Genetic changes linked to two different syndromic forms of autism enhance reinforcement learning in adolescent male but not female mice

Juliana Chase, PhD[1*], Jing-Jing Li[1], Wan Chen Lin, PhD[1], Lung-Hao Tai, PhD[1], Fernanda Castro[2,3], Anne GE Collins, PhD[3,4], Linda Wilbrecht, PhD[1,3]

1. Department of Neuroscience, University of California, Berkeley, Berkeley, CA, 94720
2. Current address: Cellular & Molecular Pharmacology, University of California, San Francisco, Mission Bay, CA 94143
3. Department of Psychology, University of California, Berkeley, Berkeley, CA, 94720
4. Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA, 94720

Email: Linda Wilbrecht, PhD, wilbrecht@berkeley.edu

Abstract

Autism Spectrum Disorder (ASD) is characterized by restricted and repetitive behaviors and social differences, both of which may manifest, in part, from underlying differences in corticostriatal circuits and reinforcement learning. Here, we investigated reinforcement learning in mice with mutations in either *Tsc2* or *Shank3*, both high-confidence ASD risk genes associated with major syndromic forms of ASD. Using an odor-based two-alternative forced choice (2AFC) task, we tested adolescent mice of both sexes and found male *Tsc2* and *Shank3B* heterozygote (Het) mice showed enhanced learning performance compared to their wild type (WT) siblings. No gain of function was observed in females. Using a novel reinforcement learning (RL) based computational model to infer learning rate as well as policy-level task engagement and disengagement, we found that the gain of function in males was driven by an enhanced positive learning rate in both *Tsc2* and *Shank3B* Het mice. The gain of function in Het males was absent when mice were trained with a probabilistic reward schedule. These findings in two ASD mouse models reveal a convergent learning phenotype that shows similar sensitivity to sex and environmental uncertainty. These data can inform our understanding of both strengths and challenges associated with autism, while providing further evidence that sex and experience of uncertainty modulate autism-related phenotypes.

Significance Statement

Reinforcement learning is a foundational form of learning that is widely used in behavioral interventions for autism. Here, we measured reinforcement learning in adolescent mice carrying genetic mutations linked to two different syndromic forms of autism. We found that males showed strengths in reinforcement learning compared to their wild type siblings, while females showed no differences. This gain of function in males was no longer observed when uncertainty was introduced into the reward schedule for correct choices. These findings support a model in which diverse genetic changes interact with sex to generate common phenotypes underlying autism. Our data further support the idea that autism risk genes may produce strengths as well as challenges in behavioral function.

Main Text

Introduction

Autism Spectrum Disorder (ASD) is a common neurodevelopmental condition that is characterized by restricted and repetitive behaviors in addition to differences in social communication. Autism is genetically heterogeneous, but it is thought that mutations to single genes can cause disruptions in brain development, leading to the condition. These genetic differences can be introduced into mouse lines for study (Tsai et al., 2012; Peça et al., 2011).

Restricted and repetitive behaviors observed in autism are likely related to differences in learning and corticostriatal circuit function (Fuccillo, 2016), differences which may also influence social learning. Discrete regions of the striatum are thought to support both cognitive and motor learning with separate cortical, limbic, and neuromodulatory inputs controlling different aspects of action generation, cue and action evaluation, and reinforcement (Cox & Witten, 2019; Peters et al., 2021). Underscoring the striatum's critical role in ASD, studies consistently link changes in striatal development and function with ASD symptoms and severity (Qiu et al., 2016; Fuccillo, 2016; Rothwell, 2016; Comparan-Meza et al., 2021; Dölen & Malenka, 2014). Mutations to autism risk genes have also been shown to alter corticostriatal circuit function and structure (Rothwell et al., 2014; Platt et al., 2017; Le Merrer et al., 2023; Cording & Bateup, 2023). Notably, many mouse models with mutations to ASD risk genes show enhanced motor learning on the rotarod task, a well-established assay of striatal function (Cording & Bateup, 2023). However, whether such gains extend to reinforcement learning tasks, which also depends on the striatum (Peters et al., 2021; Reinhold et al., 2023), remains unclear.

Here we focus on learning in two lines of mice, *Tsc2*$^{+/-}$ and *Shank3B*$^{+/-}$, that have been engineered to mimic genetic differences found in two syndromic forms of autism: Tuberous Sclerosis Complex (TSC) and Phelan-McDermid Syndrome, respectively. Syndromic forms of autism, including TSC and Phelan-McDermid Syndrome, account for around 1% of *de novo* genetic causes of ASD (de la Torre-Ubieta et al., 2016). TSC is defined by a mutation to either the *TSC1* or *TSC2* gene. People with either TSC mutation often exhibit brain lesions called tubers, epilepsy, and cognitive deficits of varying degrees, and, in 25-60% of cases, they also receive an ASD diagnosis (Gipson et al., 2019; Boronat, Thiele, & Caruso, 2017; Baker et al., 1998; Capal et al., 2021). The protein products of the *TSC1* and *TSC2* genes, hamartin and tuberin respectively, form a protein complex that negatively regulates the mammalian target of rapamycin complex 1 (mTORC1) (Huang & Manning, 2008). mTORC1 is a central and conserved intracellular signaling pathway that coordinates essential cellular processes such as autophagy, cellular metabolism, and protein and lipid synthesis (Laplante & Sabatini, 2012). Phelan-McDermid Syndrome is driven by either a deletion of the terminal end of chromosome 22 or a mutation to *SHANK3* (Costales & Kolevzon, 2015). It is characterized by intellectual disability, neuromuscular effects, and is accompanied by an autism diagnosis in approximately 85% of cases (Durand et al., 2006) which account for up to 2% of all ASD cases (Leblond et al., 2014). The protein encoded by *SHANK3,* SH3 and multiple ankyrin repeat domains 3 (SHANK3), is one of a family of proteins that bind directly with synapse-associated protein 90/postsynaptic density 95-associated protein (SAPAP) and postsynaptic density protein (PSD) 95 (Naisbitt et al., 1999; Kim & Sheng, 2004). Together they form the PSD-95-SAPAP-SHANK complex and play an integral role in synaptic formation and development (Sala et al., 2001; Baron et al., 2006; Peça et al., 2011). SHANK3 is predominantly found in the striatum and is localized postsynaptically at glutamatergic synapses onto spiny projection neurons (SPNs) (Peça et al., 2011), synapses known to be important for learning (Hwang et al., 2022). Mice with a deletion of the PDZ region of Shank3*,* called *Shank3B*$^{-/-}$, conditionally knocked out of indirect pathway SPNs (iSPNs) substantially increased grooming, a mark of innate repetitive behavior in rodents (Wang et

al., 2017). Importantly, haploinsufficiency of *SHANK3* in Phelan-McDermid patients has been shown to be responsible for the neurological symptoms associated with the syndrome (Durand et al., 2006).

TSC1/2 and SHANK3 have distinct cellular functions, yet they may have a convergent impact on the development of circuits, including corticostriatal circuits which are implicated in cognitive and motor learning (Thorn et al., 2010; Koralek et al., 2012; Fisher et al., 2020; Peters et al., 2021; Reinhold et al., 2023; Cording and Bateup, 2023). Importantly, *Tsc1* and *Shank3* manipulations in mice have both been shown to enhance cortical connectivity onto striatal spiny projection neurons (SPNs) (Benthall et al., 2018; Peixoto et al., 2016;2019; Benthall et al., 2021). In *Shank3B*$^{-/-}$ mice, hyperconnectivity was greater earlier in development and transitioned to hypoconnectivity in adulthood (Peixoto et al., 2019). This convergent circuit phenotype could plausibly lead to convergent behavioral differences, particularly in diverse forms of learning supported by striatal functions. Therefore, we chose to examine reinforcement learning in adolescent *Tsc2* and *Shank3B* mice in the first days of learning about the correct actions to perform in response to specific odor cues.

Our goal was to test whether adolescent mice, haploinsufficient for *Tsc2* or *Shank3B,* would show a gain or loss of function in an odor-based two alternative forced choice (2AFC) task where they were trained to associate an odor with a correct action to receive a water reward. Based on previous rotarod data, we may expect *Tsc2* Het animals to show strong motor coordination, but whether reinforcement learning is stronger or weaker than WT was unknown.

We chose to focus on initial stages of learning in developing mice (starting odor learning around postnatal day (P)30) to enhance translational value. In humans, development is a prime period for intervention and the numbers of trials for training are more limited Yet, in animal studies, rodents with altered autism risk genes are typically tested in adulthood and examined after extensive overtraining.

We also chose to study both male and female mice and to separately analyze results in each sex. Gene expression unfolds within a larger context of the developing body, affecting cellular and behavioral phenotypes in dynamic ways, particularly in adolescence when puberty is beginning (Ferri et al., 2018; Delevich et al., 2021; Larsen and Luna, 2018). There has historically been a male bias in autism diagnoses (Baio et al., 2018), and male biased sex differences are also observed in mice with mutations to ASD risk genes (Ferri et al., 2018; Grissom et al., 2018). However, in the past few decades, the male:female ratio is decreasing in human diagnosis (Hull, Petrides, & Mandy, 2020) and there is growing awareness of a need to understand putative female resilience and/or how ASD phenotypes may present differently in females.

To examine the latent processes that support mice with different genotypes or sexes as they solve the odor-based 2AFC task, we fit trial-by-trial data using multiple reinforcement learning (RL) models, including a novel hybrid model that we recently developed that accounts for learning and exploration noise, as well as periods of disengagement at the policy level (Li et al., 2024).

Results

Adolescent male *Tsc2* Het mice show gain of function in early learning

We first tested adolescent male mice with a loss-of-function mutation in *Tsc2* in the odor-based 2AFC. To begin, *Tsc2* Het (n = 15) and WT (n = 13) mice were water restricted and habituated to the chamber, the central initiation port, and its two peripheral water delivery ports. After 4 days of habituation (Figure 1A), the mice were exposed to two odors (odors “A” and “B”) that were stably paired with a water reward on

either the left or right peripheral port (Figure 1B). In the first odor learning session (Session 1), male *Tsc2* Hets steadily increased performance throughout the session to reach over 70% correct, on average, in the last quarter of the session, whereas their WT littermates maintained performance near to chance (Figure 1C; 2-way RM ANOVA: geno: $F(1, 26)=5.10$, $p = 0.03$; time: $F(1.62, 42.28) = 17.49$, $p < 0.0001$; geno x time: $F(3, 78) = 2.79$, $p = 0.04$; post-hoc Tukey in quartile 4 showed Het significantly higher than WT ($p<0.01$)). In the second session of odor learning (the following day), Het animals again significantly outperformed their WT littermates by choosing the correct side more consistently following each odor presentation (Figure 1D; 2-way RM ANOVA: geno: $F(1,26)=8.56$, $p = 0.007$; time: $F(2.165, 56.29)=10.05$, $p = 0.0001$; geno x time: $F(3, 78) = 1.17$, $p =0.16$; with a significant post-hoc difference in quartiles 2, 3, and 4).

In addition to measuring performance, we quantified the trial reinitiation times over the intertrial interval (ITI) – the time from the reward port to the center port – as a reaction-time-like metric that may reflect differences in motivation, motor coordination, or both. In session 1, the ITI metric was comparable between WT and Het, but in session 2 Het mice were faster than WT (Supplemental Figure 1F: Unpaired t-test: Session 1: $t = 1.29$, df = 26, $p = 0.13$; Session 2: $t = 2.92$, df = 25, $p = 0.007$). We observed a significant negative correlation between performance and ITI in quartiles in both *Tsc2* Het and WT groups in the second session (Spearman's correlation: WT: $r = -0.52$, $p < 0.0001$; HET: $r = -0.42$, $p = 0.0004$). Together, these data reveal a gain of function in *Tsc2* males in the earliest sessions of odor-action reinforcement learning.

Adolescent female *Tsc2* Het mice have no gain of function in early learning

To test whether we would see a similar or different phenotype in *Tsc2* females, we trained female *Tsc2* Het (n = 10) and WT (n = 10) animals in the same odor 2AFC task. *Tsc2* Het and WT females learned the odor-based 2AFC task at a comparable rate (Figure 1E: Session 1: 2-way RM ANOVA: geno: $F(1,18)=1.69$, $p = 0.20$; time: $F(1.90, 34.31)=7.5$, $p = 0.002$; geno x time: $F(3,54)=0.30$, $p =0.82$) and both *Tsc2* Het and *Tsc2* WT groups reached 70% average performance by the end of the second session (Figure 1F, Session 2: 2-way RM ANOVA: geno: $F(1,18)=0.31$, p=58; time: $F(2.18,39.33)$, $p=0.0007$; geno x time: $F(3,54) = 0.89$, $p = 0.89$).

Trial reinitiation times quantified by ITI were also comparable between female *Tsc2* Het and WT animals across both sessions (Supplemental Figure 1G: Session 1, Mann-Whitney test: $U = 41$, $p = 0.52$; Session 2, Unpaired *t*-test: $t = 0.22$, df = 16, $p = 0.82$). Performance in session 2 was also negatively correlated with ITI in both genotypes (Pearson's correlation: WT: $r = -0.91$, $p < 0.0001$; HET: $r= -0.47$, $p = 0.002$).

To put these sex differences in context we also compared males and females in a supplemental analysis. *Tsc2* WT males and WT females had comparable performance in both sessions (Supplemental Figure 1B-C, 2-way RM ANOVA: Session 1: sex: $F(1,21)=3.21$, $p =0.08$; time: $F(1.70, 35.75)=8.18$, $p =0.001$; sex x time: $F(3,63)=1.337$, $p =0.27$; Session 2: sex: $F(1,21)=0.66$, $p =0.42$; time: $F(2.35, 49.46)=6.12$, $p =0.002$; sex x time: $F(3,63)=0.27$, $p=0.84$), while female *Tsc2* Het and male *Tsc2* Het showed distinctly different learning across both sessions with Het males showing a significant gain of function over Het females (Supplemental Figure 1D-E, 2-way RM ANOVA: Session 1: sex: $F(1,22)=4.97$, $p =0.03$; time: $F(1.84,40.6) = 6.12$, $p < 0.0001$; sex x time: $F(3,66)=2.74$, $p =0.04$; Session 2: sex: $F(1,23)=6.17$, $p =0.02$; time: $F(2, 46.11)=13.77$, $p < 0.0001$; sex x time: $F(3,69)=0.11$, $p = 0.95$). As a control, we also tested whether daily weight of male and female *Tsc2* Het and WT mice was correlated with performance. We found no evidence of a significant relationship between these variables in either sex (Supplemental Figure 1H-K).

Gain of function in adolescent *Tsc2* Het males is captured by higher learning rates

We next turned to computational modeling to see whether an exploration of latent variables in trial-by-trial learning could help inform which cognitive mechanisms were responsible for differences in performance between male *Tsc2* Het and WT mice. This also allowed us to explore potential latent differences in females. We compared five models that extended a basic reinforcement learning (RL) framework that posits, based on research from our lab (Chase et al., 2024), that mice rely on RL in an odor-based associative learning task. RL assumes that animals use reward outcomes to estimate the value of making a choice and iteratively update the values by the discrepancy between an obtained and an expected reward. To test if performance differences emerged from differences in attention or task disengagement, we also investigated a hybrid model we recently developed (Li et al., 2024). In this model, the agent alternates between two latent states that govern distinct policies: an RL-based policy when in an "engaged" state and a non-RL, side-biased policy when in a "disengaged" latent state (probabilistically choosing a single side while disregarding trial stimuli and outcomes). A hidden Markov model regulates transitions between these states (Figure 2A).

Our model comparison results, based on Bayesian Information Criterion (BIC), showed that hybrid models combining RL and biased policies significantly outperformed the RL component alone (Figure 2B, Wilcoxon signed-rank test: z= - 3.31, p < 0.001 between a0b3s and a0b3s_hybrid; see Materials & Methods for detailed model specifications). Among the multiple hybrid models we evaluated, a0b1s_hybrid had the best fit to choice behavior (Wilcoxon signed-rank test: z=2.10, p=0.036 between a0b1s_hybrid and a0b3s_hybrid; z=2.98, p=0.0028 between a0b1s_hybrid and a0b2s_hybrid). The best fitting model (a0b1s_hybrid) included a single learning rate (α +), an inverse temperature parameter (□) that controls choice stochasticity, a single strategy parameter (ξ_1) that measures "stickiness" (the likelihood of a mouse to repeat their previous action), latent state transition probabilities (lapse and recover), a bias parameter, and a fixed parameter set at 0 for learning from unrewarded trials. Model parameters were found to be identifiable (Supplemental Figure 2). At the group level, the hybrid model captured decision accuracy across learning (Figure 2C), including the highly variable trajectories observed during early learning sessions in individual animals (Figure 2D-E).

Using the best fit model, we compared parameters and variables across groups. In Session 1, male *Tsc2* Het mice exhibited a significantly higher alpha learning rate, α+, which signifies reliance on previous reward outcomes to guide future choices, compared to WT animals (Figure 2G: Mann-Whitney test, *U* =54, p = 0.04). This significant effect was maintained in an examination of a post-hoc variable, log (α * □), that captures the incremental change in policy (Daw, 2011) (Figure 2I: Mann-Whitney test, *U* = 41, p = 0.008).

All other parameters were comparable between male *Tsc2* Het and WT mice. There was no difference between male *Tsc2* Het and WT mice in the normalized bias parameter, which measures the degree to which behavior is biased towards a specific choice (Figure 2F: Unpaired t-test, *t* = 1.94, df = 26, p = 0.06) or in the softmax □ parameter (Figure 2H: Mann-Whitney test, *U* = 66, p = 0.15) or ξ_1 (Supplemental Figure 2K, Mann-Whitney test: *U* = 80, p = 0.43). The probability to transition from a biased to an engaged state (recover) and back (lapse) was also comparable between both *Tsc2* Het and WT mice (Supplemental Figure 1L-M: Mann-Whitney test, recover: *U* = 96, p =0.96; lapse: *U* = 87, p = 0.65). Finally, using the fitted parameters including latent state transition probabilities, we calculated post-hoc each animal's probability of being in an engaged state, "p(engaged)", on each trial and found no difference in the mean p(engaged) value between male *Tsc2* Het and WT mice (Supplemental Figure 1N: Mann-Whitney test, *U* = 87, p = 0.65) or in the p(engaged) frequency distribution (Supplemental Figure 8A).

When female *Tsc2* Het and WT mouse data were fit with the same winning model, a0b1s_hybrid, we found no significant difference in log(α+) learning rate, the inverse temperature parameter □, or in the post-hoc variable, log(α+ * □) (Figure 2K-M: Mann-Whitney tests: α+: *U* = 43.5, p = 0.64; ⊡, *U* = 34, p = 0.24; α+ * □: *U* = 42, p = 0.57). There was also no significant difference between female *Tsc2* Het and WT mice in the normalized bias parameter, indicating that neither Het nor WT mice exhibited a stronger bias towards a particular choice (Figure 2J: Mann-Whitney test, *U* = 46.5, p = 0.81). Likewise, no differences were observed in other fitted parameters or post-hoc calculations, including stickiness (ξ_1), transition probabilities (recover and lapse), and mean p(engaged), which estimates the likelihood of staying in the engaged state (Supplemental Figure 1O-R: Mann-Whitney test, ξ_1: *U* = 44, p = 0.66; recover: *U* = 44, p = 0.68; lapse: *U* = 40, p = 0.48; p(engaged): Mann-Whitney test, *U* = 40, p = 0.48).

Adolescent male *Shank3B* Het mice show a gain of function in early learning

Shank3B mice show corticostriatal hyperconnectivity in development (Peixoto et al., 2016; 2019), a phenotype that overlaps with observations in TSC models (Benthall et al., 2021). Therefore, we were interested in testing if there was a convergent reinforcement learning phenotype between *Tsc2* and *Shank3B* Hets. Adolescent male *Shank3B* Het mice (n = 12) along with WT littermates (n = 13) were water-restricted and trained in the odor-based 2AFC task as described (Materials and Methods). In the first odor learning session (Session 1), male *Shank3B* Het mice showed stronger performance (fraction correct) than WT littermates (Figure 3A: 2 way RM ANOVA: geno: F(1,23) = 3.3, p = 0.08; time: F(1.96, 45.2) = 26.08, p < 0.0001; geno x time: F(3,69) = 5.3, p = 0.002). ITI movement times were comparable between genotypes (Supplemental Figure 3A: Mann-Whitney test, *U* = 65, p = 0.50).

In the second session, male *Shank3B* Het sustained their higher level of performance in the first quartile, but *Shank3B* WT improved and closed the gap (Figure 3B: 2-way RM ANOVA: geno: F(1, 24) = 3.206, p = 0.08; time: F(1.53, 36.94)=7.46, p = 0.003; geno x time: F(3,72)=1.59, p = 0.19; post-hoc Tukey p<0.05 first quartile). ITIs in male *Shank3B* Het and WT were again comparable (Supplemental Figure 3A: Mann-Whitney test, *U* = 56, p = 0.56). Performance showed a negative correlation with ITI that was statistically significant in *Shank3B* Het, but not WT (Spearman's correlation: Het: r = -0.78, p < 0.0001, WT: r = -0.12, p = 0.48).

Adolescent female *Shank3B* Het mice have no gain of function in early learning

To test how haploinsufficiency of *Shank3B* affects females, we trained a separate cohort of adolescent female *Shank3B* Het (n = 11) and WT (n = 11) animals in the same odor-based 2AFC task. Female *Shank3B* Het and WT mice showed comparable performance in Session 1 (Figure 3F: 2-way RM ANOVA: geno: F(1,20) = 0.03, p = 0.85; time: F(2.22, 44.44)=6.26, p = 0.003; geno x time: F(3,60)=1.96, p = 0.12) and Session 2 (Figure 3G: 2-way RM ANOVA: geno: F(1,20)=0.20, p = 0.65; time: F(1.77, 35.43)=9.45, p = 0.0008; geno x time: F(3,60)=0.29, p = 0.83).

Female *Shank3B* WT animal performance was comparable to male *Shank3B* WT performance across both days (Supplemental Figure 3B-C: Session 1: 2-way RM ANOVA: sex: F(1,22) = 0.19, p = 0.66; time: F(2.00, 44.07)=7.37, p = 0.001; sex x time: F(3, 66) = 0.61, p = 0.60; Session 2: 2-way RM ANOVA: sex: F(1,23)=2.8, p = 0.10; time: F(1.47, 33.98) = 11.88, p = 0.0004; F(3,69)=0.28, p = 0.83). Performance observed in male *Shank3B* Hets was higher compared to female *Shank3B* Hets with a significant main effect of sex and sex x time interaction in session 1 (Supplemental Figure 3D-E: Session 1: 2-way RM ANOVA: sex: F(1,21)=4.75, p = 0.04; time: F(2.12, 44.53)=21.95, p < 0.0001; sex x time: F(3,63)=5.55, p = 0.001) and sex in session 2(Session 2: 2-way RM ANOVA: sex: F(1,21) = 7.26, p = 0.01; time: F(1.83, 38.55)=5.0, p = 0.01; sex x time: F(3,63)=0.48, p =0.69).

Female *Shank3B* Het and WT animals also showed comparable ITIs in both the first and second odor learning session (Supplemental Figure 3F: Session 1, Mann-Whitney: $U = 23$, $p = 0.07$; Session 2, Unpaired *t*-test, $t = 1.87$, df = 15, $p = 0.08$). We observed a significant negative relationship between fraction correct (performance) and ITI in both Het and WT *Shank3B* females (Spearman's correlation: $r = -0.66$, $p < 0.0001$; WT: Pearson's correlation: $r = -0.48$, $p = 0.002$).

Male *Shank3B* Het mice had significantly faster ITIs on average compared to female Het mice (Session 1: $t = 4.54$, $df = 19$, $p = 0.0002$; Session 2: $U = 15$, $p = 0.01$). In contrast, ITI speeds were comparable between *Shank3B* male and female WTs (Session 1: $U = 40$, $p = 0.13$; Session 2: $U = 40$, $p = 0.13$). Last, we performed an additional analysis to test whether daily weight of male and female *Shank3B* Het and WT mice was correlated with performance. We found no significant correlation between these variables in any *Shank3B* genotype (Supplemental Figure 3G-J).

Gain of function is captured by higher learning rates in adolescent male, not female, *Shank3B* Het mice

To explore latent structures in trial-by-trial learning, we next fit the *Shank3B* data using the set of models described above and detailed in our methods (Materials and Methods). Just as for the *Tsc2* mouse data, the winning model for *Shank3B* data was a0b1s_hybrid (Supplemental Figure 4A), which allowed for dynamic policy shifting between an RL-based "engaged" strategy and a biased, or "disengaged," strategy (see Figure 2A). The model correctly captured decision accuracy both at the group level (Supplemental Figure 4B) and at the individual level (Supplemental Figure 4C) for the *Shank3B* data.

Like *Tsc2* Het males, *Shank3B* Het males also showed significantly higher $\alpha+$ learning rates compared to WT littermates (Figure 3L: Mann-Whitney test $U = 47$, $p = 0.03$); this difference was also observed when examined with the post-hoc $\alpha+ *$ □ variable (Figure 3N: Mann-Whitney test $U = 41$, $p = 0.01$). Similar to *Tsc2* Hets, *Shank3B* Hets did not differ from WT animals in the normalized bias parameter, indicating that they did not prefer any particular side (Figure 3K, Mann-Whitney test, $U = 75.5$, $p = 0.46$), nor did they differ from WT in the strategy parameter, lapse, recover, or our post-hoc estimation of mean time spent in an engaged state (Supplemental Figure 4G-J: ξ_1: $U = 74$, $p = 0.42$; lapse: $U = 67$, $p = 0.25$; recover: $U = 61$; $p = 0.15$; p(engaged): unpaired t-test, $t = 0.95$, $df = 24$, $p = 0.34$). *Shank3B* Hets and WT also did not differ in their distribution of p(engaged) frequency (Supplemental Figure 8B). However, different from the *Tsc2* male results, *Shank3B* Het males also had a significantly higher □ parameter than *Shank3B* WT males (Figure 3M: Mann-Whitney test, □: $U = 35$, $p = 0.005$), indicating that, in addition to learning rapidly from outcomes, they may also have a less noisy or exploratory policy than WT littermates.

Female *Shank3B* Het and WT mice were also fit with the same set of models. The winning model for *Shank3B* females was the same a0b1s_hybrid model (Supplemental Figure 4D) and was validated both in group learning curves (Supplemental Figure 4E) and individual learning trajectories (Supplemental Figure 4F). Female *Shank3B* Hets did not differ significantly from WT mice in any fit parameter or post-hoc calculation included in our comparisons (Figure 3O-R: normalized bias: Unpaired *t*-test, $t = 0.20$, df = 17, $p = 0.84$; $\alpha+$: Mann-Whitney test, $U = 37.5$, $p = 0.54$; □: Mann-Whitney test, $U = 36$, $p = 0.49$; $\alpha+*$□: Mann-Whitney test, $U = 31$, $p = 0.27$; Supplemental Figure 4K-N: ξ_1: Mann-Whitney, $U = 32.5$, $p = 0.32$; recover: Unpaired *t*-test, $t = 0.06$, df = 17, $p = 0.95$; lapse: Mann-Whitney, $U = 43$, $p = 90$; p(engaged): $t = 0.66$, $df = 17$, $p = 0.51$).

Gain of function in learning in *Tsc2* and *Shank3B* Het mice is reward schedule dependent

We next investigated whether altering the reward schedule could influence the enhanced performance and learning rates observed in *Tsc2* and *Shank3B* Het males. Past work has suggested that neurocognitive processes captured by RL models, such as α learning rates or □ values, are not simply intrinsic to the individual and instead, can vary depending on the context in which the learning takes place (Eckstein, Wilbrecht & Collins, 2021; Eckstein et al., 2022). This especially applies to learning across development (Wilbrecht & Davidow, 2024) and there is evidence that autistic individuals learn differently from probabilistic feedback (Koegel et al., 1979; Solomon et al., 2011; South et al., 2014; Lawson, Mathys, & Rees, 2017).

In new cohorts of *Tsc2* and *Shank3B* Het and WT mice (*Tsc2* Het, n = 10; *Tsc2* WT, n = 10; *Shank3B* Het, n = 11; *Shank3B* WT, n = 11) we trained mice in the same odor-based 2AFC task with one key difference: correct choices were unrewarded 10-20% of the time, introducing probabilistic outcomes (see Materials & Methods).

All mice were able to learn stimulus-action contingencies, despite the introduction of uncertainty into the environment. But, unlike the genotype differences observed in our first experiment, under a probabilistic reward schedule, task performance of male *Tsc2* Hets and male *Shank3B* Hets was comparable to their WT littermates in both Session 1 (Figure 4A: *Tsc2,* Session 1: 2-way RM ANOVA: geno: $F(1,18) = 0.47$, $p = 0.50$; time: $F(1.61, 29.1)=15.3$, $p < 0.0001$; geno x time: $F(3,54) = 0.13$, $p = 0.93$; Figure 4D: *Shank3B*, Session 1: geno: $F(1,20)=0.25$, $p = 0.61$; time: $F(1.87,37.42)=13.16$, $p < 0.0001$; geno x time: $F(3,60)=2.45$, $p = 0.07$), and Session 2 (Figure 4B: *Tsc2,* Session 2: 2-way RM ANOVA: geno: $F(1,18) = 0.16$, $p = 0.68$; time: $F(2.01, 36.22) = 5.4$, $p = 0.008$; geno x time: $F(3, 54) = 0.25$, $p = 0.85$; Figure 4E: *Shank3B*, Session 2: geno: $F(1,20)=0.08$, $p = 0.77$; time: $F(2.15,43.15)=13.63$, $p <0.0001$; geno x time: $F(3,60)=2.29$, $p = 0.08$).
ITIs in Session 1 and 2 were also comparable between Het and WT from both lines in the probabilistic schedule (*Tsc2* Figure 4C: Session 1: Unpaired *t*-test, $t = 1.03$, df = 18, $p = 0.31$; Session 2: Unpaired *t*-test, $t = 0.08$, df = 14, $p = 0.93$) (*Shank3B* Figure 4F: Session 1: Mann-Whitney test, $U = 42$, $p = 0.38$; Session 2: Mann-Whitney test, $U = 19$, $p = 0.08$).

Adolescent male *Tsc2* and *Shank3B* Het mice have comparable learning rates to WT males when trained with a probabilistic reward schedule

We next fit our family of RL models to the data from males trained in the probabilistic context. Despite a lack of difference between genotypes in basic performance, it is still possible that there could be differences in latent parameters. We found that the winning model (a0b1s_hybrid) for *Tsc2* and *Shank3B* males trained in the probabilistic context was the same winning model as for the deterministic contexts (Supplemental Figure 5A,D) and was validated at both the group level (Supplemental Figure 5B,E) and for individual trajectories (Supplemental Figure 5C,F).

When trained in the probabilistic context, *Tsc2* Het and WT male mice had statistically comparable parameter values (Figure 4G-J: normalized bias: Mann-Whitney, $U = 38$, $p = 0.39$; α+: Mann-Whitney test, $U = 41$, $p = 0.51$; □: Mann-Whitney test, $U = 45$, $p = 0.73$; α+*□: Mann-Whitney test, $U = 44$, $p = 0.68$; Supplemental Figure 5G-J: ξ_1: Unpaired t-test, $t = 1.9$, $df = 18$, $p = 0.07$; recover: Mann-Whitney test, $U = 32$, $p = 0.18$; lapse: Mann-Whitney test, $U = 31$, $p = 0.16$; mean p(engaged): Mann-Whitney test, $U = 31$, $p = 0.16$).

Similarly, *Shank3B* Het and WT male mice trained in the probabilistic context also showed statistically comparable parameter values (Figure 4K-N: normalized bias: Mann-Whitney, $U = 43$, $p = 0.25$; α+: Mann-Whitney test, $U = 48$, $p = 0.42$; □: Mann-Whitney test, $U = 58$, $p = 0.89$; α+*□: Mann-Whitney test,

U = 55, p = 0.74; Supplemental Figure 5K-N: ξ_1: Unpaired t-test, *t* = 1.03, *df* = 20, p = 0.31; recover: Unpaired t-test with welch's correction, *t* = 1.67, *df* = 14.08, p = 0.11; lapse: Mann-Whitney test, *U* = 45, p = 0.33; p(engaged): Mann-Whitney test, *U* = 54, p = 0.69). These data suggest that the gain of function in learning rate observed in male *Tsc2* and *Shank3B* Het mice compared to WT was sensitive to uncertainty in reward schedule.

A subset of adolescent male *Shank3B* KO exhibit high trial numbers and gain of function in learning

While *Shank3B* Het mice more closely reflect the mutations reported in humans (Leblond et al., 2014; Cammarata-Scalisi et al., 2022), many animal studies focus on *Shank3B* KO mice (Mei et al., 2016; Jaramillo et al., 2016; Wang et al., 2017; Dhamne et al., 2017; Rendall et al., 2019). Thus, we also chose to train *Shank3B* KO male mice in the deterministic odor-based 2AFC task (100% reward).

We found that, on average, male *Shank3B* KO (n = 13) had a lower fraction correct performance than *Shank3B* WT (littermates whose data was shown above) in the first session (Figure 5A: 2-way RM ANOVA: genotype: F(1,23)=9.5, p = 0.005; time: F(2.01,46.44)=35.85, p < 0.0001; F(3,69)=2.20, p = 0.09) and similar performance to *Shank3B* WT across the second session (Figure 5B: 2-way RM ANOVA: F(1,25)=0.91, p=0.34; time: F(1.63, 40.94)=7.2, p = 0.003; geno x time: F(3,75)=1.13, p = 0.33).

In our experience, most animals, WT or otherwise, complete well over 600 trials in each behavioral session. However, in examining total trials completed by individual *Shank3B* KO mice across each session, we observed a bimodal distribution in the number of trials completed by *Shank3B* KO mice (see Supplemental Figure 6A-B). The first sub-group of KO animals (n = 5) (sg 1) showed typical numbers of trial completed in a session (compared to other *Shank3B* genotypes), with 840.5 ± 125.50 (SD) total trials per session (*Shank3B* Het averaged 890.21 ± 100.36 (SD) and *Shank3B* WT averaged 841.29 ± 110.45 (SD)). In contrast, the second sub-group of KO mice (n = 8) (sg 2) completed only 310.46 ± 160.55 (SD) trials, on average, per AB session (Figure 5C-D). *Shank3B* KO sub-groups came from multiple breeder pairs and individuals from the same litter could be found in the two sub-groups.

When we compared the task performance of these two *Shank3B* KO sub-groups to WT, we found that sg 1 significantly out performed WT while sg 2 showed a deficit compared to WT (Figure 5E: Session 1, 2-way RM ANOVA: subgroup: F(2,23)=6.12, p=0.007; time: F(2.07,47.68)=26.93, p<0.0001; subgroup x time: F(6,69)=4.57, p=0.0006; asterisks indicate Dunnett post hoc tests comparing sub-group vs WT for that quartile) (Figure 5F: Session 2: subgroup: F(2,23)=14.14, p < 0.0001; time: F(1.62, 37.32)=5.00, p =0.01; subgroup x time: F(6,69)=0.80, p = 0.56; asterisks indicate Dunnett post hoc tests comparing sub-group vs WT for that quartile). Further analysis of the KO revealed sub-group 1 KO mice weighed more than the sub-group 2 KO mice (Figure 5G: Mann-Whitney test: *U*=4, p = 0.001) and had faster ITIs in both sessions (Figure 5H, Session 1: Unpaired *t*-test, *t* = 4.59, df = 11, p = 0.0008; Session 2: Unpaired *t*-test, *t* = 7.37, df = 10, p < 0.0001). ITIs from sub-group 1 KOs were comparable to both *Shank3B* Het and *Shank3B* WT littermates (One-way ANOVA: Session 1: group: F(2,27)=0.95, p = 0.39; Session 2: sub-group: F(2,25)=0.14, p = 0.86).

In sum, our analysis of male *Shank3B* KO sub-groups suggests that a subset of these mice show deficits in all of our metrics, while another subset that appears more typical in basic metrics shows a gain of function in reinforcement learning, similar to *Shank3B* Het and *Tsc2* Het mice.

Discussion

In the present study, we used an odor-based 2AFC task and reinforcement learning models to examine learning in mice with mutations in two autism risk genes, *TSC2* and *SHANK3*. We chose these genes for two reasons. First, single gene mutations in *TSC2* or *SHANK3* in humans cause syndromes that are associated with autism diagnoses and together, may account for approximately 1% of ASD cases (de la Torre-Ubieta et al., 2016). Second, mice with mutations in *Tsc1*, *Tsc2*, or *Shank3B* have been shown to have an increase in corticostriatal connectivity (Peixoto et al., 2016; Benthall et al., 2018). This hyperconnectivity is potentially linked to enhanced motor learning in TSC lines (Benthall et al., 2021; Cording and Bateup, 2023) and grooming behaviors in mice that have parallels to repetitive, restricted behaviors in humans (Wang et al., 2017). Using an odor-based 2AFC, we found a convergent gain of function in the first days of odor learning for both adolescent male *Tsc2* Het and *Shank3B* Het mice that was not observed in females and was dependent on a deterministic (100%) reward schedule. When we analyzed trial-by-trial learning using an RL-based mixture model that accounts for periods of task disengagement, we found that the gain of function in early learning performance in both *Tsc2* and *Shank3B* males was driven by a significant difference in the α+ learning rate parameter – a difference that was not observed in females. This gain of function in learning rate in male Hets was no longer apparent when *Tsc2* and *Shank3B* males were trained with a probabilistic schedule, suggesting the learning phenotype was modulated by environmental uncertainty.

It may seem surprising that we see a gain of function in mice with mutations in autism risk genes. While it is true that previous studies of rodents with mutations in autism risk genes, including both *Tsc2* and *Shank3*, have shown deficits in hippocampal-mediated, perceptual, and social learning (Ehninger et al., 2008; Copping et al., 2017; Dhamne et al., 2017), there is a large literature showing a gain of function in motor learning in mice with mutations to a variety of autism risk genes, including *Tsc2* Hets (Penagarikano et al., 2011; Rothwell et al., 2014; Hulbert et al., 2020; Benthall et al., 2021; rev'd in Cording & Bateup, 2023; Cording et al., 2024). Gains in motor learning have been found using the accelerating rotarod, a task that engages the dorsal and ventral striatum (Cording & Bateup, 2023). One study found that a conditional knockout of *Tsc1* in direct pathway neurons of the dorsolateral striatum reduced endocannabinoid dependent LTD onto these neurons, enhanced their cortical inputs, and was sufficient to drive a gain of function in the rotarod (Benthall et al., 2021).

In *Shank3B* mice, the rotarod data are more complex as they vary by genotype (Het vs KO) and may also depend on the age of testing. Briefly, in adult *Shank3B* Het mice, rotarod performance has been found to be comparable to WT in two studies (Rendall et al., 2019; Cerilli et al., 2024). In *Shank3B* KO mice one study reports a day 1 gain of function then WT and Het "catch up" (Rendall et al., 2019) while another reports an overall loss of function compared to HET and WT (Cerilli et al., 2024). Notably, these rotarod tests were performed in adulthood when corticostriatal circuits may weaken in *Shank3B* mice (Peixoto et al., 2019). Future work should examine rotarod performance across different ages in *Shank3B* mice to examine if it is developmentally regulated.

In humans, ASD related restricted and repetitive behaviors have been correlated with changes in connectivity and structure of the dorsal striatum (Cerliani et al., 2015; Schuetze et al., 2016). It has been proposed that an increase in innate repetitive grooming behaviors in adult *Shank3B* KO mice may model these symptoms. Interestingly, this enhanced grooming in *Shank3B* KO has also been linked to disrupted endocannabinoid-dependent regulation of cortical inputs onto spiny projection neurons in the dorsal striatum (Wang et al., 2017) with some clear parallels to studies of *Tsc1* (Benthall et al., 2021). In a vein more similar to our 2AFC data, a gain of function in perceptual discrimination learning has also been reported in *Shank3B* Hets and KO that is limited to the initial days of training (Rendall et al., 2019).

While rotarod learning, repetitive grooming, perceptual discrimination, and 2AFC learning look very different in practice, they all rely on the striatum and corticostriatal circuits (Balleine, Delgado, & Hikosaka, 2007; Lee et al., 2015; Znamenskiy et al., 2013; Cox & Witten, 2019; Tang et al., 2022). Building from this literature, we support and extend a working model in which enhanced corticostriatal connectivity (generated by diverse mechanisms), can facilitate a gain of function that can manifest in rotarod learning (Benthall et al., 2021; Cording & Bateup, 2023), repetitive behaviors like excess grooming (Wang et al., 2017), and 2AFC learning with an enhanced α learning rate (data presented here). The fact that two very different ASD risk genes produced a similar gain of function in early 2AFC learning in males, supports the models which suggest ASD may emerge from a convergent change in circuit function, possibly at the level of corticostriatal synapses, downstream of a variety of genetic changes (Fucillo, 2016; Cording and Bateup, 2023; Cording et al., 2024).

The fact that significant behavioral phenotypes and RL parameter differences were observed in males, but not females in both lines, also suggest convergent sex differences. This aligns with the human literature, where a male bias in autism diagnosis is commonly found (Maenner et al., 2023). Post-diagnosis, in humans, autistic phenotypes are also known to differ significantly between sexes, with males often exhibiting more pronounced restricted and repetitive behaviors than females (Hartley & Sikora, 2009; Mandy et al., 2012; Schuck, Flores, & Fung, 2019). In rodent models of ASD, similar sex differences have been reported. For instance, some pharmacological models show that female rodents display less severe behavioral differences from wild-type animals than males, particularly in sociability and repetitive behaviors like grooming or errors in reversal learning tasks (Murta et al., 2023). In multiple genetic models of autism, male mice have shown more putative autism related phenotypes compared to female mice (Ferri et al., 2021), but some risk genes in mice have been associated with ASD related phenotypes with no evidence of sex differences (Murta et al., 2023).

Our findings strongly support a role for sex in the expression of learning related phenotypes, but they do not help us isolate when genes and/or hormonal influences were critical to the generation of these differences. We collected data at P30, which is just before the time of first estrus in mice in female C57 Bl/6 mice reared in our lab (Piekarski et al., 2017). We did not monitor pubertal milestones or gonadal hormone levels in our mice, but it is possible that adult phenotypes differ again and are sensitive to gonadal hormones at puberty (Delevich et al., 2021). Further research is also needed to determine how development may also be of importance in learning in people with and without an ASD diagnosis (Dahl et al 2018; Nussenbaum & Hartley, 2019; Corbett et al., 2020; Muscatello et al., 2022). There are of course numerous differences between human and mouse brains, but reinforcement learning in a 2AFC task may enable us to isolate latent variables comparable across species (Chase et al., 2024; Master et al., 2020).

Recent research indicates that learning rates are not necessarily intrinsic to an individual but instead are shaped by interactions between the individual and the context of a learning environment (Nussenbaum & Hartley, 2019; Eckstein, Wilbrecht & Collins, 2021; Eckstein et al., 2022; Heald, Wolpert, & Lengyel, 2023). Our probabilistic schedule experiments were intended to probe the gain of function in *Tsc2* and *Shank3B* Het males by using a new cohort and training them with a different reinforcement schedule in the 2AFC task where correct choices were not rewarded 10-20% of the time. With this contextual change that increased uncertainty, we no longer observed a gain of function in performance or α learning rate in male *Tsc2* and *Shank3B* Hets compared to their WT littermates. From a bird's eye perspective, reducing the reinforcement schedule could be seen as a phenotypic "rescue" that could be useful for training and interventions where enhanced learning may be disadvantageous. These data may be consistent with data from human participants that found changes in reinforcement schedule led to appearance and/or disappearance of performance differences between autistic and neurotypical participants (Solomon et al., 2011; Lawson, Mathys, & Rees, 2017; Koegel, 1972).

In conclusion, our data suggest diverse genetic changes associated with autism can interact with sex to lead to convergent behavioral outcomes. While these behavioral phenotypes stand out in males, they are still sensitive to environmental context and experience. Future studies may find there are alternate manifestations in females and that age and/or stage of testing also plays a large role. Human learning profiles may also differ from mice. One takeaway is that learning proceeds differently with ASD-related biology and may be sculpted by different rules. These observations support the celebration of some strengths in autism and may inform the design of future interventions and outcome metrics.

Materials and Methods

Animals

Two lines of mice were used for this experiment: *Tsc2* (JAX, #004686) and *Shank3B* (Jackson Laboratories Repository (JAX), #017688). Breeding colonies were maintained at the University of California, Berkeley for both *Shank3B* and *Tsc2* animals by breeding heterozygous males and females for *Shank3B* and a heterozygous male/female with a WT male/female for *Tsc2* (homozygous deletion of *Tsc2* is embryonically lethal). All mice used in experiments were bred in the animal facility under a reverse light/dark cycle (12h:12h) and were group housed. Adolescent (postnatal (P) 27-32) male (n = 108) and female (n = 42) *Tsc2*$^{+/-}$, *Shank3B*$^{+/-}$ *, Shank3B*$^{-/-}$ *,* and WT mice from both strains were used for these experiments. All procedures were approved by University of California, Berkeley Institutional Animal Care and Use Committee (IACUC).

Odor-based Reinforcement Learning Task

Custom-built operant chambers that have an initiation port in the center and two lateral ports (Figure 1A) were used for 2AFC behavioral testing. Each port contains an infrared photodiode/phototransistor pair to accurately time port entry and exit (Island Motion, Tappan, NY). Side ports release water through valves (Neptune Research) and were regularly calibrated to ensure that exactly 2 μL of water was delivered for correct choices. Constant air flows through the center port at ~0.5L/min and is redirected through one of the two possible odorant filters (Target2 GMF F2500-19) at trial initiation. Our behavioral rig allows for the delivery of multiple odors independently, allowing for seamless transition between odors without experimenter intervention. For these experiments, the task was controlled using custom-built MATLAB scripts that also store outcome and action information.

Inexperienced adolescent (postnatal day 27) male and female mice were first water-restricted and pre-trained for four days to initiate trials via nose poke to the center port and next to associate side ports with water reward. Next, animals entered Session 1 where they first encountered odor pair A (cinnamon: McCormick & Co, Hunt Valley, MD) and B (vanilla: McCormick Culinary, Hunt Valley, MD), which stably predicted the left and right port respectively for reward. Each session lasted 6-12 hours. Animals were all weighed at the end of each session and, if they did not gain weight within the session and/or across sessions, they were given supplemental water. All animals gained weight throughout their training (Supplemental Figure 1F) at a similar developmental trajectory to WT littermates (Supplemental Figure 7) and fraction correct performance was not correlated with weight in males or females (Supplemental Figure 1B-C, G-H; Supplemental Figure 3G-J). The training protocol was identical for the probabilistic reward schedule cohorts, except that 10-20% of correct choice trials were unrewarded. All mice experienced 10% unrewarded trials for Session 1 (90-0 protocol) and most mice also experienced 10% unrewarded trials for Session 2. A small subset of mice (3 *Tsc2* WT, 2 *Tsc2* Het, and 1 *Shank3B* WT)

experienced both an 90-0 and 80-0 protocol on Session 2. No animals in the probabilistic cohort experienced deterministic rewards in the odor task prior to training.

Data and Statistical Analysis

Behavioral data was collected in MATLAB. Individual sessions were of variable length so each session for each mouse was divided into quartiles and the mean fraction correct and intertrial interval (ITI) was calculated as time between side port exit and center port entry for each quartile. To better reflect task-related behavior but to remove grooming sessions or other off-task behavior, ITIs that were longer than 10 seconds (typically less than 5% of trials for each animal) were removed before this ITI calculation. Note, all trials were included in fraction correct and trial by trial modeling. Fraction correct and ITI data for each group are shown as arithmetic mean and standard error of the mean (SEM) for each quartile in each session (Jupyter notebook, Python 3; GraphPad Prism 10). We checked assumptions of normality and variance using Shapiro-Wilk and Anderson-Darling tests (Graphpad Prism 10) and then used repeated measures 2-way analysis of variance (RM 2-way ANOVA) to examine performance over time in a session, Unpaired *t*-tests or Mann-Whitney tests for single time point comparisons, and Pearson and Spearman for tests of correlation (Graphpad Prism 10 or Jupyter Notebook Python 3). The significance level was set to $p < 0.05$ for all statistical tests.

Computational Modeling

Model specifications

We evaluated five RL-based models with variants of learning and decision-making mechanisms. All models learn by tracking and updating the values of both actions trial-by-trial. On any trial $t \in [1, T]$, we denote the state (odor identity) as s_t, action (go left or right) as $a_t \in \{L, R\}$, magnitude of reward as $r_t \in \{0, 1\}$, and value of an action a in state s as $Q(s, a)$. On trial t, the model samples an action according to π_t, a softmax policy that returns the probability of sampling the "go left" action given the current state s_t:

$$\pi_t(L|s_t) = \frac{1}{1 + exp(\beta^* (Q_t(s_t, R) - Q_t(s_t, L)) + stickiness)} \cdot (1 - \epsilon) + \epsilon \cdot \frac{1}{2},$$

where β is the softmax inverse temperature parameter and ϵ parameterizes uniform decision noise. Stickiness $= \xi \cdot \mathbb{1}_{a_{t-1}=L}$ is a choice kernel defined by a set of parameters ξ (varying across models) that models the extent to which the animal sticks to the previously chosen action conditioned on the outcome of the previous action (i.e., r_{t-1}) and whether the current odor is the same as the previous odor (i.e., $s_t = s_{t-1}$); see below for further details on the "stickiness" kernels. When the chosen action a_t is rewarded, the action value is updated incrementally with a reward prediction error, following the standard delta rule:

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \alpha (r_t - Q_t(s_t, a_t)),$$

where α is the learning rate parameter, whose value depends on whether the chosen action a_t is rewarded. If it is rewarded, the learning rate takes on the value α_+, whereas if the chosen action is unrewarded, the learning rate is α_-.

Our baseline models include a standard RL model and the winning model in Chase et al. (2024), named RL_epsilon and a0b3s, respectively. RL_epsilon has five parameters: positive and negative learning rates (α_+ and α_-), softmax inverse temperature (β), uniform decision noise (ϵ), and one-back choice stickiness independent of states and outcomes (ξ). The a0b3s model also includes five parameters: α_+, β, and three choice stickiness parameters conditioned on the outcome of the previous action and whether the state has changed from the previous trial to the current trial (ξ_1 if the previous outcome is positive, ξ_2 if the previous outcome is negative and the current state differs from the previous state, and ξ_3 if the previous outcome is negative and the current state is the same as the previous state). Following Chase et al. (2024), when the chosen action is not rewarded, no action values are updated (i.e., the learning rate for negative rewards α_- is zero).

Extending the a0b3s model using dynamic noise estimation (Li et al., 2024), we tested three additional models with hybrid policies between a latent RL policy state and a latent biased policy state: a0b3s_hybrid, a0b2s_hybrid, and a0b1s_hybrid. We named our models based on the parameters in the RL policy: the "a0-" prefix indicates that the positive learning rate α_+ is a free parameter while the negative learning rate α_- is fixed to 0, "-b-" denotes the softmax inverse temperature parameter β, and the "-3s", "-2s", and "-1s" suffixes specify the number of stickiness parameters. Specifically, a0b3s_hybrid includes the a0b3s model as a policy state, a0b2s_hybrid only includes ξ_1 and ξ_2 (fixed $\xi_3 = 0$), and a0b1s_hybrid only includes ξ_1 (fixed ξ_2=0 or $\xi_3 = 0$).

The biased policy is a heuristic (i.e., not learned in the task) defined by a "bias" parameter between 0 and 1, which quantifies the degree of bias for the left action:

$$\pi_t(L|s_t) = bias .$$

The hybrid model uses a hidden Markov model to construct a mixture policy between the RL and biased policies. It assumes the existence of two discrete hidden latent states (engaged or biased), which determine whether the agent follows the engaged RL-based policy or the biased heuristic policy. The transition probability from an engaged to a biased (lapse) and from a biased to an engaged state (recover) is described in Figure 2A, is equal to the transition probability from the current state into the same state, and can be summarized with the below relationship:

$$T_{Biased}^{RL} + T_{RL}^{Biased} = 1.$$

Its likelihood is estimated by extending the dynamic noise estimation framework described in (Li et al., 2024).

Model fitting

All models were fitted with maximum a posteriori (MAP) estimation using a normal prior on the softmax inverse temperature parameter β with a mean of 5 and a standard deviation of 7, and uninformative priors

for all other parameters. We used a prior for the β parameters as it is known to otherwise trade off with the learning rate parameter at fitting (Daw, 2011). All models were fit at the individual level.

Model comparison

The fitted models were compared using the Bayesian Information Criterion (BIC) (Schwarz, 1978). BIC Figures can be found in Figure 2 for *Tsc2* mice as well as Supplementary Figure 4 and 5 for *Shank3B* mice and probabilistic training schedules, respectively. The best-fit parameters and latent variables were compared at the group-level for interpretation. We report the log-transformed values of the learning rates due to small values. The bias parameter is normalized by $|bias - 0.5|$ to indicate the un-sided magnitude of bias in choices. We additionally report $log(\alpha_+ * \beta)$, a composite index between the positive learning rate and softmax inverse temperature, as a measure of the incremental change in policy that accounts for both the rate of learning updates and the stochasticity in decision-making (Daw, 2011).

Model validation

To validate the model, we simulated choice behavior for each animal repeatedly for 1000 times using the MAP parameters obtained from model-fitting. For simulations with hybrid models, we used the latent state probability—P(RL) and P(Biased)—trajectories inferred from real data to estimate latent state occupancy. To validate how well the models captured behavior, we compared learning curves between these model simulations and the animal data that the models were fitted to. Group and individual validation curves can be found in Figure 2 for *Tsc2* mice as well as Supplementary Figure 4 and 5 for *Shank3B* mice and probabilistic training schedules, respectively.

Parameter recovery

We performed parameter recovery analysis as a robustness check of the winning model, a0b1s_hybrid. To that end, we simulated choice data for each animal using the best-fit parameter values as ground-truth given the model-estimated latent state occupancy trajectory based on animal data. We then fitted the model back to the simulated data to obtain best-fit parameters, which were compared to the ground-truth parameter values (Supplementary Figure 2).

Quantification of group-level p(engaged) distribution differences

To assess differences in the distributions of model-estimated engagement probability (p(engaged)) between groups (e.g., WT and Hets), we employed Kullback-Leibler (KL) divergence as a measure of difference between two probability distributions. KL divergence quantifies the extent to which one distribution diverges from another, with values close to zero indicating high similarity. Group-level p(engaged) distributions were constructed by dividing trial-level probabilities into 10 bins for each animal and averaging across animals within each group.

For statistical comparison, we performed a permutation test. Animal-level *p(engaged)* distributions were randomly reassigned to animals across groups, and KL divergence was recalculated for each permutation (10,000 iterations). The observed KL divergence between the original group distributions was compared to the distribution of permuted KL divergences to compute a p-value, defined as the proportion of permuted KL values greater than or equal to the observed value. This procedure provided a non-parametric assessment of whether observed group-level differences were significant. Error bars represent the standard error of the mean (SEM) for binned *p(engaged)* values.

Acknowledgements

We thank Tory Benson, Anna Jahng, Katrina Manaloto, and Katrina Wong for their assistance in animal behavior and handling. We thank Kristen Delevich for helping to initiate these experiments and Helen Bateup, Dan Feldman, Rui Peixoto, Jonathan Tarbox, Katie Cording, and Katie Benthall for discussion.

This work was supported by a Pilot grant from the Simons Foundation Autism Initiative (SFARI) Award #613972 to L.W. and the National Institutes of Health (R01MH134514) to L.W.

Author Contributions:

Juliana Chase: Investigation, Conceptualization, Data Curation, Methodology, Writing - Original Draft, Visualization, Supervision, Analysis, Validation; **Jing-Jing Li**: Methodology, Software, Analysis, Investigation, Data Curation, Visualization, Writing; **Wan Chen Lin**: Software, Analysis; **Lung-Hao Tai**: Methodology, Software; **Fernanda Castro**: Investigation; **Anne GE Collins**: Resources, Supervision, Conceptualization, Funding acquisition, Writing - Editing; **Linda Wilbrecht:** Project Administration, Funding acquisition, Conceptualization, Supervision, Writing - Editing

Competing Interest Statement: The authors have no competing interests.

References

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Kurzius-Spencer, M., Zahorodny, W., Robinson Rosenberg, C., White, T., Durkin, M. S., Imm, P., Nikolaou, L., Yeargin-Allsopp, M., Lee, L.-C., Harrington, R., Lopez, M., Fitzgerald, R. T., Hewitt, A., & Pettygrove, S. (2018). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *MMWR. Surveillance Summaries*, *67*(6), 1–23. https://doi.org/10.15585/mmwr.ss6706a1

Baker, P., Piven, J., & Sato, Y. (1998). *Journal of Autism and Developmental Disorders*, *28*(4), 279–285. https://doi.org/10.1023/a:1026004501631

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *The Journal of Neuroscience*, *27*(31), 8161–8165. https://doi.org/10.1523/jneurosci.1554-07.2007

Baron, M. K. (2006). An Architectural Framework That May Lie at the Core of the Postsynaptic Density. *Science*, *311*(5760), 531–535. https://doi.org/10.1126/science.1118995

Benthall, K. N., Cording, K. R., Agopyan-Miu, A. H. C. W., Wong, C. D., Chen, E. Y., & Bateup, H. S. (2021). Loss of Tsc1 from striatal direct pathway neurons impairs endocannabinoid-LTD and enhances motor routine learning. *Cell Reports*, *36*(6), 109511. https://doi.org/10.1016/j.celrep.2021.109511

Benthall, K. N., Ong, S. L., & Bateup, H. S. (2018). Corticostriatal transmission is selectively enhanced in striatonigral neurons with postnatal loss of TSC1. *Cell Reports*, *23*(11), 3197–3208. https://doi.org/10.1016/j.celrep.2018.05.037

Boronat, S., Elizabeth Anne Thiele, & Caruso, P. (2017). Cerebellar lesions are associated with *TSC2* mutations in tuberous sclerosis complex: a retrospective record review study. *Developmental Medicine and Child Neurology/Developmental Medicine & Child Neurology*, *59*(10), 1071–1076. https://doi.org/10.1111/dmcn.13499

Cammarata-Scalisi, F., Callea, M., Martinelli, D., Willoughby, C. E., Tadich, A. C., Araya Castillo, M., Lacruz-Rengel, M. A., Medina, M., Grimaldi, P., Bertini, E., & Nevado, J. (2022). Clinical and genetic aspects of Phelan–McDermid Syndrome: An interdisciplinary approach to management. *Genes*, *13*(3), 504. https://doi.org/10.3390/genes13030504

Capal, J. K., Williams, M. E., Pearson, D. A., Kissinger, R., Horn, P. S., Murray, D., Currans, K., Kent, B., Bebin, M., Northrup, H., Wu, J. Y., Sahin, M., & Krueger, D. A. (2021). Profile of Autism Spectrum Disorder in Tuberous Sclerosis Complex: Results from a Longitudinal, Prospective, Multisite Study. *Annals of Neurology*, *90*(6), 874–886. https://doi.org/10.1002/ana.26249

Cerilli, E., Dall'O, G. M., Chelini, G., Catena, B., Weinberger, B., Bozzi, Y., & Pangrazzi, L. (2024). Immune system dysfunction and inflammation in aging Shank3b mutant mice, a model of autism spectrum disorder. *Frontiers in Immunology*, *15*. https://doi.org/10.3389/fimmu.2024.1447385

Cerliani, L., Mennes, M., Thomas, R. M., Di Martino, A., Thioux, M., & Keysers, C. (2015). Increased functional connectivity between subcortical and cortical resting-state networks in autism spectrum disorder. *JAMA Psychiatry*, *72*(8), 767. https://doi.org/10.1001/jamapsychiatry.2015.0101

Chase, J., Xia, L., Tai, L.-H., Lin, W. C., Collins, A. G., & Wilbrecht, L. (2024). Adolescent and adult mice use both incremental reinforcement learning and short term memory when learning concurrent stimulus-action associations. *PLOS Computational Biology*, *20*(12). https://doi.org/10.1371/journal.pcbi.1012667

Comparan-Meza, M., Vargas de la Cruz, I., Jauregui-Huerta, F., Gonzalez-Castañeda, R. E., Gonzalez-Perez, O., & Galvez-Contreras, A. Y. (2021). Biopsychological correlates of repetitive and restricted behaviors in autism spectrum disorders. *Brain and Behavior*, *11*(10), e2341. https://doi.org/10.1002/brb3.2341

Copping, N. A., Berg, E. L., Foley, G. M., Schaffler, M. D., Onaga, B. L., Buscher, N., Silverman, J. L., & Yang, M. (2017). Touchscreen learning deficits and normal social approach behavior in the Shank3B model of Phelan–McDermid Syndrome and autism. *Neuroscience*, *345*, 155–165. https://doi.org/10.1016/j.neuroscience.2016.05.016

Corbett, B. A., Vandekar, S., Muscatello, R. A., & Tanguturi, Y. (2020). Pubertal timing during early adolescence: Advanced pubertal onset in females with autism spectrum disorder. *Autism Research*, *13*(12), 2202–2215. https://doi.org/10.1002/aur.2406

Cording, K. R., & Bateup, H. S. (2023). Altered motor learning and coordination in mouse models of autism spectrum disorder. *Frontiers in Cellular Neuroscience*, *17*. https://doi.org/10.3389/fncel.2023.1270489

Cording, K. R., Tu, E. M., Wang, H., Agopyan-Miu, A. H., & Bateup, H. S. (2024). Cntnap2 loss drives striatal neuron hyperexcitability and behavioral inflexibility. *bioRxiv 2024.05.09.593387*. https://doi.org/10.1101/2024.05.09.593387

Costales, J. L., & Kolevzon, A. (2015). Phelan–McDermid syndrome and Shank3: Implications for treatment. *Neurotherapeutics*, *12*(3), 620–630. https://doi.org/10.1007/s13311-015-0352-z

Cox, J., & Witten, I. B. (2019). Striatal circuits for reward learning and decision-making. *Nature Reviews Neuroscience*, *20*(8), 482–494. https://doi.org/10.1038/s41583-019-0189-2

Dahl, R. E., Allen, N. B., Wilbrecht, L., & Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature*, *554*(7693), 441–450. https://doi.org/10.1038/nature25770

Daw, N. D. (2011). Trial-by-trial data analysis using computational models . In *Decision Making, Affect, and Learning: Attention and Performance XXIII* (pp. 3–38). essay, Oxford University Press.

de la Torre-Ubieta, L., Won, H., Stein, J. L., & Geschwind, D. H. (2016). Advancing the understanding of autism disease mechanisms through genetics. *Nature Medicine*, *22*(4), 345–361. https://doi.org/10.1038/nm.4071

Delevich, K., Klinger, M., Okada, N. J., & Wilbrecht, L. (2021). Coming of age in the frontal cortex: The role of puberty in cortical maturation. *Seminars in Cell & Developmental Biology*, *118*, 64–72. https://doi.org/10.1016/j.semcdb.2021.04.021

Dhamne, S. C., Silverman, J. L., Super, C. E., Lammers, S. H. T., Hameed, M. Q., Modi, M. E., Copping, N. A., Pride, M. C., Smith, D. G., Rotenberg, A., Crawley, J. N., & Sahin, M. (2017). Replicable in vivo physiological and behavioral phenotypes of the Shank3B null mutant mouse model of autism. *Molecular Autism*, *8*(1). https://doi.org/10.1186/s13229-017-0142-z

Dölen, G., & Malenka, R. C. (2014). The Emerging Role of Nucleus Accumbens Oxytocin in Social Cognition. *Biological Psychiatry*, *76*(5), 354–355. https://doi.org/10.1016/j.biopsych.2014.06.009

Durand, C. M., Betancur, C., Boeckers, T. M., Bockmann, J., Chaste, P., Fauchereau, F., Nygren, G., Rastam, M., Gillberg, I. C., Anckarsäter, H., Sponheim, E., Goubran-Botros, H., Delorme, R., Chabane, N., Mouren-Simeoni, M.-C., de Mas, P., Bieth, E., Rogé, B., Héron, D., … Bourgeron, T. (2006). Mutations in the gene encoding the synaptic scaffolding protein Shank3 are associated with autism spectrum disorders. *Nature Genetics*, *39*(1), 25–27. https://doi.org/10.1038/ng1933

Eckstein, M. K., Wilbrecht, L., & Collins, A. G. (2021). What do reinforcement learning models measure? interpreting model parameters in cognition and Neuroscience. *Current Opinion in Behavioral Sciences*, *41*, 128–137. https://doi.org/10.1016/j.cobeha.2021.06.004

Eckstein, Maria Katharina, Master, S. L., Xia, L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. (2022). The interpretation of computational model parameters depends on the context. *eLife*, *11*. https://doi.org/10.7554/elife.75474

Ehninger, D., Han, S., Shilyansky, C., Zhou, Y., Li, W., Kwiatkowski, D. J., Ramesh, V., & Silva, A. J. (2008). Reversal of learning deficits in a Tsc2+/− mouse model of tuberous sclerosis. *Nature Medicine*, *14*(8), 843–848. https://doi.org/10.1038/nm1788

Ferri, S. L., Abel, T., & Brodkin, E. S. (2018). Sex Differences in Autism Spectrum Disorder: a Review. *Current Psychiatry Reports, 20*(2). https://doi.org/10.1007/s11920-018-0874-2

Ferri, S. L., Dow, H. C., Schoch, H., Lee, J. Y., Brodkin, E. S., & Abel, T. (2021). Age- and sex-specific fear conditioning deficits in mice lacking PCDH10, an autism associated gene. *Neurobiology of Learning and Memory*, *178*, 107364. https://doi.org/10.1016/j.nlm.2020.107364

Fisher, S. D., Ferguson, L. A., Bertran-Gonzalez, J., & Balleine, B. W. (2020). Amygdala-cortical control of striatal plasticity drives the acquisition of goal-directed action. *Current Biology*, *30*(22). https://doi.org/10.1016/j.cub.2020.08.090

Fuccillo, M. V. (2016). Striatal circuits as a common node for autism pathophysiology. *Frontiers in Neuroscience*, *10*. https://doi.org/10.3389/fnins.2016.00027

Gipson, T. T., Poretti, A., Kelley, S. A., Carson, K. A., Johnston, M. V., & Thierry A.G.M. Huisman. (2019). Characterization of the Basal Ganglia Using Diffusion Tensor Imaging in Children with Self‐Injurious Behavior and Tuberous Sclerosis Complex. *Journal of Neuroimaging*, *29*(4), 506–511. https://doi.org/10.1111/jon.12628

Grissom, N. M., McKee, S. E., Schoch, H., Bowman, N., Havekes, R., O'Brien, W. T., Mahrt, E., Siegel, S., Commons, K., Portfors, C., Nickl-Jockschat, T., Reyes, T. M., & Abel, T. (2018). Male-specific deficits in natural reward learning in a mouse model of neurodevelopmental disorders. *Molecular Psychiatry*, *23*(3), 544–555. https://doi.org/10.1038/mp.2017.184

Hartley, S. L., & Sikora, D. M. (2009). Sex differences in autism spectrum disorder: An examination of developmental functioning, autistic symptoms, and coexisting behavior problems in Toddlers. *Journal of Autism and Developmental Disorders*, *39*(12), 1715–1722. https://doi.org/10.1007/s10803-009-0810-8

Heald, J. B., Wolpert, D. M., & Lengyel, M. (2023). The computational and neural bases of context-dependent learning. *Annual Review of Neuroscience*, *46*(1), 233–258. https://doi.org/10.1146/annurev-neuro-092322-100402

Huang, J., & Manning, B. D. (2008). The TSC1–TSC2 complex: A molecular switchboard controlling cell growth. *Biochemical Journal*, *412*(2), 179–190. https://doi.org/10.1042/bj20080281

Hulbert, S. W., Wang, X., Gbadegesin, S. O., Xu, Q., Xu, X., & Jiang, Y. (2020). A novel Chd8 mutant mouse displays altered ultrasonic vocalizations and enhanced motor coordination. *Autism Research*, *13*(10), 1685–1697. https://doi.org/10.1002/aur.2353

Hull, L., Petrides, K. V., & Mandy, W. (2020). The female autism phenotype and camouflaging: A narrative review. *Review Journal of Autism and Developmental Disorders*, *7*(4), 306–317. https://doi.org/10.1007/s40489-020-00197-9

Hwang, F.-J., Roth, R. H., Wu, Y.-W., Sun, Y., Kwon, D. K., Liu, Y., & Ding, J. B. (2022). Motor learning selectively strengthens cortical and striatal synapses of motor engram neurons. *Neuron*. https://doi.org/10.1016/j.neuron.2022.06.006

Jaramillo, T. C., Speed, H. E., Xuan, Z., Reimers, J. M., Escamilla, C. O., Weaver, T. P., Liu, S., Filonova, I., & Powell, C. M. (2016). Novel Shank3 Mutant exhibits behaviors with face validity for autism and altered striatal and hippocampal function. *Autism Research*, *10*(1), 42–65. https://doi.org/10.1002/aur.1664

Jackson Laboratory. Body weights for 24 inbred strains of mice and F1 hybrids, by age. MPD:Jax1. Mouse Phenome Database web resource (RRID:SCR_003212), The Jackson Laboratory, Bar Harbor, Maine USA. https://phenome.jax.org

Kim, E., & Sheng, M. (2004). PDZ domain proteins of synapses. *Nature Reviews Neuroscience*, *5*(10), 771–781. https://doi.org/10.1038/nrn1517

Koegel, R. L., Schreibman, L., Britten, K., & Laitinen, R. (1979). The effects of schedule of reinforcement on stimulus overselectivity in autistic children. *Journal of Autism and Developmental Disorders*, *9*(4), 383–397. https://doi.org/10.1007/bf01531446

Laplante, M., & Sabatini, D. M. (2012). MTOR signaling in growth control and disease. *Cell*, *149*(2), 274–293. https://doi.org/10.1016/j.cell.2012.03.017

Larsen, B., & Luna, B. (2018). Adolescence as a neurobiological critical period for the development of higher-order cognition. *Neuroscience & Biobehavioral Reviews*, *94*(1), 179–195. https://doi.org/10.1016/j.neubiorev.2018.09.005

Lawson, R. P., Mathys, C., & Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. *Nature Neuroscience*, *20*(9), 1293–1299. https://doi.org/10.1038/nn.4615

Le Merrer, J., Detraux, B., Gandía, J., De Groote, A., Fonteneau, M., de Kerchove d'Exaerde, A., & Becker, J. A. J. (2024). Balance between projecting neuronal populations of the nucleus accumbens controls social behavior in mice. *Biological Psychiatry*, *95*(2), 123–135. https://doi.org/10.1016/j.biopsych.2023.05.008

Leblond, C. S., Nava, C., Polge, A., Gauthier, J., Huguet, G., Lumbroso, S., Giuliano, F., Stordeur, C., Depienne, C., Mouzat, K., Pinto, D., Howe, J., Lemière, N., Durand, C. M., Guibert, J., Ey, E., Toro, R., Peyre, H., Mathieu, A., … Bourgeron, T. (2014). Meta-analysis of shank mutations in autism spectrum disorders: A gradient of severity in cognitive impairments. *PLoS Genetics*, *10*(9). https://doi.org/10.1371/journal.pgen.1004580

Lee A.M., Tai, L-H., Zador, A., & Wilbrecht,L. (2015) Between the primate and 'reptilian' brain: Rodent models demonstrate the role of corticostriatal circuits in decision making. Neuroscience, 296, 66-74. https://doi.org/10.1016/j.neuroscience.2014.12.042

Li, J.-J., Shi, C., Li, L., & Collins, A. G. E. (2024). Dynamic noise estimation: A generalized method for modeling noise fluctuations in decision-making. *Journal of Mathematical Psychology*, *119*, 102842. https://doi.org/10.1016/j.jmp.2024.102842

Maenner, M. J., Warren, Z., Williams, A. R., Amoakohene, E., Bakian, A. V., Bilder, D. A., Durkin, M. S., Fitzgerald, R. T., Furnier, S. M., Hughes, M. M., Ladd-Acosta, C. M., McArthur, D., Pas, E. T., Salinas, A., Vehorn, A., Williams, S., Esler, A., Grzybowski, A., Hall-Lande, J., … Shaw, K. A. (2023). Prevalence and characteristics of autism spectrum disorder among children aged 8 years — autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2020. *MMWR. Surveillance Summaries*, *72*(2), 1–14. https://doi.org/10.15585/mmwr.ss7202a1

Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2012). Sex differences in autism spectrum disorder: Evidence from a large sample of children and adolescents. *Journal of Autism and Developmental Disorders*, *42*(7), 1304–1313. https://doi.org/10.1007/s10803-011-1356-0

Master, S. L., Eckstein, M. K., Gotlieb, N., Dahl, R., Wilbrecht, L., & Collins, A. G. E. (2020). Disentangling the systems contributing to changes in learning during adolescence. *Developmental Cognitive Neuroscience*, *41*, 100732. https://doi.org/10.1016/j.dcn.2019.100732

Mei, Y., Monteiro, P., Zhou, Y., Kim, J.-A., Gao, X., Fu, Z., & Feng, G. (2016). Adult restoration of Shank3 expression rescues selective autistic-like phenotypes. *Nature*, *530*(7591), 481–484. https://doi.org/10.1038/nature16971

Murta, V., Seiffe, A., & Depino, A. M. (2023). Sex differences in mouse models of autism spectrum disorders: Their potential to uncover the impact of brain sexual differentiation on gender bias. *Sexes*, *4*(3), 358–391. https://doi.org/10.3390/sexes4030024

Muscatello, R. A., Rafatjoo, E., Mirpuri, K. K., Kim, A., Vandekar, S., & Corbett, B. A. (2022). Salivary testosterone in male and female youth with and without autism spectrum disorder: Considerations of development, sex, and diagnosis. *Molecular Autism*, *13*(1). https://doi.org/10.1186/s13229-022-00515-4

Naisbitt, S., Kim, E., Tu, J. C., Xiao, B., Sala, C., Valtschanoff, J., Weinberg, R. J., Worley, P. F., & Sheng, M. (1999). Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and Cortactin. *Neuron*, *23*(3), 569–582. https://doi.org/10.1016/s0896-6273(00)80809-0

Nussenbaum, K., & Hartley, C. A. (2019). Reinforcement learning across development: What insights can we draw from a decade of research? *Developmental Cognitive Neuroscience*, *40*, 100733. https://doi.org/10.1016/j.dcn.2019.100733

Peça, J., Feliciano, C., Ting, J. T., Wang, W., Wells, M. F., Venkatraman, T. N., Lascola, C. D., Fu, Z., & Feng, G. (2011). Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature*, *472*(7344), 437–442. https://doi.org/10.1038/nature09965

Peixoto, Rui T, Wang, W., Croney, D. M., Kozorovitskiy, Y., & Sabatini, B. L. (2016). Early hyperactivity and precocious maturation of corticostriatal circuits in Shank3b−/− mice. *Nature Neuroscience*, *19*(5), 716–724. https://doi.org/10.1038/nn.4260

Peixoto, R. T., Chantranupong, L., Hakim, R., Levasseur, J., Wang, W., Merchant, T., Gorman, K., Budnik, B., & Sabatini, B. L. (2019). Abnormal striatal development underlies the early onset of behavioral deficits in Shank3b mice. *Cell Reports*, *29*(7). https://doi.org/10.1016/j.celrep.2019.10.021

Peñagarikano, O., Abrahams, B. S., Herman, E. I., Winden, K. D., Gdalyahu, A., Dong, H., Sonnenblick, L. I., Gruver, R., Almajano, J., Bragin, A., Golshani, P., Trachtenberg, J. T., Peles, E., & Geschwind, D. H. (2011). Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits. *Cell*, *147*(1), 235–246. https://doi.org/10.1016/j.cell.2011.08.040

Peters, A. J., Fabre, J. M., Steinmetz, N. A., Harris, K. D., & Carandini, M. (2021). Striatal activity topographically reflects cortical activity. *Nature*, *591*(7850), 420–425. https://doi.org/10.1038/s41586-020-03166-8

Piekarski, D. J., Boivin, J. R., & Wilbrecht, L. (2017). Ovarian hormones organize the maturation of inhibitory neurotransmission in the frontal cortex at puberty onset in female mice. *Current Biology*, *27*(12). https://doi.org/10.1016/j.cub.2017.05.027

Platt, R. J., Zhou, Y., Slaymaker, I. M., Shetty, A. S., Weisbach, N. R., Kim, J.-A., Sharma, J., Desai, M., Sood, S., Kempton, H. R., Crabtree, G. R., Feng, G., & Zhang, F. (2017). Chd8 Mutation Leads to Autistic-like Behaviors and Impaired Striatal Circuits. *Cell Reports*, *19*(2), 335–350. https://doi.org/10.1016/j.celrep.2017.03.052

Qiu, T., Chang, C., Li, Y., Qian, L., Xiao, C. Y., Xiao, T., Xiao, X., Xiao, Y. H., Chu, K. K., Lewis, M. H., & Ke, X. (2016). Two years changes in the development of caudate nucleus are involved in restricted repetitive behaviors in 2–5-year-old children with autism spectrum disorder. *Developmental Cognitive Neuroscience*, *19*, 137–143. https://doi.org/10.1016/j.dcn.2016.02.010

Rendall, A. R., Perrino, P. A., Buscarello, A. N., & Fitch, R. H. (2019). Shank3B mutant mice display pitch discrimination enhancements and learning deficits. *International Journal of Developmental Neuroscience*, *72*(1), 13–21. https://doi.org/10.1016/j.ijdevneu.2018.10.003

Reinhold, K., Iadarola, M., Tang, S., Kuwamoto, W., Sun, S., Hakim, R., Zimmer, J., Wang, W., & Sabatini, B. L. (2023). Striatum Supports Fast Learning but Not Memory Recall. *bioRxiv*. [Preprint] https://doi.org/10.1101/2023.11.08.566333 (accessed 11 January 2025).

Rothwell, Patrick E., Fuccillo, Marc V., Maxeiner, S., Hayton, Scott J., Gokce, O., Lim, B., Fowler, Stephen C., Malenka, Robert C., & Südhof, Thomas C. (2014). Autism-Associated Neuroligin-3 Mutations Commonly Impair Striatal Circuits to Boost Repetitive Behaviors. *Cell*, *158*(1), 198–212. https://doi.org/10.1016/j.cell.2014.04.045

Rothwell, P. E. (2016). Autism spectrum disorders and drug addiction: Common pathways, common molecules, distinct disorders? *Frontiers in Neuroscience*, *10*. https://doi.org/10.3389/fnins.2016.00020

Sala, C., Piëch, V., Wilson, N. R., Passafaro, M., Liu, G., & Sheng, M. (2001). Regulation of Dendritic Spine Morphology and Synaptic Function by Shank and Homer. *Neuron*, *31*(1), 115–130. https://doi.org/10.1016/s0896-6273(01)00339-7

Schuck, R. K., Flores, R. E., & Fung, L. K. (2019). Brief report: Sex/gender differences in symptomology and camouflaging in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *49*(6), 2597–2604. https://doi.org/10.1007/s10803-019-03998-y

Schuetze, M., Park, M. T., Cho, I. Y., MacMaster, F. P., Chakravarty, M. M., & Bray, S. L. (2016). Morphological alterations in the thalamus, striatum, and pallidum in autism spectrum disorder. *Neuropsychopharmacology*, *41*(11), 2627–2637. https://doi.org/10.1038/npp.2016.64

Solomon, M., Smith, A. C., Frank, M. J., Ly, S., & Carter, C. S. (2011). Probabilistic reinforcement learning in adults with autism spectrum disorders. *Autism Research*, *4*(2), 109–120. https://doi.org/10.1002/aur.177

South, M., Chamberlain, P. D., Wigham, S., Newton, T., Le Couteur, A., McConachie, H., Gray, L., Freeston, M., Parr, J., Kirwan, C. B., & Rodgers, J. (2014). Enhanced decision making and risk avoidance in high-functioning autism spectrum disorder. *Neuropsychology, 28*(2), 222–228. https://doi.org/10.1037/neu0000016

Tang, H., Costa, V. D., Bartolo, R., & Averbeck, B. B. (2022). Differential coding of goals and actions in ventral and dorsal corticostriatal circuits during goal-directed behavior. *Cell Reports*, *38*(1), 110198. https://doi.org/10.1016/j.celrep.2021.110198

Thorn, C. A., Atallah, H., Howe, M., & Graybiel, A. M. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron*, *66*(5), 781–795. https://doi.org/10.1016/j.neuron.2010.04.036

Tsai, P. T., Hull, C., Chu, Y., Greene-Colozzi, E., Sadowski, A. R., Leech, J. M., Steinberg, J., Crawley, J. N., Regehr, W. G., & Sahin, M. (2012). Autistic-like behaviour and cerebellar dysfunction in Purkinje cell TSC1 mutant mice. *Nature*, *488*(7413), 647–651. https://doi.org/10.1038/nature11310

Wang, W., Li, C., Chen, Q., van der Goes, M.-S., Hawrot, J., Yao, A. Y., Gao, X., Lu, C., Zang, Y., Zhang, Q., Lyman, K., Wang, D., Guo, B., Wu, S., Gerfen, C. R., Fu, Z., & Feng, G. (2017). Striatopallidal dysfunction underlies repetitive behavior in Shank3-deficient model of autism. *Journal of Clinical Investigation*, *127*(5), 1978–1990. https://doi.org/10.1172/jci87997

Wilbrecht, L., & Davidow, J. Y. (2024). Goal-directed learning in adolescence: Neurocognitive Development and contextual influences. *Nature Reviews Neuroscience*, *25*(3), 176–194. https://doi.org/10.1038/s41583-023-00783-w

Wilson, R.C., & Collins, A.G.E. (2019). Ten simple rules for the computational modeling of behavioral data *eLife* 8:e49547. https://doi.org/10.7554/eLife.49547

Znamenskiy, P., Zador, A. (2013) Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* 497, 482−485. https://doi.org/10.1038/nature12077

Figures

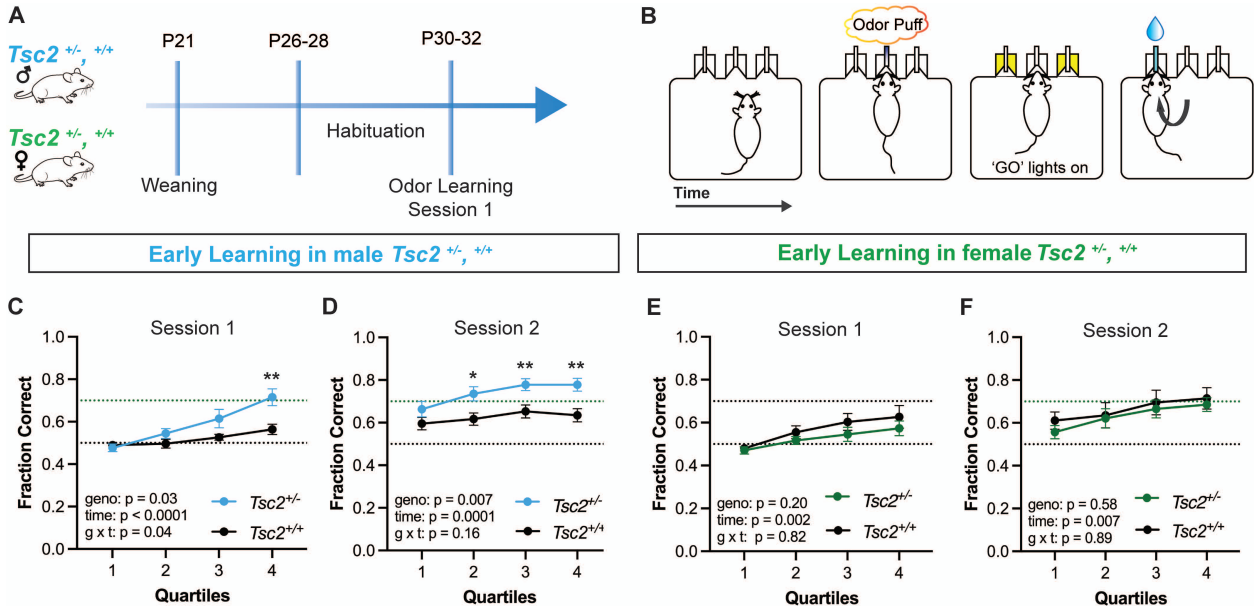


Figure 1: Adolescent male, but not female *Tsc2* Het mice show a gain of function in learning in early training in an odor based 2AFC task. (A) Timeline of training. One week after weaning, male and female *Tsc2* Het and WT animals were habituated to the ports and chamber by shaping them to first poke into the center port and then the peripheral ports to receive water rewards. Odor cues were not present during habituation. At P30-P32 mice began session 1 of the odor-based 2AFC task. (B) In this task, mice self-initiate a trial by nose poke into the center port, which releases one of two odor cues. Bilateral "Go" lights then indicate the availability of reward, prompting mice to choose a peripheral side port. Choosing the correct port for that odor, results in a 2μL water reward. (C-D) Performance in Male *Tsc2* Het mice was significantly greater than WT males in both the first and second session of odor learning. This was captured by a significant main effect of genotype in both sessions and a genotype x time interaction in session 1. Significant Tukey post-hoc tests between Het and WT for each quartile are indicated by asterisks. (E-F) We also tested female *Tsc2* Het and WT mice in the same odor-based 2AFC task on the same schedule. Female *Tsc2* Het and WT did not show differences in learning performance for Session 1 or Session 2. Full statistical results are reported in the main text. Male *Tsc2* Het, n = 15; WT, n = 13; Female *Tsc2* Het, n = 10; WT, n = 10. Learning curves represent mean ± SEM. *p<0.05, **p<0.01.

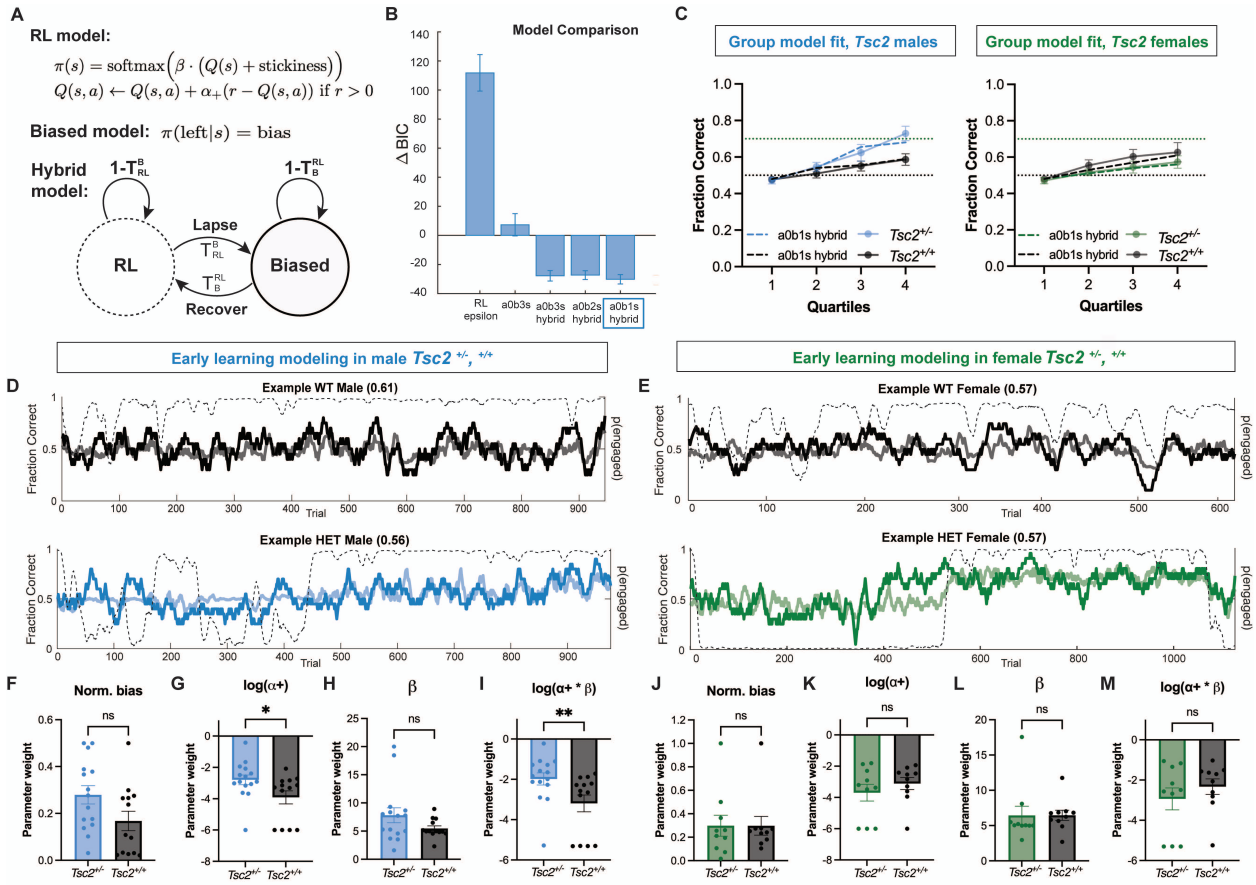


Figure 2: Adolescent male, but not female, *Tsc2* Het mice have a higher learning rate when fit with the winning hybrid model. We examined a series of models that use a reinforcement learning (RL) framework to account for differences in choice behavior between genotypes (see Materials & Methods). (A) Several of the candidate models that were tested were mixture models that combined a standard Q-learning RL model, representing an engaged state, with a biased model, representing a disengaged state, to account for periods of task disengagement at the policy level. The transition probabilities of these separate policy states are governed by a hidden Markov model. (B) Using Bayesian Information Criterion (BIC), we determined that models with the mixture policy (named hybrid here) had a lower BIC score. The change in BIC is calculated by the difference in each model's BIC and the mean BIC of all models examined. All hybrid models were comparable, but the model with the lowest BIC (a0b1s_hybrid, blue box) contained a single α positive learning rate, a single inverse temperature parameter, □, and a single strategy parameter,ξ_1, that indicated the likelihood for an animal to stick with their previous choice if they had received a reward. The full winning model is written in the schematic in (A). (C) We generated data by simulating the model with parameters fit on individual sessions. Then, we inspected the group similarity between simulated data (blue dotted line for male *Tsc2* Het mice and green dotted line for female *Tsc2* Het mice) and found that the model simulated with fit model parameters captures averaged mouse learning curve data. (D-E) Examples of individual trial-by-trial model fits for both male (D) and female mice (E). The blue for males (D) and the green for females (E) indicate the fraction correct performance for session 1 for two example animals from each sex. The lighter, transparent line indicates the simulated data from the winning hybrid model and the thin black dashed line line represents the inferred probability that the mouse is in the latent "engaged" state, i.e. following the RL policy, as opposed to the biased state. (F) There was no significant difference in normalized bias between male Het and WT groups. (G) There was a significant difference between *Tsc2* Het and WT males in learning rate α+

(plotted in log scale for visibility), Mann-Whitney test, U =54, p = 0.04, as well as α+ *□, a post-hoc variable that captures the overall impact of an outcome on the policy, (I) Mann-Whitney test, U = 41, p = 0.008. There was no difference in the □ parameter (H) in males. In females, there were no differences between *Tsc2* Het and WT mice in normalized bias (J), log(α)+ (K), □ (L), or log(α+* □) (M). Full statistical results are reported in the main text. Error bars represent SEM. *p<0.05, **p<0.01.

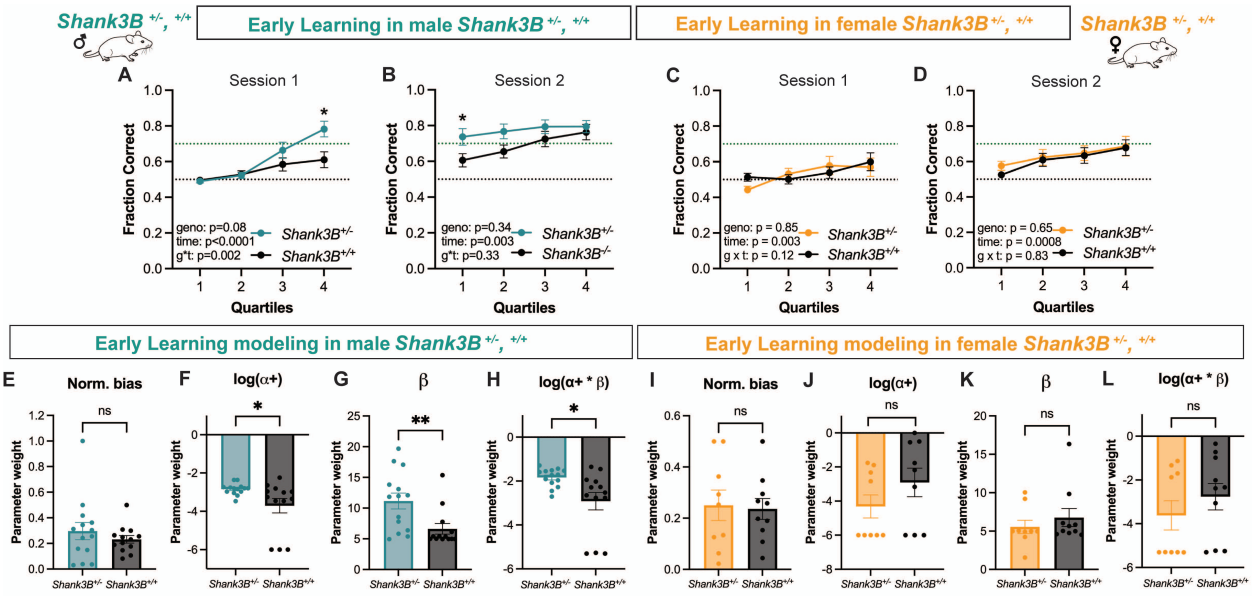


Figure 3: Adolescent male, but not female *Shank3B* Het mice show a gain of function in learning and a higher learning rate during early training in an odor based 2AFC task. (A) Male *Shank3B Het* showed stronger learning performance than WT littermates in the first training session, Session 1. This was captured by a significant main effect of time and genotype x time interaction. A Tukey post-hoc comparison test revealed a significant difference between Het and WT performance in the fourth quartile, indicated by the asterisk (p = 0.01). (B) There was a main effect of time but no genotype or genotype x time interaction in Session 2. A Tukey post-hoc comparison revealed a significant difference between WT and Het mice in the first quartile (p = 0.03) We also tested female *Tsc2* Het and WT mice in the same odor-based 2AFC task on the same schedule. (C-D) Female Het and WT mice did not differ in learning performance for Session 1 or for Session 2 as evidenced by lack of significant main effect of genotype or an interaction. We next fit computational models to *Shank3B* male and female early learning data. Male *Shank3B* Hets and WT did not differ in normalized bias parameter (E), but did differ significantly in α+ (F), □ (G), and the α+*□ variable (H). Female *Shank3B* Hets and WT did not differ significantly for any modeling parameter: (I) normalized bias parameter, (J) α+, (K) □, (L) α+*□. Full statistical results are reported in the main text. Male *Shank3B* Het, n = 12; Male *Shank3B* WT, n = 13; Female *Shank3B* Het, n = 11; Female *Shank3B* WT, n = 11. Error bars represent SEM. *p<0.05, **p<0.01.

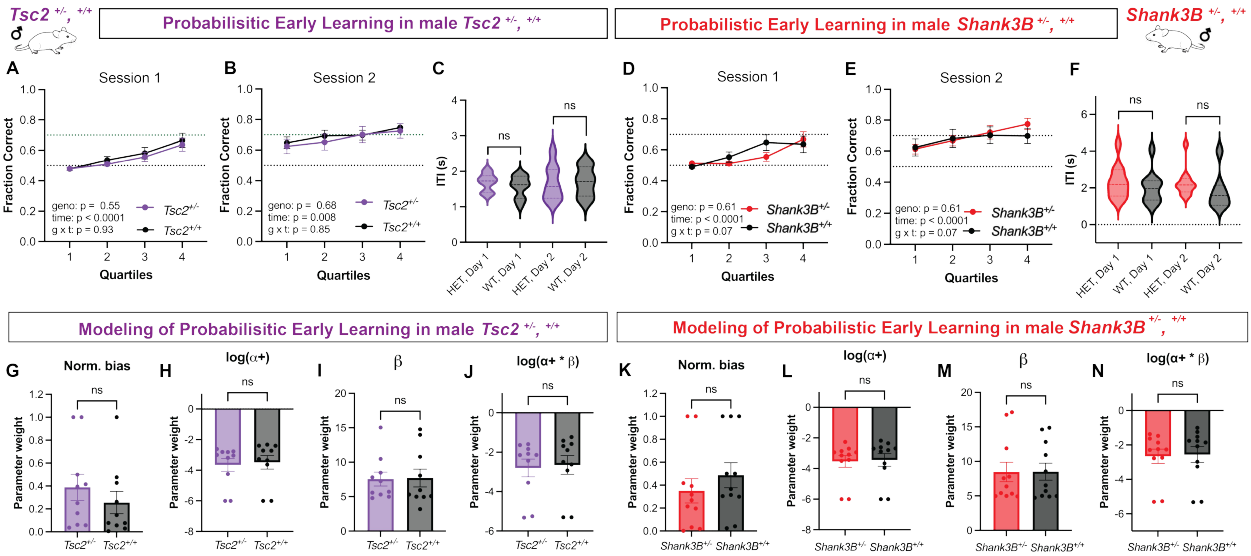


Figure 4: Gain of function in adolescent male *Tsc2* and *Shank3B* Het is no longer apparent when reward schedule is probabilistic. We tested male *Tsc2* (A-C) and male *Shank3B* (D-F) in a probabilistic version of the odor-based 2AFC task where 10-20% of correct choices in session 1 and 2 were not rewarded (compared to Figure 1-3 data where 100% of correct choices were rewarded). Male *Tsc2* Het and WT showed comparable learning performance in this probabilistic context in both session 1 and 2 (A,B). ITI speeds were also comparable (C). (D,E) Similarly, in male *Shank3B* mice, Het and WT learning performance was comparable between genotypes for both sessions. (F) There was also no significant difference between groups in ITI for either session. We next fit our series of models to probabilistic learning trial-by-trial data. We found that there were no significant differences between model parameters in (G-J) male *Tsc2* Het and WT mice nor between (K-N) male *Shank3B* Het and WT mice when performing in the probabilistic task context. Full statistical results are reported in the main text. *Tsc2* Het, n = 10; *Tsc2* WT, n = 10; *Shank3B* Het, n = 11; *Shank3B* WT, n = 11. Error bars show SEM. Violin plots show quartiles (25% and 75%), mean, and range.

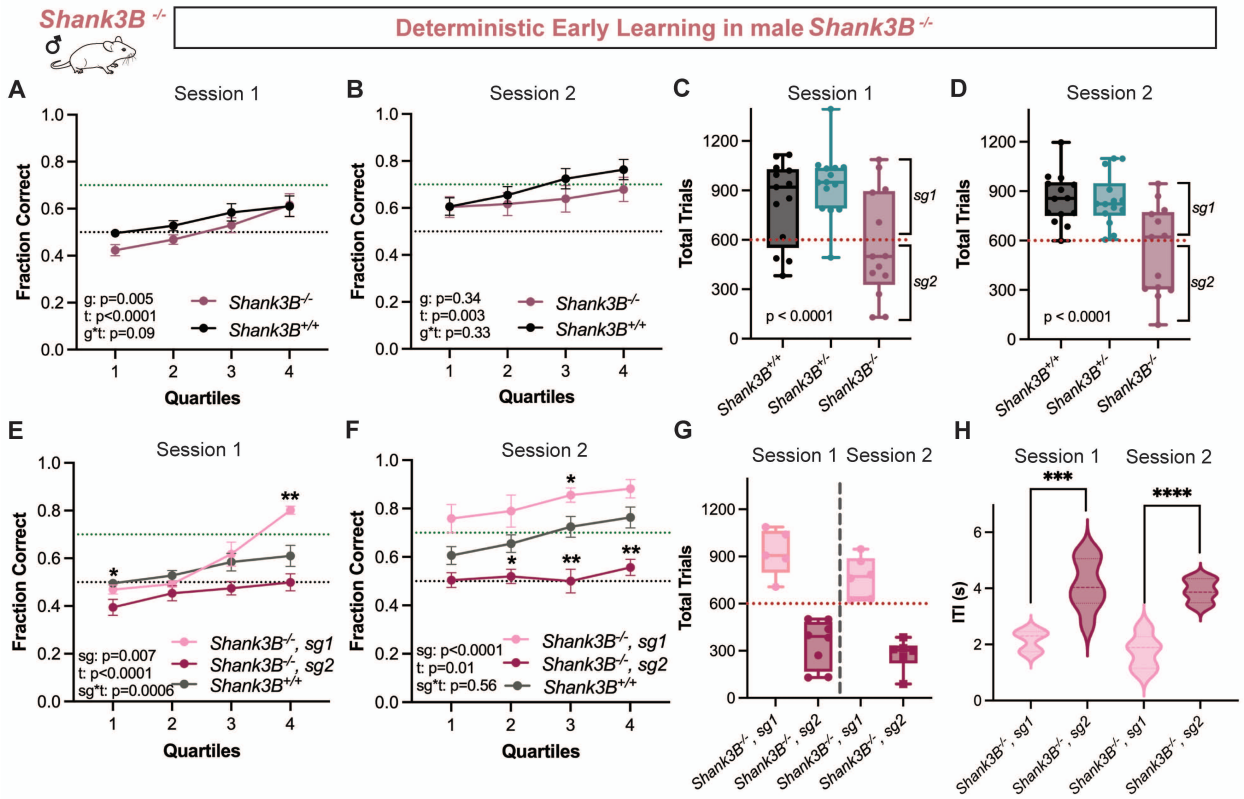
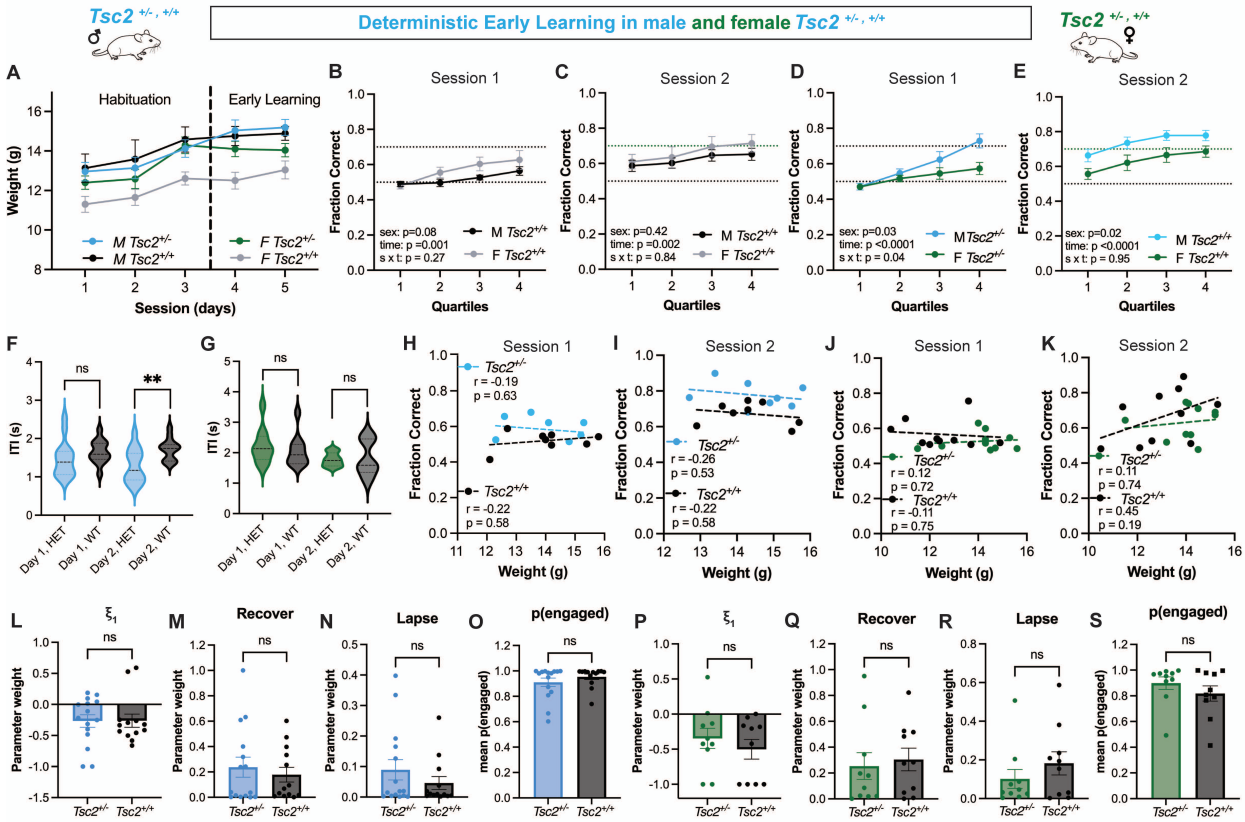
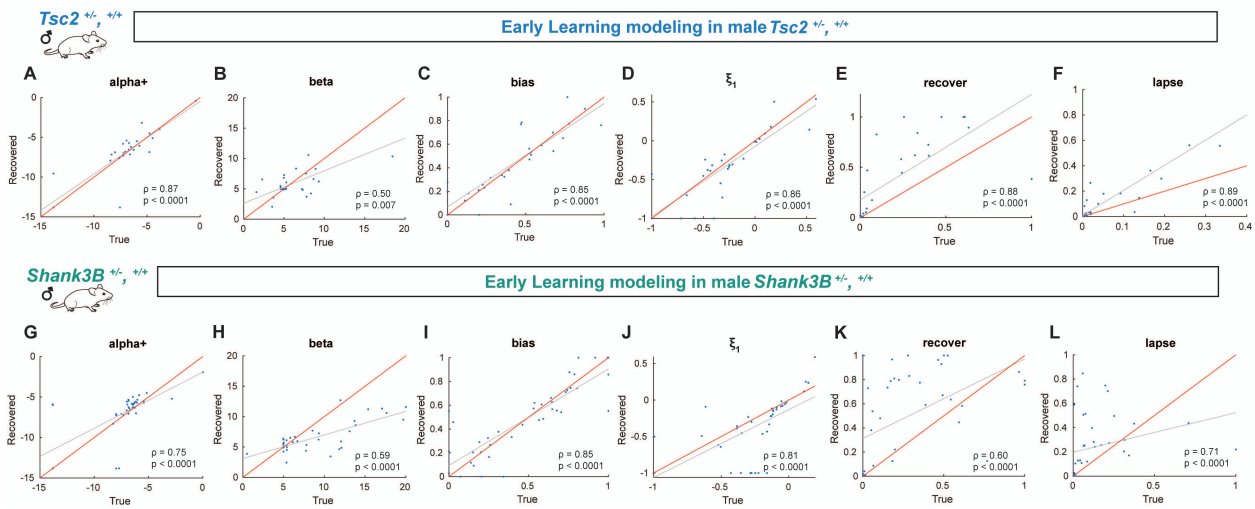


Figure 5: Adolescent male *Shank3B* KO mice exhibit bimodal levels of trial completion and the more active sub-group shows gain of function in learning performance. (A-B) The full group of male *Shank3B* KO mice (n=13) showed significantly lower learning performance than male *Shank3B* WT animals (n=13) in Session 1. In Session 2, KO performance reached levels comparable to WT. (C,D) Looking at total trials completed in each session, we noted that the total number of completed trials was bimodal for KO mice. Total trials completed were comparable to WT and Het siblings for one sub-group of KO (n=5) (sg1), but there was also a second sub-group (sg2) (n=8) that completed substantially fewer trials per session (see also G). Box and whisker plots show 25%, mean, 75%, and range with individual points indicating total number of trials for each individual mouse. (E-H) Based on these differences we decided to split the KO group into two sub-groups and compare their learning to WT. (E,F) Sg1 KOs, who completed a typical range of trials, showed significantly better learning performance than WT in the odor based 2AFC in both Session 1 and Session 2. Sg2 KOs, who completed substantially fewer trials, performed at chance levels, significantly lower than WT in both sessions. Asterisks indicate post-hoc comparisons of sub-groups to WT. In Session 1, quartile one, sg2 KOs had significantly lower performance than WT ($p = 0.03$) and in quartile four, sg1 KOs had significantly higher performance than WT ($p = 0.002$). In Session 2, sg2 KOs had significantly lower performance than WT in quartile two ($p = 0.01$), three ($p = 0.007$), and four ($p = 0.002$). Sg1 outperformed WT in quartile three ($p = 0.04$). (G) Trials completed for each sub group in box and whisker plots, indicating range, mean, and 25%, 75% quartiles. (H) Examination of intertrial interval movement times showed that sg2 was also significantly slower to move between ports than sg1 in both sessions. Violin plots show quartiles (25% and 75%), mean, and range. Error bars represent SEM. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

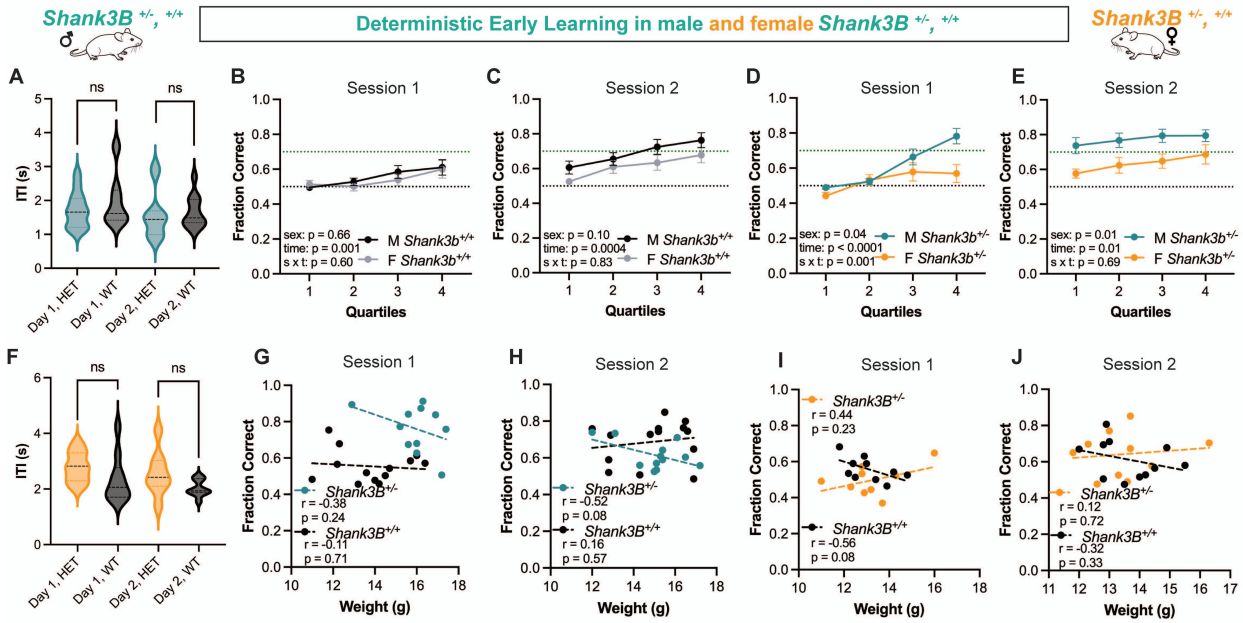
Supplemental Information:



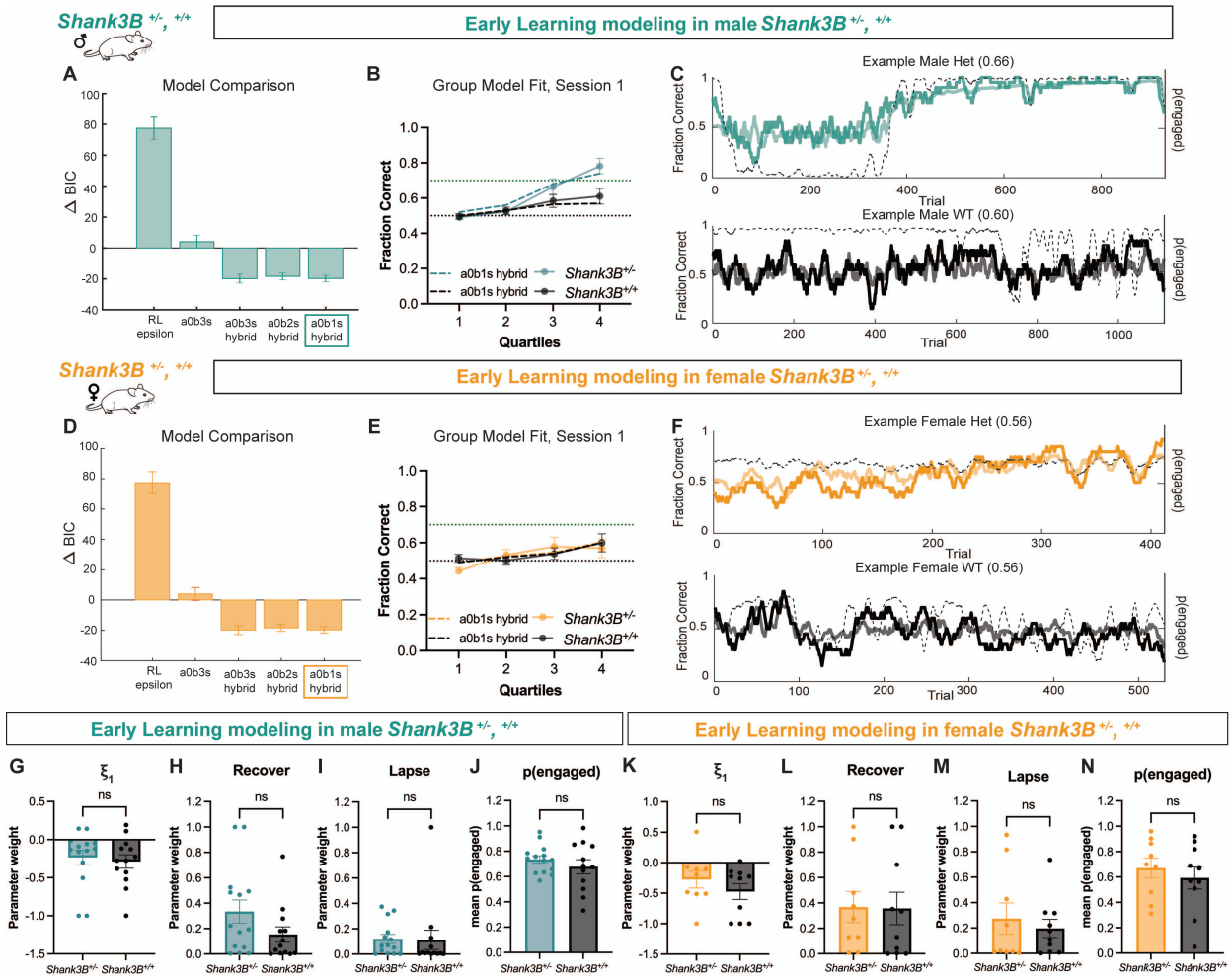
Supplemental Figure 1: Additional analyses for *Tsc2* mice trained in the 2AFC with a deterministic schedule show sex differences in learning in Het, but not WT mice. Additional modeling parameters to support the main text (Fig. 1 and Fig. 2) are also shown. (A) There was a significant gain in weight across habituation and early learning for all groups of animals. Mixed-effect model, using REML, fixed effect of time: $F(2.59, 102.3)=36.50$, $p < 0.0001$; fixed effect of group: $F(3,40)=4.30$, $p = 0.01$. (B-C) Comparison of male (black) and female (grey) *Tsc2* WT animals in Session 1 and in Session 2 revealed no differences in performance or interaction. (D-E) However, there was a significant main effect of sex between male and female *Tsc2* Het for both Session 1 and Session 2. (F-G) Group average ITI comparisons for Session 1 and Session 2 for male *Tsc2* Hets (blue) and female *Tsc2* Hets (green) compared to sex-matched WTs (black) with pairwise comparison results described in the text. Violin plots show quartiles (25% and 75%), mean, and range. (H-I) Performance in Session 1 and 2 for male *Tsc2* Het or WT was not correlated with weight, using Pearson's or Spearman's correlation dependent on normality tests. (J-K) Neither *Tsc2* WT or Het female mice had performance in the task that was correlated with weight, using Pearson's or Spearman's correlation dependent on normality tests. (L-N) Male *Tsc2* Het and WT groups showed no significant differences in additional model parameters, ξ_1, recover, and lapse, all fit from the winning model. Others are reported in main figures and text. (O) In a post-hoc analysis, male *Tsc2* Het and WT animals showed a similar level of mean engagement throughout the task. (P-R) Female *Tsc2* Het and WT mice showed no significant group differences in additional model parameters (ξ_1, recover, and lapse) fit from the winning model. Others are reported in main figures and text. (S) In a post-hoc analysis, female *Tsc2* Het and WT animals showed a similar level of mean engagement throughout the task. See results section of main text for full statistics. Error bars represent SEM.



Supplemental Figure 2: Parameters from winning model, a0bs1 hybrid, are identifiable. We tested a series of models that explored which strategies mice may use to learn the odor-based 2AFC task. The winning model (as determined by BIC, see Figure 2 and Supplemental Figures 4 and 5) had 6 free parameters that included an α+ (learning from positive outcomes), an inverse temperature parameter □, a bias parameter that described an animal's likelihood to continuously choose one side, a stickiness parameter, ξ_1, that indicated the animal's likelihood to repeat the previous trial's action, and two parameters, lapse and recover, that marked the transition probabilities between an engaged and a disengaged (biased) state. In order to validate the winning model we generated data by simulating the model with parameters fit to individual sessions and fit the simulated data to obtain recovered parameters. We found that generated parameters (x-axis) were highly correlated with recovered parameters (y-axis) for both *Tsc2* (A-F) and *Shank3B* (G-L) male Session 1 data. This correlation indicates that the model parameters are identifiable (Wilson & Collins, 2019).

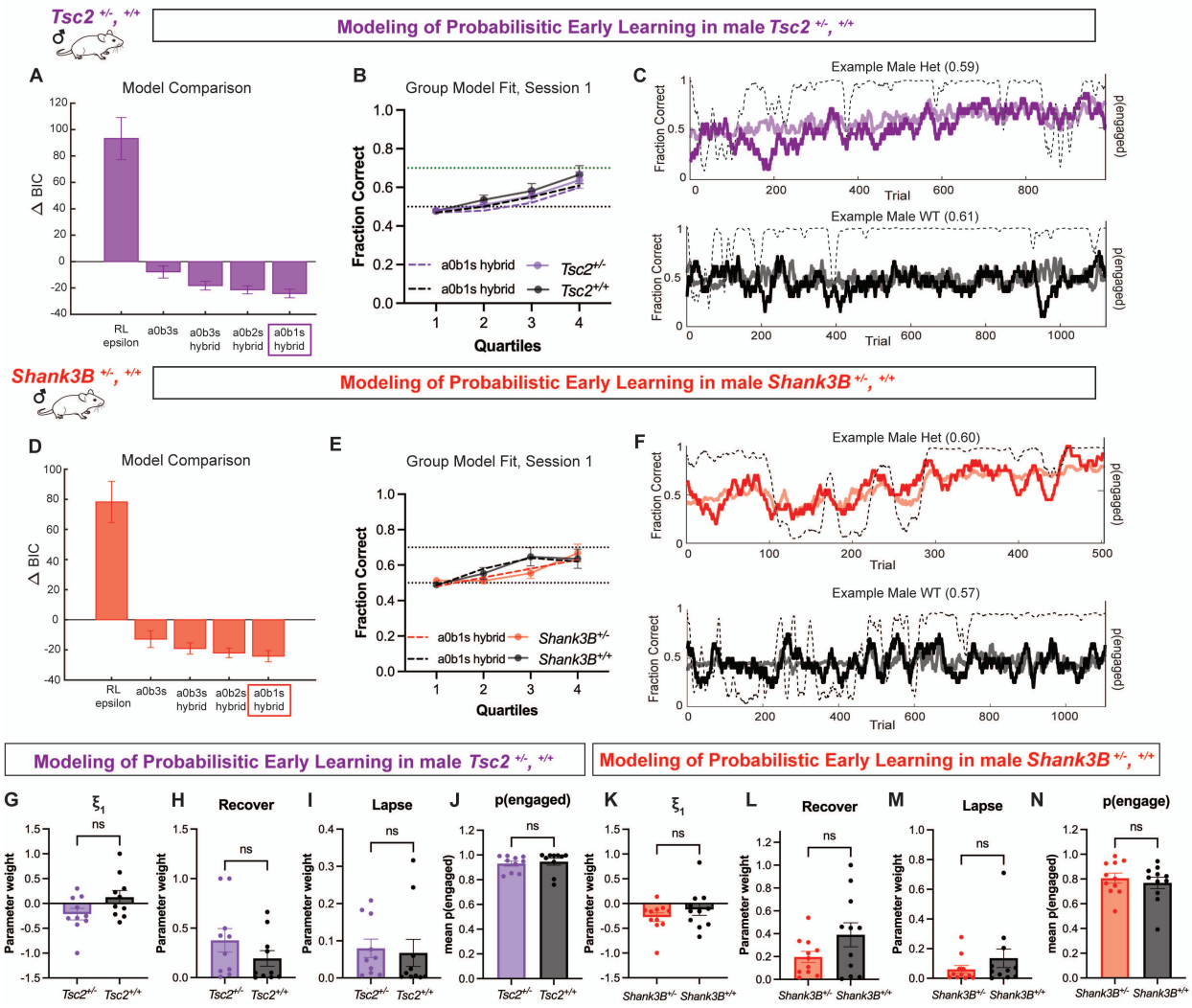


Supplemental Figure 3: Additional analyses of male and female *Shank3B* mice trained in the 2AFC with a deterministic schedule show sex differences in learning in Het, but not WT mice. (A) **A**verage ITI was comparable between Het (teal) and WT (black) in Session 1 and Session 2. (B) Comparison of male and female *Shank3B* WT animals in Session 1 and in Session 2 (C) revealed no sex differences in learning performance. However, when comparing male and female *Shank3B* Hets, there was a significant difference between male and female animals in session 1(D) and session 2 (E) learning performance. (F) Average ITI was comparable for female *Shank3B* Het (yellow) and WT (black) mice in Session 1 and Session 2. There were no significant correlations between fraction correct performance and weight in Session 1 and Session 2 for males (G-H) or for females (I-J).



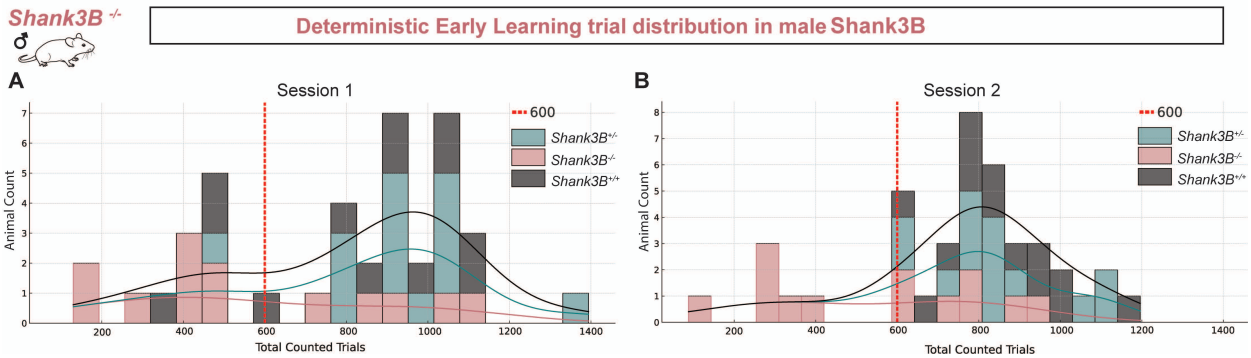
Supplemental Figure 4: Additional data from the male and female *Shank3B* Het and WT mice trained in the 2AFC with a deterministic schedule: model comparison, validation, and additional fit parameters. (A,D) We compared a series of reinforcement learning and mixture models that can account for periods of task disengagement at the policy level. Using Bayesian Information Criterion (BIC), we determined that models that contained the mixture policy (hybrid) had a lower BIC score. The change in BIC is calculated by the difference in each model's BIC and the mean BIC of all models examined. All hybrid models were comparable, but the model with the lowest BIC (colored square) was a0b1s_hybrid for both males (A) and female (D) *Shank3B.* (B,E) We generated data by simulating the model with parameters fit on individual sessions. Then, we inspected the group similarity between simulated data (teal dotted line for male *Shank3B* Het and yellow dotted line for female *Shank3B* Het) and found that the model simulated with fit model parameters captures group mouse learning curve data. (C,F) Examples of individual trial-by-trial model fits for both males (C) and females (F) with the teal or yellow line reflecting male or female trial-by-trial performance, respectively, the light teal/yellow transparent line reflects the model fit and the thin black dashed line is the likelihood that the animal is in an engaged state. (G-I) Male *Shank3B* Het and WT mice showed no significant differences in additional model parameters (ξ_1, recover, and lapse) fit from the winning model, pairwise comparison results described in main text results. (J) In a post-hoc analysis, male *Shank3B* Het and WT animals showed a similar level of mean engagement throughout the task. (K-M) Female *Shank3B* Het and WT mice showed no significant differences in additional model parameters (ξ_1, recover, and lapse) fit from the winning model, pairwise comparison

results described in main text results. (N) In a post-hoc analysis, female *Shank3B* Het and WT animals showed a comparable level of mean engagement throughout the task. Error bars throughout represent SEM.

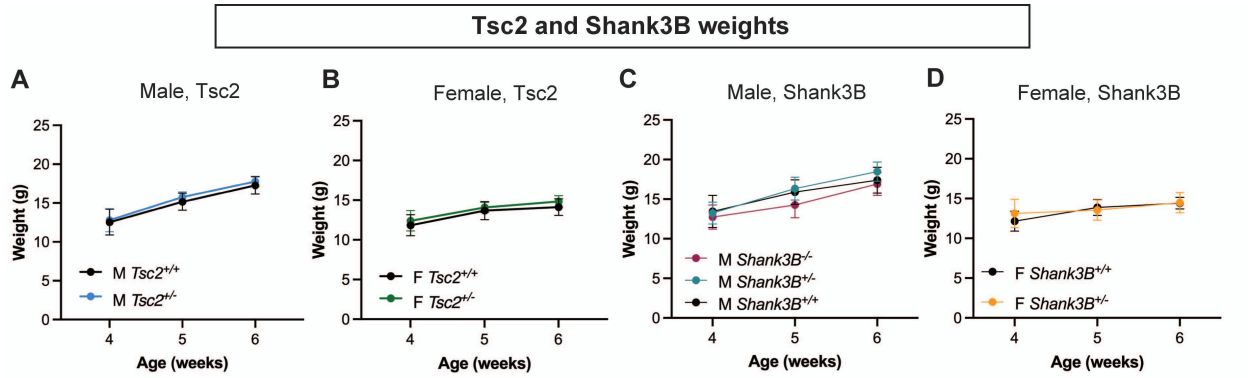


Supplemental Figure 5: Additional data from the male *Tsc2* and male *Shank3B* Het and WT mice trained in the 2AFC with a probabilistic schedule: model comparison, validation, and additional fit parameters. (A,D) We examined a series of mixture models that can better account for periods of task disengagement at the policy level. Using Bayesian Information Criterion (BIC), we determined that models that contained the mixture policy (hybrid) had a lower BIC score, as calculated by the difference in each model's BIC and the mean BIC of all models examined. All hybrid models were comparable, but the model with the lowest BIC (colored square) was a0b1s_hybrid for both male *Tsc2* (A) and male *Shank3B* (D) mice tested in the probabilistic schedule (where 10-20% of correct trials were not rewarded). This was the same winning model found above for the deterministic schedule (when 100% of correct trials rewarded) (B,E) We generated data by simulating the model with parameters fit on individual sessions. Then, we inspected the group similarity between simulated data (purple dotted line for male *Tsc2* Hets and red dotted line for male *Shank3B* Hets) and found that the model simulated with fit model parameters captures group mouse learning curve data. (C,F) Examples of individual trial-by-trial model fits for both *Tsc2* Het males (C) and *Shank3B* Het males (F) with the purple or red lines reflecting male *Tsc2* or *Shank3B* trial-by-trial performance, respectively, the light/transparent purple and red lines reflects the

model fit and the thin black dashed line is the likelihood that the animal is in an engaged state. (G-I) Male *Tsc2* Het and WT mice had no differences in additional model parameters (ξ_1, recover, and lapse) fit from the winning model, pairwise comparison results described in text. (J) In a post-hoc analysis, male *Tsc2* Het and WT animals showed a similar level of mean engagement throughout the task. (K-M) Male *Shank3B* Het and WT mice had no differences in additional model parameters (ξ_1, recover, and lapse) fit from the winning model, pairwise comparison results described in main text. (N) In a post-hoc analysis, male *Shank3B* Het and WT animals showed a similar level of mean engagement throughout the task. Error bars represent SEM.

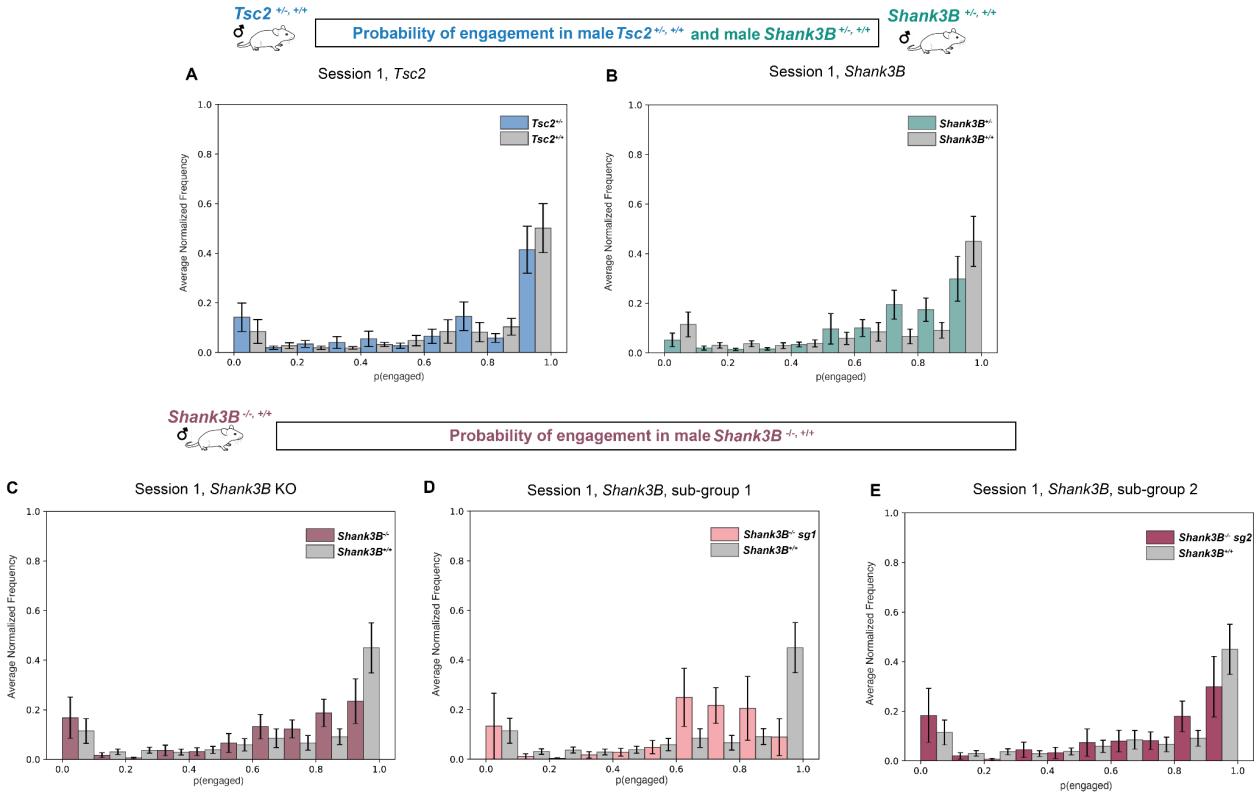


Supplemental Figure 6: Distribution of trial counts for deterministic early learning in male Shank3B mice. (A-B) Histograms showing the total counted trials for male *Shank3B* WT (black), Het (teal), and KO (salmon) mice across (A) Session 1 and (B) Session 2 of the deterministic task. Data are binned by 50 trials. Overlaid density curves represent the distribution within each group. The dashed red line indicates the 600-trial threshold. Differences in trial distribution highlight variation in trials completed among genotypes.



Supplemental Figure 7: Weight trajectories in male and female *Tsc2* and *Shank3B* mice across genotype. We examined whether there was any difference between WT and Het weights for *Tsc2* and *Shank3B* mice across weeks 4, 5, and 6. Statistical comparisons were made using a 2-way ANOVA with factors for group (genotype) and week (4, 5, and 6), including the interaction between group and week. (A) There was no significant group difference between *Tsc2* WT and Het males (main effect of group: $F(1, 58) = 2.37$, $p = 0.12$; interaction: $F(2, 58) = 0.10$, $p = 90$), but there was a main effect of time ($F(2, 58) = 90.3$, $p < 0.0001$) indicating weight gain over time. (B) For female *Tsc2* WT and Hets, there was also no significant group difference or interaction (main effect of group: $F(1, 50) = 3.92$, $p = 0.05$; interaction: $F(2, 50) = 0.11$, $p = 0.89$). However, there was a significant main effect of time ($F(2, 50) = 23.33$, $p < 0.0001$),

reflecting weight gain across weeks. (C) There was a significant main effect of genotype for male *Shank3B* mice ($F(2, 105) = 7.78$, $p = 0.0007$). Post-hoc Tukey tests showed that KO mice weighed significantly less than Het mice at Weeks 5 and 6 ($p < 0.05$) and less than WT mice at Week 5 ($p < 0.05$). No significant differences were observed between WT and Het at any time point. A significant main effect of time ($F(2, 105) = 78.10$, $p < 0.0001$) indicated weight increases across weeks, consistent across genotypes (no significant interaction: $F(4, 105) = 1.23$, $p = 0.3018$). (D) For female WT and Het *Shank3B* mice there was significant weight gain across weeks (main effect of time: $F(2, 60) = 11.47$, $p < 0.0001$) but there was no significant difference in weight or weight gain between genotypes (main effect of genotype: $F(1, 60) = 0.55$, $p = 0.46$; genotype and time interaction: $F(2, 60) = 1.60$, $p = 0.20$). Data represents mean ± SEM.



Supplemental Figure 8: In Session 1 of the deterministic version of the 2AFC task, male *Tsc2* Het, *Shank3B* Het, and *Shank3B* KO mice showed similar distributions of engagement probability when compared to WT counterparts. To quantitatively assess differences in group distributions of p(engaged), we computed the Kullback-Leibler (KL) divergence for each pair of group-level distributions and performed a permutation test (see Materials & Methods). KL divergence measures the difference between two probability distributions, with lower values indicating higher similarity and values close to zero suggesting negligible divergence. Each panel shows the average normalized frequency that each group spent in p(engaged) bins during the first session. (A) *Tsc2* WT and Het males showed a low observed KL divergence of 0.129, with a non-significant p-value of 0.433. (B) *Shank3B* WT and Het males had an observed KL divergence of 0.264, $p = 0.20$, further indicating high overlap in their distributions. We also analyzed *Shank3B* KO males as a single group and divided into sub-groups described in the text, compared to *Shank3B* WT males. (C) The comparison between *Shank3B* WT and all *Shank3B* KO males yielded a KL divergence of 0.241, $p = 0.266$. (D) *Shank3B* WT and *Shank3B* KO sub-group 1 had an observed KL divergence of 0.601, $p = 0.161$, while (E) *Shank3B* WT and *Shank3B* KO sub-group 2 had an observed KL divergence of 0.121, $p = 0.766$. Together, these results highlight that the KL divergence

values remain low across comparisons and, coupled with non-significant p-values, suggest no meaningful differences in the group-level distributions of p(engaged). Error bars represent SEM.