.** Predicted dynamics of expected value and attention allocation in the TST. Panels I and J present the predicted differences in expected value (I) and attention allocation (J) between struggling and immobility in the CRS experiment, while panels K and L show the same predictions for the mPFC inhibition experiment. As in the FST panels, solid lines indicate group-level medians, with shaded areas representing the 95% HDI. Asterisks indicate significant differences compared to the control groups, as determined by the Mann–Whitney U test; ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

We further examined intermediate variables such as the expected value, adapted consequence sensitivity, and attention allocation over the course of the tests (Fig. 5E-L). In general, the differences in expected value between swimming/struggling and immobility decreased over time, reaching a stable, lower level. In the FST, both the CRS and hM4D groups showed lower value differences at later stages compared to the control group. In the TST, the CRS group exhibited a more rapid decline in expected value differences in the early stages, while there were no significant differences between the hM4D and control groups throughout the test. For consequence sensitivity, both the CRS and hM4D groups showed lower sensitivity to consequences in the early stages of the FST. In terms of attention allocation, the CRS group allocated slightly more attention to struggling behavior early in the TST compared to the control group.

The decrease in $lb$ suggests that CRS and mPFC inhibition in the FST may lead to heightened feelings of despair in the mice after failed escape attempts. The increase in $\beta$ likely accelerated this sense of despair, as illustrated by the yellow and magenta curves in Fig. 5F and H, which consistently fall below the control group's (blue) curve, showing a steeper decline in the early stages. This shift resulted in a lower expected value for swimming (Fig. 5E and G), ultimately contributing to a higher incidence of immobility. In the TST, the CRS group's early increase in $\eta$ suggests that these mice more quickly 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 $\kappa$, this led to quicker learning of the negative consequences associated with struggling (Fig. 5I, where the yellow line declines more rapidly), ultimately resulting in a faster onset of immobility.

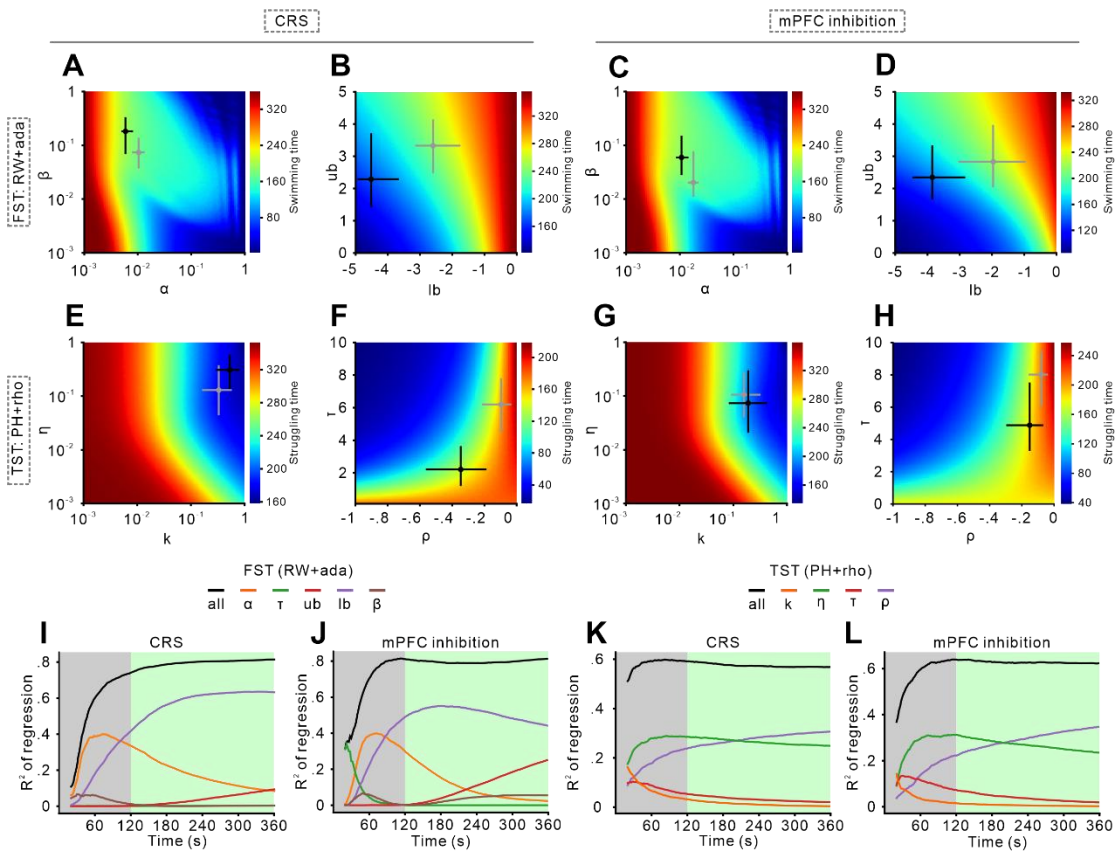


**Fig. 6** | Linking model parameters to behavioral patterns. **A-D.** Simulated swimming durations in the RW+ada model. Heatmaps depict simulated total swimming times across the parameter space of the RW+ada models. Each heatmap explores the effects of two selected parameters while holding others fixed at their posterior median values. The x- and y-axes represent the sampled parameter values, with brighter colors indicating longer swimming durations. Circles mark the posterior median estimates, while horizontal and vertical lines indicate the 95% HDI. Black markers represent the CRS or hM4D groups, and gray markers indicate the control group. Panels A and B show results for the CRS experiment, while panels C and D correspond to the mPFC inhibition experiment. **E-H.** Simulated struggling durations in the PH+rho model. Similar to panels A-D, these heatmaps illustrate simulated struggling behavior in the PH+rho model. Panels E and F display results for the CRS experiment, while panels G and H correspond to the mPFC inhibition experiment. **I-L.** Explanatory power of model parameters for behaviors. Panels I-L present the explanatory power ($r^2$) of individual and combined model parameters from linear regression analyses predicting cumulative swimming (I, J) and struggling (K, L) times in the FST and TST. Black curves represent $r^2$ values for models incorporating all parameters, while colored curves show the explanatory power of individual parameters. Panels I and J correspond to the CRS experiment, while panels K and L display results for the mPFC inhibition experiment. Parameter contributions evolved over time: learning rate-related parameters ($\alpha$, $\eta$) were most influential in the early phases, whereas parameters tied to consequence sensitivity ($lb$, $ub$, $\rho$) gained importance in later stages.

## 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 FST and TST were shown in Fig. 6 A-D and E-H, respectively. In the FST, we found that higher values of both the lower bound ($lb$) and upper bound ($ub$) of consequence sensitivity adaptation were associated with longer swimming durations. The relationship between $\alpha$ and $\beta$ was more complex: very low $\alpha$ values prolonged swimming, while moderate $\beta$ values, when combined with higher $\alpha$, promoted swimming behavior. In the TST, longer struggling durations were linked to higher values of $\rho$, along with lower values of $\kappa$, $\eta$ and $\tau$.

To further 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 ($r^2$). This analysis revealed that different parameters exerted varying degrees of influence at different stages of the experiment. In the early stages of the FST, $\alpha$ from the RW+ada model was the strongest predictor of cumulative swimming time. However, as the experiment progressed, the influence of $\alpha$ diminished, while the bounds of consequence sensitivity adaptation ($lb$ and $ub$) became more significant (Fig. 6I, J). In the TST, $\eta$ from the PH+rho model was the dominant factor early in the test, while $\rho$ gained importance in the later stages (Fig. 6K, L).

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 when interpreting animal behavior.

# Discussion

In this study, we introduced a custom-developed tool for analyzing animal behavior, enabling the efficient collection of detailed behavioral trajectories in the FST and TST. Our computational modeling analysis revealed that both behaviors in these tasks are driven by reinforcement learning processes involving learning and decision-making. However, the underlying cognitive processes in the FST and TST differ, challenging the common assumption that these tasks are equivalent or mutually validating. By analyzing model parameters that represent different cognitive components, we found that CRS and mPFC inhibition primarily increased immobility in both tests by altering the learning rates and sensitivity of consequence. Additionally, our analysis showed that different cognitive components had varying impacts at different stages of the experiment. This suggests that traditional methods, which often exclude early-phase data, may underestimate the influence of learning-related factors while overemphasizing sensitivity to consequence. Overall, our findings emphasize that behavioral trajectories are a rich source of information, which, when carefully investigated through computational modeling, can reveal the complex cognitive processes driving behavior.

## The necessity of using fine-grained behavioral trajectories

The FST and TST are commonly used to study depression-like behavior in rodents. As research into depression deepens, there is increasing interest in the cognitive insights these tests can offer. However, traditional metrics, which typically focus on immobility time in the final minutes as a single measure, fail to capture the complexity and evolution of behavior throughout the test. This limitation has hindered a comprehensive understanding of the cognitive processes involved in the FST and TST, highlighting the

need for fine-grained behavioral trajectory data.

Currently, the identification of behaviors in the FST and TST relies heavily on labor-intensive manual annotation, often requiring second-by-second or frame-by-frame analysis. This process is not only time-consuming but also limits the broader adoption of behavioral trajectory analysis in research. To address these challenges, we developed a user-friendly tool for automatically tracking behavioral trajectories in the FST and TST. This Python-based tool provides frame-level temporal resolution, allowing researchers to identify animal action states with high precision. By using this tool to capture detailed behavioral trajectories, researchers can detect subtle changes in behavior, apply computational models, and gain deeper insights into the cognitive processes underlying animal behavior.

## Using computational models to analyze behaviors in FST and TST

This study represents a pioneering effort to analyze mouse behavior in the FST and TST using reinforcement learning models, offering new insights into the cognitive processes at play. We framed the continuous decision-makings between swimming/struggling and immobility as reinforcement learning processes, and applying the classic Rescorla-Wagner (RW) and Pearce-Hall (PH) models, along with their variants, to model behavior.

While both RW and PH models predict that the expected values of actions gradually converge with perceived consequences over time, our observations of behavioral trajectories in the FST revealed a more dynamic pattern. Rather than stabilizing in the later stages, mouse behavior exhibited rebounds and fluctuations, with swimming behaviors resurfacing after extended periods of immobility. To better capture this complexity, we incorporated the adaptation of consequence sensitivity into our models.

Consequence adaptation consists of two components: positive adaptation, which alleviates the negative subjective experiences associated with failed escape attempts after prolonged immobility, and negative adaptation, which amplifies the negative experiences resulting from repeated failures in swimming or struggling. Models that incorporated this adaptive factor emerged as the winning models for the FST, accurately predicting immobility durations and capturing the dynamic evolution of behavior. This highlights the effectiveness of integrating adaptive processes into computational models for simulating behavior in this task.

## Behavior in the FST and TST can be interpreted as reinforcement learning processes

Despite their widespread use, the cognitive processes assessed by the FST and TST remain insufficiently understood. Although the adaptive learning hypothesis is promising, it has not been rigorously supported with quantitative evidence. In this study, reinforcement learning models effectively captured the evolution of mouse behavior, predicting behavioral metrics both at the individual and group levels, as well as the behavioral trajectories in the FST and TST. These findings suggest that behavior in these tasks can indeed be interpreted as reinforcement learning processes, involving both learning and decision-making.

Specifically, when initially placed in an uncomfortable environment, mice evaluate swimming and struggling as potential escape strategies, assigning higher expected value to these actions. As a result, these actions are more frequently chosen in the early stages of the test. However, repeated failed escape

leads to prediction errors—discrepancies between the expected value of these actions and the actual consequences. These prediction errors drive a continuous reassessment of the expected values of each behavior. Over time, the expected value of swimming and struggling diminishes, potentially falling below that of immobility, which results in the increased immobility observed in the later phases of the experiment.

## FST and TST exam different cognitive processes

Although these tests share similar methodologies, our findings indicate that they engage different cognitive processes. The models that best fit the behavioral data for each task differ, suggesting that mice rely on different cognitive mechanisms in each task. In the FST, the RW+ada model provides the most accurate representation of behavior, implying that mice dynamically adjust their subjective experiences based on previous actions and consequences, updating their expected value at a fixed rate. Conversely, in the TST, the PH+rho model best captures the observed behavior, suggesting that mice alter their rate of learning based on attention allocation, which is influenced by discrepancies between expected and perceived consequences, while maintaining a stable perception of those consequences.

These results challenge the common practice of using the FST and TST interchangeably for cross-validation, as they underscore the differences in the cognitive processes driving behavior in each task. The FST appears to emphasize how animals dynamically perceive and adapt to the consequences of their actions, while the TST may be more focused on how attention modulates learning processes. This distinction may help explain the inconsistent results often observed when using these tests to evaluate depression-like behavior. Our findings point to the need for a more nuanced understanding of the cognitive processes measured by these behavioral assays, and we hope that future research will further elucidate these processes and develop more appropriate methods for cross-validation.

## Limitation of traditional methods: Overlooking learning and overemphasizing consequence sensitivity

Traditional metrics in the FST and TST often focus primarily on the final minutes of the tests, a practice that can lead to incomplete or potentially misleading conclusions. Our findings suggest that this approach may fail to capture the full complexity of behavior. Using computational models, we observed that, in the FST, the learning rate was the most influential factor in shaping behavior during the early stages. However, as the experiment progressed, its influence diminished, and the bounds of consequence sensitivity adaptation became more prominent. A similar pattern was observed in the TST. In the early phases, the rate of attention updating had the greatest influence, but as time went on, the role of consequence sensitivity ($\rho$) became more dominant.

These shifts highlight the critical role of learning in the early stages of the tests, with consequence sensitivity taking a more prominent role in later phases. By disregarding the first two minutes of the experiment—a common practice—traditional methods risks underestimating the importance of learning, while overemphasizing the roles of consequence sensitivity. This oversight may limit our understanding of depressive-like behaviors and hinder the development of effective antidepressant treatments. To address this issue, we recommend that future studies focus on analyzing the full behavioral trajectory, utilizing computational models to gain a more comprehensive understanding of the cognitive processes

driving behavior.

# 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.

~~frame calibration and region of interest (ROI) selection, passive motion elimination, active motion detection, and.~~

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. **Behavioral Trajectory Smoothing and Binning 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 8 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, $\delta_t$ 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. $\phi$ 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 $\rho$ .

**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 $\rho_t^{ada}$ replaces the fixed parameter ρ from the previous model, with $\beta$ 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. 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, $\kappa$ is a normalization parameter for learning rate. The attention $\alpha_{a,t}$ is updated based on the prediction error $\delta_t$ as follows:

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

Where $\eta$ refers to the updating rate for attention.

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

**7. PH model with consequence sensitivity (PH+ρ)**: Combines the PH learning rule with

consequence sensitivity, as defined in the RW+ρ model.

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

## Hierarchical ~~bayesian~~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, $\mu_{group}$ and $\sigma_{group}$ 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. The Gelman-Rubin convergence diagnostic (R-hat) was applied to assess MCMC stability. With all R-hat statistics falling within the recommended threshold of 1.05, the chains exhibited sufficient mixing and stationarity, supporting the reliability of posterior estimates.

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 manipulations 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.