Behavioral/Cognitive

# Multidimensional Analysis of a Social Behavior Identifies Regression and Phenotypic Heterogeneity in a Female Mouse Model for Rett Syndrome

**Michael Mykins, Benjamin Bridges, Angela Jo, and Keerthi Krishnan**
Department of Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee

**Regression is a key feature of neurodevelopmental disorders such as autism spectrum disorder, Fragile X syndrome, and Rett syndrome (RTT). RTT is caused by mutations in the X-linked gene methyl-CpG–binding protein 2 (*MECP2*). It is characterized by an early period of typical development with subsequent regression of previously acquired motor and speech skills in girls. The syndromic phenotypes are individualistic and dynamic over time. Thus far, it has been difficult to capture these dynamics and syndromic heterogeneity in the preclinical *Mecp2*-heterozygous female mouse model (Het). The emergence of computational neuroethology tools allows for robust analysis of complex and dynamic behaviors to model endophenotypes in preclinical models. Toward this first step, we utilized DeepLabCut, a marker-less pose estimation software to quantify trajectory kinematics and multidimensional analysis to characterize behavioral heterogeneity in Het in the previously benchmarked, ethologically relevant social cognition task of pup retrieval. We report the identification of two distinct phenotypes of adult Het: Het that display a delay in efficiency in early days and then improve over days like wild-type mice and Het that regress and perform worse in later days. Furthermore, regression is dependent on age and behavioral context and can be detected in the initial days of retrieval. Together, the novel identification of two populations of Het suggests differential effects on neural circuitry, opens new avenues to investigate the underlying molecular and cellular mechanisms of heterogeneity, and designs better studies for stratifying therapeutics.**

*Key words:* **behavioral heterogeneity; female mice; pose estimation; pup retrieval; regression; Rett syndrome**

### Significance Statement

**A long-standing problem in the field of neuropsychiatric disorders is the reliable identification of heterogeneous endophenotypes in animal models for the disorders. This problem has clear implications for identifying the etiology and therapeutic targets. The emergence of accessible computational neuroethology tools is a powerful solution in the systematic characterization of animal behaviors in resolving this heterogeneity. Using DeepLabCut and multidimensional analysis of an ethologically relevant social cognition task, we identify two distinct populations exhibiting delays in efficient behavior and regression in a female mouse model for Rett syndrome. The novel identification of two populations in genotypically identical mice has profound implications for both etiology and personalized therapeutic approaches.**

Received June 9, 2023; revised Nov. 1, 2023; accepted Nov. 17, 2023.

Author contributions: M.M. and K.K. designed research; M.M. performed research; M.M., B.B., and A.J. analyzed data; M.M. and K.K. wrote the paper.

We thank undergraduates Alexandra McBryar and Trinity Rose Schultz for their technical help; graduate students Logan Dunn and Jacob Elrod and postdoctoral researcher Dr. Billy Lau for their technical help and intellectual discussion; and the OpenBehavior community for sharing their tools and advice in growing this computational neuroethology field. This work was supported by BCMB Chancellor's fellowships (M.M.); the 2021 Scholarly and Research Incentive Funds through the Office of Research, Innovation & Economic Development at the University of Tennessee, Knoxville (UTK, M.M.); BCMB James and Dora Wright Fellowship (M.M.); UTK 2021 Summer Undergraduate Research Group Experience Program (K.K., M.M., B.B.); UTK Rocky Top Summer Research Mentor Program (K.K., M.M., A.J.); startup funds from the University of Tennessee, Knoxville (K.K.); and the National Institute of Mental Health of the National Institutes of Health under Grant Number R15MH124042 (K.K.).

The authors declare no competing financial interests.

Correspondence should be addressed to Keerthi Krishnan at krishnan@utk.edu.

https://doi.org/10.1523/JNEUROSCI.1078-23.2023



## Introduction

Regression is defined as a loss of previously acquired motor skills over time and is a behavioral hallmark of neurodevelopmental disorders (Charman et al., 2002; McVicar and Shinnar, 2004; Goin-Kochel et al., 2014; Thurm et al., 2018). Despite decades of clinical observations that regression is prevalent in many neurodevelopmental disorders and that the occurrence of regression is higher than previously reported, the underlying cause of regression in most neurodevelopmental disorders remains elusive (Goldberg et al., 2008; Goin-Kochel et al., 2014; Ozonoff et al., 2018, Ozonoff and Iosif, 2019). Genetic mutations in proteins important for transcriptional regulation and experience-dependent synaptic plasticity in the brain are associated with a higher risk for regression (Goin-Kochel et al.,

2017; Tammimies, 2019). Among these targets is the X-linked gene methyl-CpG–binding protein 2 (*MECP2*), the monogenic cause for Rett syndrome (RTT), and a "hotspot" gene associated with other disorders such as severe neonatal-onset encephalopathy, PPM-X syndrome, and microcephaly (Rett, 1966; Amir et al., 1999; Schüle et al., 2008; Gonzales and LaSalle, 2010; Lambert et al., 2016).

Most surviving RTT patients are girls and women who are heterozygous for *MECP2* mutations (Kirby et al., 2010). Female patients typically survive into middle age and exhibit impairments in sensory processing, cognitive function, and motor skills throughout life (Nomura, 2005; Neul et al., 2010; Djukic and Valicenti McDermott, 2012; Djukic et al., 2012; LeBlanc et al., 2015; Peters et al., 2015; Buchanan et al., 2019; Key et al., 2019; Stallworth et al., 2019; Symons et al., 2019; Merbler et al., 2020). In these females, random X-chromosome inactivation leads to mosaic wild-type *MECP2* expression and consequently, a syndromic phenotype. Broadly, RTT is diagnosed from two types of clinical presentations: (1) girls present with a short period of typical development, followed by regression and expression of stereotypic sensory, motor, speech, and cognitive impairments or (2) developmental delay and expression of stereotypic sensory, motor, speech, and cognitive impairments (Hagberg et al., 1983; Charman et al., 2002; Neul et al., 2010, 2023; Djukic and Valicenti McDermott, 2012; Han et al., 2012; Cosentino et al., 2019; Einspieler and Marschik, 2019; Symons et al., 2019). Preregression developmental delays or other abnormalities were noted in a large majority of patients (Kerr, 1987, 1995; Nomura and Segawa, 1990; Leonard and Bower, 1998; Charman et al., 2002). Regression typically occurs between 6 and 18 months of age and as late as 4–8 years from individual case studies (Han et al., 2012; Buchanan et al., 2019). These prolonged timelines of vulnerability suggest that the increased cognitive load required during early development while mastering motor control through exploration and interactions within the social context might reveal specific features over ages. Though RTT is diagnosed in early development, regression in specific skills or learning is also observed throughout life. Currently, the underlying neural pathways that manifest in developmental delay or regression after typical development and throughout different phases of life in patients with RTT are unknown.

Preclinical *Mecp2* rodent models recapitulate many of the phenotypic features of RTT such as sensory processing, social communication, and motor deficits (Durand et al., 2012; Goffin et al., 2012; Ito-Ishida et al., 2015; Krishnan et al., 2015, 2017; Su et al., 2015; Lo et al., 2016; Orefice et al., 2016; Veeraragavan et al., 2016; Lee et al., 2017; Lau et al., 2020a; Stevenson et al., 2021; Xu et al., 2022; Mykins et al., 2023). Thus far, three other published papers have shown possible regression phenotypes for breathing abnormalities in a mouse conditional knock-out model, learned forepaw skill involving seed opening in a female rat model, and skilled motor learning in female mice (Huang et al., 2016; Veeraragavan et al., 2016; Achilly et al., 2021). In these models, deterioration in learned skills was noted over multiple weeks. However, reliably identifying regression within a social behavioral task in apt preclinical models, with shorter time resolution in order to measure the impact of therapeutics, has remained a challenge.

Toward this end, we have established an ethologically relevant social behavioral assay called pup retrieval task as a model to study cellular and neural circuitry basis for social cognition, complex sensorimotor integration, and experience-dependent plasticity (Krishnan et al., 2015, 2017; Lau et al., 2020a, b; Stevenson et al., 2021; Dvorkin and Shea, 2022; Mykins et al., 2023). During this task, female mice integrate sensory cues to execute goal-directed motor sequences to retrieve scattered pups back to their home nest (Beach and Jaynes, 1956; Hemel, 1973; Stern and Mackinnon, 1978; Koch and Ehret, 1989; Stern, 1996; Champagne et al., 2007; Alsina-Llanes et al., 2015; Lonstein et al., 2015; Marlin et al., 2015; Champagne and Curley, 2016; Dunlap et al., 2020). Using end-point metrics such as latency index and errors during retrieval, we found that 6-week-old adolescent *Mecp2*-heterozygous female mice (Het) were as efficient at pup retrieval as wild-type littermate controls (WT), and 12-week-old adult Het were inefficient at retrieval (Krishnan et al., 2017; Stevenson et al., 2021; Mykins et al., 2023). However, movements during social behaviors are dynamic over time, with distinct behavioral sequences that are often ignored due to their inherent complexity (Stevenson et al., 2021).

The emergence of computational neuroethology tools allows for fast, reliable, and systematic detection of recorded movement over time to extract multiple metrics in different contexts. This extraction is critical for identifying unique subpopulations and assessing the therapeutic value for preclinical animal studies for psychiatric disorders (Yamamoto et al., 2018; Popovitz et al., 2021; Tanas et al., 2022; Luxem et al., 2023; Shemesh and Chen, 2023). Thus, we used DeepLabCut, a marker-less pose estimation software, to capture the complexity of individual variations in a mosaic *Mecp2*-heterozygous population over multiple days of pup retrieval behavior (Mathis et al., 2018; Nath et al., 2019). From the derived animal trajectories, we performed a multidimensional analysis of trajectory kinematics in adolescent and adult female WT and Het. The goal of this study was to determine if DeepLabCut can provide better behavioral characterization and aid in identifying regression in this heterogeneous population of Het.

Through this approach, we report the identification of robust phenotypic variation and regression in genotypic adult Het. Particularly, we identified genotypic Het that had milder phenotypes (Het-nonregressors, Het-NR), similar to adult WT in many features, and another distinct group of genotypic adult Het that exhibited severe inefficiency in consolidating the pup retrieval movements over days (Het-regressors, Het-R). Interestingly, regression was dependent on context as we observed that WT, Het-NR, and Het-R clustered together in Principal Component (PC) space during habituation and isolation phases, suggesting that Het-R have specific issues likely in integrating sensory and motor sequences during the decision/execution phase of retrieval. Additionally, this inefficiency in consolidation was specific to adulthood as adolescent Het behaved similarly to adolescent WT, indicating that genotypic Het are able to perform this complex pup retrieval task at an earlier time point in life and are not able to maintain goal-directed movement skills as they age. The novel identification of two populations, Het-NR and Het-R, suggests differential effects on neural circuitry and opens new directions of exploration on cellular mechanisms for compensation and maladaptive plasticity. Together, these results emphasize the need to analyze the dynamics of individual variations and context-dependency while considering therapeutic and treatment options.

## Materials and Methods

### Animals

We used the following mouse strains: CBA/CaJ (JAX:000654) and *Mecp2*-heterozygous (Het; C57BL/6J, B6.129P2©-Mecp2$^{tm1.1Bird}$/J, JAX:003890) and WT littermate controls (Guy et al., 2001). All animals were group-housed by sex after weaning, raised on a 12/12 h light/dark cycle (lights on at 7 A.M.), and received food and water ad libitum.

Behavioral experiments were performed using 6-week-old (adolescent) and 10- to 12-week-old (adult) Het and WT between the hours of 9 A.M. and 6 P.M. (during the light cycle). All procedures were conducted in accordance with the National Institutes of Health's Guide for the Care and Use of Laboratory Animals and approved by the Institutional Animal Care and Use Committee at the University of Tennessee–Knoxville.

### Pup retrieval task

*Behavioral paradigm*

The pup retrieval task was performed as previously described (Stevenson et al., 2021; Mykins et al., 2023). Two 6-week-old or 10- to 12-week-old naïve female littermates (one WT and one Het) with no prior experience of pup retrieval were cohoused with a pregnant CBA/CaJ female (10–12 weeks of age) 3–5 d before birth. Once the pups were born (Postnatal Day 0; D0), we performed the pup retrieval task once a day for 6 consecutive days in the home cage that was placed inside a sound- and light-proof box. The behavioral task was performed as follows (Fig. 1*a*.): one mouse was placed in the home cage with 3–5 pups (habituation, 5 min), the pups were removed from the nest (isolation, 2 min), and the pups were scattered by placing them in the corners and center, allowing the mouse to retrieve pups back to the nest (retrieval, 5 min at max). The nest was left empty if there were fewer than five pups. The assay was performed in the dark and recorded using an infrared camera (Foscam Wired IP Camera) during all phases of each trial. If all pups were not retrieved to the nest within 5 min, we removed the mouse and placed the pups back in the nest in preparation for the next mouse. While one of the females was performing the pup retrieval task, the mother, her pups, and the other adult were housed in a separate new cage. The adult females were chosen in a random order, by cage and by day. After the task was completed, all mice and pups were returned to the home cage.

*Behavior analysis*

All recorded videos were coded so the analyzers were blind to the identity of the mice. Each video was manually scored using a latency index (the amount of time to retrieve all pups back to the nest out of 5 min) and the number of errors that represent adult–pup interactions that did not result in a successful retrieval.

Latency index was calculated as follows: latency index = $[(t1-t0) + (t2-t0) + \ldots + (tn-t0)]/(n \times L)$,

where $n$ = number of pups outside the nest, $t0$ = start time of trial (s), $tn$ = time of the $n$th pup's successful retrieval to the nest (s), and L = total trial length (300 s). Statistical analyses and figures were generated using R studio.

*Pose estimation generated from DeepLabCut.* For body part tracking, we used DeepLabCut (version 2.2.0.3; Fig. 1*b*; Mathis et al., 2018; Nath et al., 2019). We defined features of interest as the mouse's nose, left ear, right ear, shoulder, spine 1, spine 2, left hindlimb, right hindlimb, and tail base. We labeled 200 frames of features of interest taken from 25 videos of adult WT and Het during the pup retrieval phase, and then 95% was used for training. We used a ResNet-50–based neural network with default parameters and trained for 720,000 iterations on Google Colab (Insafutdinov et al., 2016). We performed outlier correction to correct inaccurate predictions from the DeepLabCut model. In the Graphical User Interface, we refined 50 frames of predicted labels by moving the predicted label's location to the actual position of the body part of interest (Fig. 1*b*). We then retrained, validated with 1 number of shuffles, and found a test error of 1.67 pixels and a train error of 1.66 pixels (image size was 640 × 480 pixels). We then used a $p$-cutoff of 0.9 to condition the $X$ and $Y$ coordinates for future analysis. This model network was then used to analyze novel behavioral videos of adult animals. We then repeated this process for 6-week-old adolescent pup retrieval behavior. We labeled 200 frames taken from 20 videos of 6-week-old pup retrieval behavior, trained for 800,000 iterations, refined 75 frames of predicted labels, retrained for 550,000 iterations, validated with 1 number of shuffles, and repeated this process to achieve a test error of 1.89 pixels and a train error of 1.82 pixels.

*Pose estimation–derived trajectory analysis.* To quantify behavioral trajectories during habituation, isolation, and retrieval, we used the

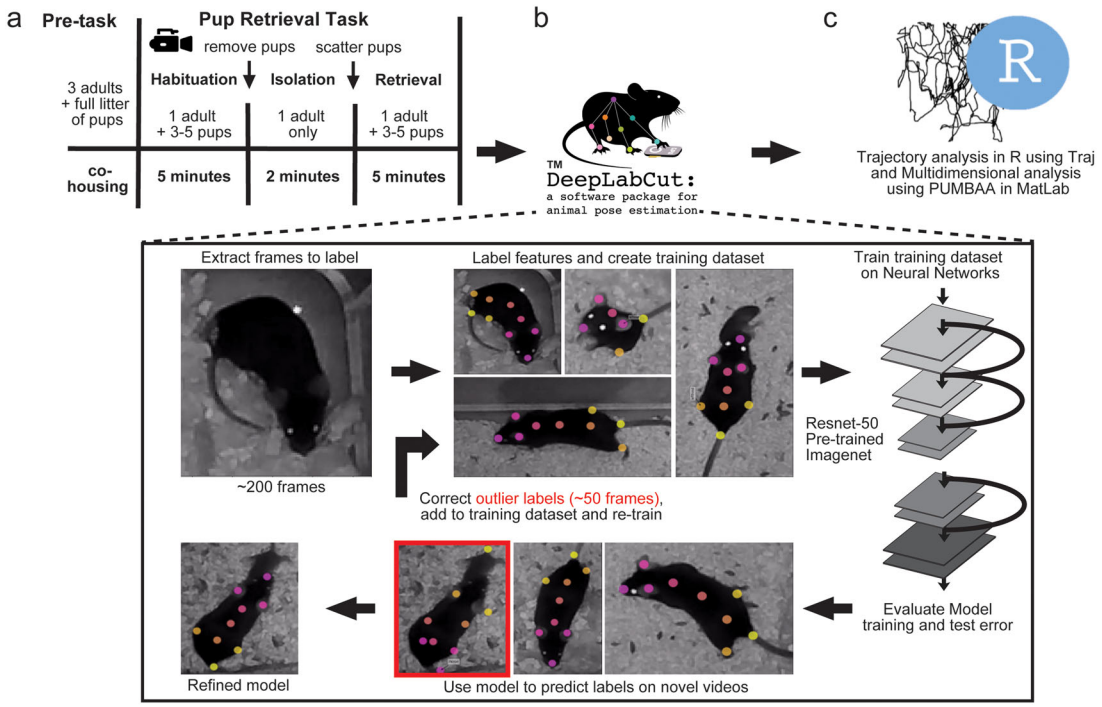


**Figure 1.** DLC-derived pipeline for multidimensional analysis of trajectories. ***a***, Schema depicting pup retrieval behavior setup. We performed and recorded habituation, isolation, and pup retrieval in 6-week-old (W0) adolescent and 10–12 W0 adult female WT and *Mecp2*-heterozygous (Het) mice for 6 consecutive days (see Materials and Methods for details). ***b***, ***c***, We trained videos of behavior in single animal DeepLabCut (DLC; ***b***, top) to predict the pose of the animal, and features of interest were labeled and trained in DLC to build pose models that were then refined with more training to correct for errors (***b***, bottom). From our models, we generate pose estimation and extract features of interest, such as the nose, to analyze trajectories of mice in R using the Traj package (*McLean and Skowron Volponi 2018*; ***c***) and perform multidimensional behavioral analysis, adapted from Tanas et al. (2022).

**Table 1. Metrics**

| Metrics | Definition |
|---|---|
| Latency index | – Normalized time it takes between retrieving *n* number of pups: latency index = $[(t1 - t0) + (t2 - t0) + \ldots + (tn - t0)]/(n \times L)$ out of the 5 min given to retrieve pups (Krishnan et al., 2017; Stevenson et al., 2021; Rupert et al., 2023) |
| Errors | – Number of physical interactions with pup during the retrieval phase that does not result in retrieval or dropping of the pup during retrieval (retrieval phase of behavior only, excluded in habituation and isolation) |
| Distance | – Total distance covered (in meters) during phase of behavior |
| Duration | – The total amount of time it takes to retrieve all pups or the max 5 min (retrieval phase only) |
| Mean speed | – Change in distance over time (scalar quantity) |
| Mean acceleration | – Acceleration—change in speed over time (scalar quantity) |
| Mean step length | – The average distance between two steps |
| Sinuosity | – Tortuosity of a random search path (Bovet and Benhamou, 1988; Benhamou, 2004)<br>– A function of both the mean cosine of turning angles and step length (Cheung et al., 2007; McLean and Skowron Volponi, 2018)<br>– Varies between 0 (straight) and 1 (random motion) |
| Straightness | – Minimum distance to get from point a to b (D) divided by displacement from origin to final destination (Mykins et al., 2023) (L; Batschelet, 1981)<br>– Closer to 1, the more efficient the trajectory. The closer to zero, the more random the walk (Batschelet, 1981; Benhamou, 2004) |
| Emax | – Dimensionless value that denotes the maximum expected displacement of a random path as a function of the number of steps (Cheung et al., 2007; McLean and Skowron Volponi, 2018)<br>– Values closer to 0 are more sinuous trajectories, and larger values (approaching infinity) are straighter trajectories (Cheung et al., 2007) |
| Emaxb | – The product of Emax times the defined step size (Cheung et al., 2007)<br>– larger values (approaching infinity) are straighter |
| Nonlinearity | – Measure of directional change<br>– Angular change (in degrees) between two steps, divided by the time difference between the two steps (Benhamou, 2004; Kitamura and Imafuku, 2015) |
| Irregularity | – Standard deviation of directional change (nonlinearity; Benhamou, 2004; Kitamura and Imafuku, 2015) |

Metric definitions for trajectory analysis.

Traj R package (McLean and Skowron Volponi, 2018; Fig. 1*c*). We filtered the *X* and *Y* coordinates of the nose from the DeepLabCut pose and quantified 13 trajectory metrics (10 metrics were used for habituation and isolation as they did not have latency index, duration, and error; Table 1). For the habituation and isolation phases, we analyzed trajectories during the entire 5 min and 2 min trial phases, respectively. For the pup retrieval phase, we analyzed trajectories during the act of retrieval; we stopped analyzing trajectories after all pups had been retrieved.

### Multidimensional analysis of trajectories

For the multidimensional analysis of behavior profiles, we performed data selection, standardization, principal component analysis (PCA), k-means clustering, and validation (Fig. 1*b*; Popovitz et al., 2021; Tanas et al., 2022). We utilized PUMBAA (phenotyping using a multidimensional behavioral analysis algorithm, https://github.com/sidorovlab/PUMBAA) for all multidimensional analyses (Tanas et al., 2022).

*Data selection for retrieval*

For individual and combined-days analyses of retrieval, 13 × the number of days of metrics were included in multidimensional analysis or PCA analysis.

*Data selection for habituation and isolation*

For all days' analysis, a total of 60 metrics (10 metrics across 6 days) were included in the multidimensional analysis (Fig. 5*b*,*e*). For combined D1 and D5 analysis, 20 metrics were included in the multidimensional analysis (Fig. 5*c*–*f*).

*Standardization*

All measures were standardized using *z*-score normalization [$z = (\text{data point} - \text{group mean}) / \text{standard deviation}$]. This accounts for different unit magnitudes across metrics similar to previous studies (Popovitz et al., 2021; Tanas et al., 2022).

*Principal component analysis and k-means clustering*

We performed PCA using PUMBAA and extracted the amount of variance explained by each PC (Fig. 4*b*). We performed k-means clustering in the PC space using PUMBAA with $k = 2$ clusters in 2 Principal Component (2PC) space. To verify the visual absence of clusters, we performed a gap statistics test using the cluster package in R. Data visualization using ggplot was performed in R studio (Wickham, 2016).

*Validation*

We compared the genotypes to their predicted k-means cluster. We color-coded groups according to their predictions (WT predicted correctly = black, Het predicted correctly = red, Het predicted incorrectly = blue, WT predicted incorrectly = purple; Fig. 3). This allowed us to determine the accuracy of predicting genotypes in our model and revealed two distinct populations with the same genotype. We set a seed value for random number generation and used R packages (cluster, factoextra) to test the optimal number of clusters using k-means and bootstrapping with a sample size of 100 and a maximum of four clusters. The data visualization using ggplot was performed in R studio (Wickham, 2016).

*Using predicted clusters to analyze behavior*

Using the identified clusters of Het from Figure 3*f*, we performed a trajectory analysis for metrics contributing significantly to the PCA. We conducted the Kruskal–Wallis test with uncorrected Dunn's test to determine the statistical significance of metrics of interest between WT and the two clusters of Het (Fig. 4) on each day and within the same group across days. We used base R, ggplot, emmeans, and dunn.test package to perform statistical analysis and data visualization in R studio (Wickham, 2016).

## Results

### Robust identification of phenotypic heterogeneity in adult Het

Previously, we reported that the adult female Het were inefficient at pup retrieval, as measured by the time taken to retrieve pups back to the nest (latency index), and increased physical interactions between the pups and adult that did not result in an efficient retrieval (errors; Krishnan et al., 2017; Stevenson et al., 2021). We also reported the individual variations with possible regression phenotype over days in adult Het, from frame-by-frame manual analysis of goal-directed movements in a small sample size ($N = 6$ females per genotype; Stevenson et al., 2021). These intriguing results identified the need for a systematic, automated, and unbiased analysis for large sample sizes over multiple days. Here, we analyzed the end-point metrics of pup retrieval with a larger sample set of $N = 21$ for adults and $N = 9$–11 for adolescents. A two-way mixed-effects ANOVA revealed the significant effects of the day of retrieval ($F = 18.68$, $p < 0.0001$) and genotype ($F = 38.39$, $p < 0.0001$) on latency, but no interaction between those variables ($F = 0.35$, $p = 0.89$; Extended Data Table 2–1). Post hoc testing revealed that, compared with baseline performance on D0, WT and Het significantly decreased in latency index, over days (Fig. 2*a*,*a'*, WT and Het; Extended Data Table 2-1). However, Het were was significantly worse than WT on all days of retrieval, and more variable (Fig. 2*a*,*a'*, WT

vs Het). Consistent with previously reported results, some of the individual Het performed on par with WT.

A two-way mixed-effects ANOVA revealed a significant effect of the genotype ($F = 29.57$, $p < 0.001$), but not the day of retrieval, ($F = 1.73$, $p = 0.13$) on the number of errors, and no interaction between those variables ($F = 0.37$, $p = 0.87$; Extended Data Table 2-1). Post hoc testing revealed that WT and Het significantly decreased in number of errors, compared with their own baseline at D0 (Fig. 2*b*,*b'*, WT and Het). However, Het were significantly worse and more variable than WT on all days except D3 (Fig. 2*b*,*b'*, WT vs Het). Collectively, these results recapitulated our previous results in a larger dataset and suggested the possibility of tackling individual variations in the Het population.

To better characterize this heterogeneity, we sought to identify newer dynamic metrics in context-specific motor sequences, in addition to our established end-point measurements. We used DeepLabCut to generate animal pose and quantified movement trajectory profiles across days in adolescent ($N = 9$ WT and $N = 11$ Het) and adult female WT and Het ($N = 21$ per genotype; Mathis et al., 2018; Fig. 2, Table 1). We took 13 metrics from each day of retrieval (13 metrics per day) and across 6 d of pup retrieval (78 metrics total) to build a behavioral profile for each mouse (Fig. 3*a*; Bovet and Benhamou, 1988; Benhamou, 2004, 2014; Cheung et al., 2007; McLean and Skowron Volponi, 2018; Tanas et al., 2022). We standardized all metrics using *z*-scores and performed a multidimensional PCA (Fig. 3*a*, scaled PCA; Popovitz et al., 2021; Tanas et al., 2022). We then used k-means to cluster mice into two groups based on genotype (Group 1: WT or Group 2: Het) and validated clustering by comparing the predicted cluster group to the known genotype of the animal (Group 1: WT; Group 2: Het), whereas incorrectly predicted genotype is labeled as Group 3 or 4 (Fig. 3*a*, k-means and Validation, respectively). Based on an average analysis across days, we identified a distinct group of Het clustered at the top of the PC space (Fig. 3*b*, Group 2) while another group of Het clustered with WT, and thus were incorrectly predicted as WT (Fig. 3*b*, Group 3, blue). The single-day analysis showed that the Group 2 cluster emerged as early as D1. However, it consolidated with Group 1 and Group 3 on D2 and D3 and emerged again on D4 and D5 (Fig. 3*c*). This result indicates dynamic variations in the movement trajectories of individual Het mice over days. The pairwise combination of D1 and D5 is sufficient to cluster WT accurately and identify two different clusters of Het (Fig. 3*d–f*). However, the pairwise comparisons between D0 and D5, D2 and D5, D3 and D5, and D4 and D5 demonstrated an inaccurate clustering as WT are predicted incorrectly (Extended Data Fig. 3-1; WT-incorrectly, purple). As pup numbers slightly vary between cohorts, we quantified the number of pups to be retrieved in each identified group to make sure this is not a major factor driving the identification of clusters [Fig. 3*g*; ANOVA ($F = 0.929$, $p = 0.404$); Extended Data Table 3-1]. Together, these novel results robustly quantify the known phenotypic heterogeneity in adult Het over multiple days.

### Identification of trajectory metrics that classify behavioral heterogeneity

In order to better characterize the three cluster groups, we plotted example trajectories from each group and noted striking differences in the movement patterns over days (Fig. 4*a*). Due to these trajectories and distinct metrics features, which are elaborated below, we adopted a new nomenclature to better describe the phenotypes in the three groups. We refer to Group 1 as WT, Group 2 as Het-regressors (Het-R), and Group 3 as Het-nonregressors (Het-NR). Similar to WT, Het-NR trajectories were random on D0 and consolidated to oriented and direct paths toward the nest by D5 (Fig. 4*a*, WT: black, Het-NR: blue). Het-R also had random trajectories on D0 but showed instability in consolidating trajectories over time compared with WT and Het-NR (Fig. 4*a*, Het-R: red).

We plotted the variable vectors in the PC space (Fig. 4*b*) for D1 and D5, as this allowed us to isolate the regressors with the

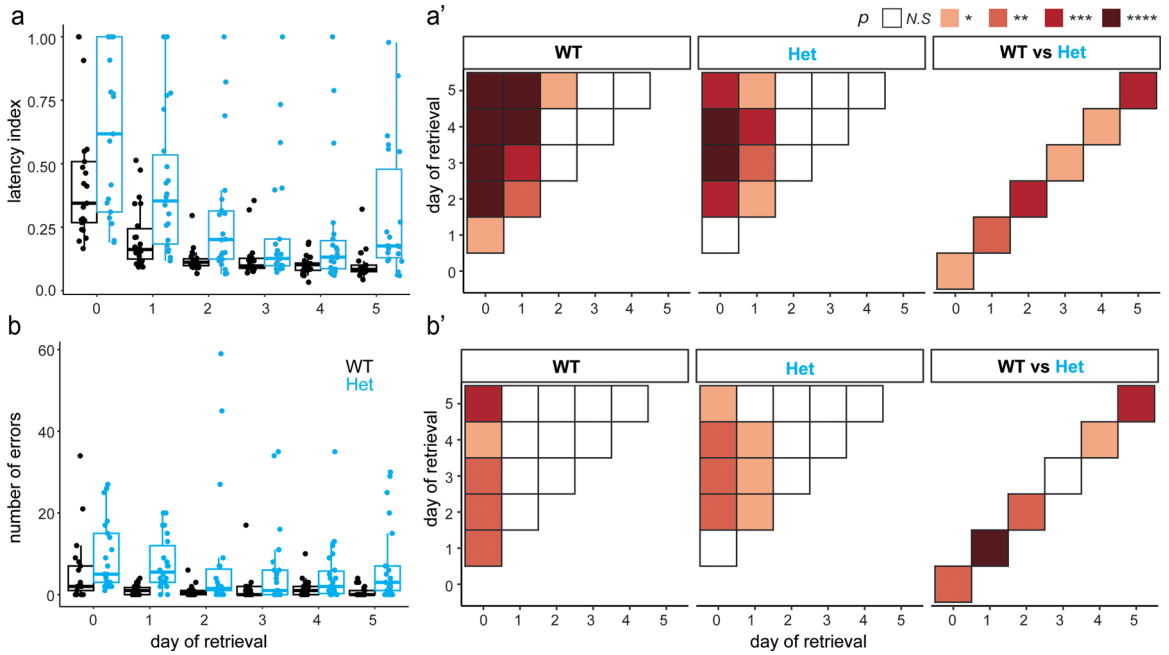


**Figure 2.** Adult *Mecp2*-heterozygous female mice (Het) are inefficient at pup retrieval task. WT and Het showed significant improvement across days, (***a***) median latency index, and (***b***) median number of errors WT (black, $n = 21$), Het (blue, $n = 21$). Box plots showing median scores for latency and error. Each dot represents an animal (***a***,***b***). ANOVA Type II analysis for multiple comparison with Šidák correction was performed and followed by post hoc Kruskal–Wallis uncorrected Dunn's (Extended Data Table 2-1) to test for statistical significance between genotypes and within genotype across days. Significance is plotted as correlation matrices color-coded by significance for (***a'***) latency index and (***b'***) number of errors. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ****$p < 0.0001$, N.S., not significant.

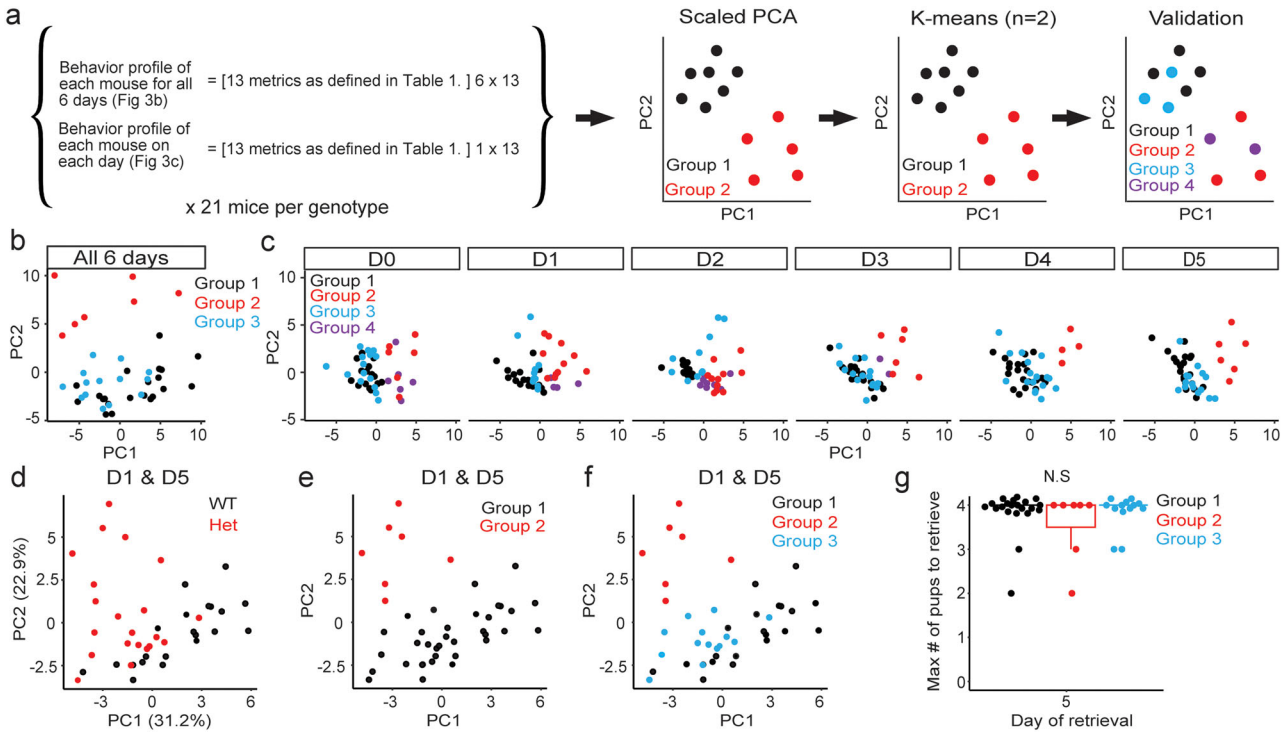


**Figure 3.** Multidimensional analysis identifies two distinct phenotypes of Het. ***a***, Thirteen trajectory metrics for all days of retrieval (78 total loadings) or 13 trajectory metrics on each day (13 total loadings) were standardized using *z*-scores for PCA (left). Example of analysis: each point represents one animal's behavioral profile in PC space, colored by group (scaled PCA). Mice were clustered into two groups (Group 1 = black, Group 2 = red) using k-means to predict genotype (k-means). The predicted genotypes were validated by comparing animals with their genotype (Group 1 = WT, black; Group 2 = Het, red; Group 3 = Het predicted incorrectly, blue; Group 4 = WT predicted incorrectly, purple; validation). ***b***, Multidimensional analysis for all days of retrieval revealed two groups of Het: Het that clustered with WT (Group 3, blue) and Het that were distinctively different from Group 1 and Group 2 (Group 3). ***c***, PCA, k-means clustering, and validation for each individual day of retrieval (D0–D5) and pairwise comparisons of each day with D5 (D1:D5, D2:D5, D3:D5, and D4:5; Extended Data Fig. 3-1) revealed that the combination of D1:D5 identified the two groups of Het. ***d***, PCA of 13 trajectory metrics for D1 and D5 of retrieval (26 loadings) color-coded by actual genotype (WT = black, Het = red). ***e***, k-means clustering ($n = 2$) of 13 trajectory metrics for D1 and D5 of retrieval (Group 1 = black, Group 2 = red). ***f***, Validation of predictions supports two distinct groups of Het with reduced data complexity. ***g***, The max number of pups to retrieve is not significantly different between the three identified groups. For all panels, each dot represents an animal, color-coded by genotype and validated cluster group. For all panels, regardless of cluster prediction, there are 21 WT and 21 Het mice. ANOVA Type II (Extended Data Table 3-1). N.S = not significant.

least number of dimensions (Fig. 3*f*). Sinuosity, the expected maximum displacement (emax) and expected maximum displacement B (emaxB) contributed to PC1. Distance, duration, nonlinearity, irregularity, latency index, and error contributed to PC2. The mean acceleration and straightness had negative correlations and did not contribute to either PC (Fig. 4*b*). The positive correlations between emax, emaxB, mean speed, mean step length, and negative correlation with sinuosity contributed to PC1 and PC2 (Fig. 4*c*). The breakdown of contributions by PC1 (Fig. 4*c*) and PC2 (Fig. 4*d*) indicated that nonlinearity, irregularity, distance, and latency contributed to the most variability in PC2.

We previously reported latency index to evaluate pup retrieval performance (Krishnan et al., 2017; Stevenson et al., 2021; Mykins et al., 2023; Fig. 2). We analyzed latency index for the three phenotypes (WT, Het-NR, Het-R). A three-way mixed-effects ANOVA revealed the significant effects of the day of retrieval ($F = 9.51$, $p < 0.0001$) and phenotype ($F = 8.45$, $p < 0.0001$) on latency, but no interaction between those variables ($F = 0.89$, $p = 0.55$; Extended Data Table 4-1). Post hoc testing revealed that WT and Het-NR significantly decreased in latency index over days, compared with D0 within their respective genotypes (Fig. 4*e*,*e'*, WT, Het-NR; Extended Data Table 4-1). This suggests learning and consolidation toward efficient pup retrieval task over multiple days for both groups. However, Het-NR were significantly worse than WT on D1, D2, D4, and D5 (Fig. 4*e*,*e'*, WT vs Het-NR;), suggesting deficiencies in efficient pup retrieval. Similar to WT and Het-NR patterns, the Het-R latency index significantly decreased from D0 to D2–D4, and increased again to D0 levels on D4–D5. Het-R were significantly worse than WT on all days (Fig. 4*e*,*e'*, WT vs Het-R;) and were significantly worse than Het-NR on D5 (Fig. 4*e*,*e'*, Het-NR vs Het-R; Extended Data Table 2-1). These results show that the Het-R phenotype is distinct from both Het-NR and WT.

Nonlinearity is a measure of directional change (Cheung et al., 2007; Kitamura and Imafuku, 2015; McLean and Skowron Volponi, 2018). A decrease in directional change is an indication that the path is more goal-directed. A three-way mixed-effects ANOVA revealed a significant effect of the day of retrieval ($F = 7.44$, $p < 0.0001$), but not phenotype ($F = 0.29$, $p = 0.75$) on nonlinearity, with a significant interaction between those variables ($F = 0.35$, $p = 0.89$; Extended Data Table 4-1). Post hoc testing revealed that WT and Het-NR significantly decreased in nonlinearity over days compared with baseline within genotype (Fig. 4*f*,*f'*, WT, Het-NR), suggesting consolidation of more goal-directed trajectories over days. There was no significant difference in nonlinearity between WT and Het-NR on all days (Fig. 4*f*,*f'*, WT vs Het-NR; Extended Data Table 4-1). However, Het-R had significantly increased nonlinearity on D4 and D5 and were significantly worse on those days compared with WT and Het-NR (Fig. 4*f*,*f'*, WT vs Het-R, Het-NR vs Het-R).

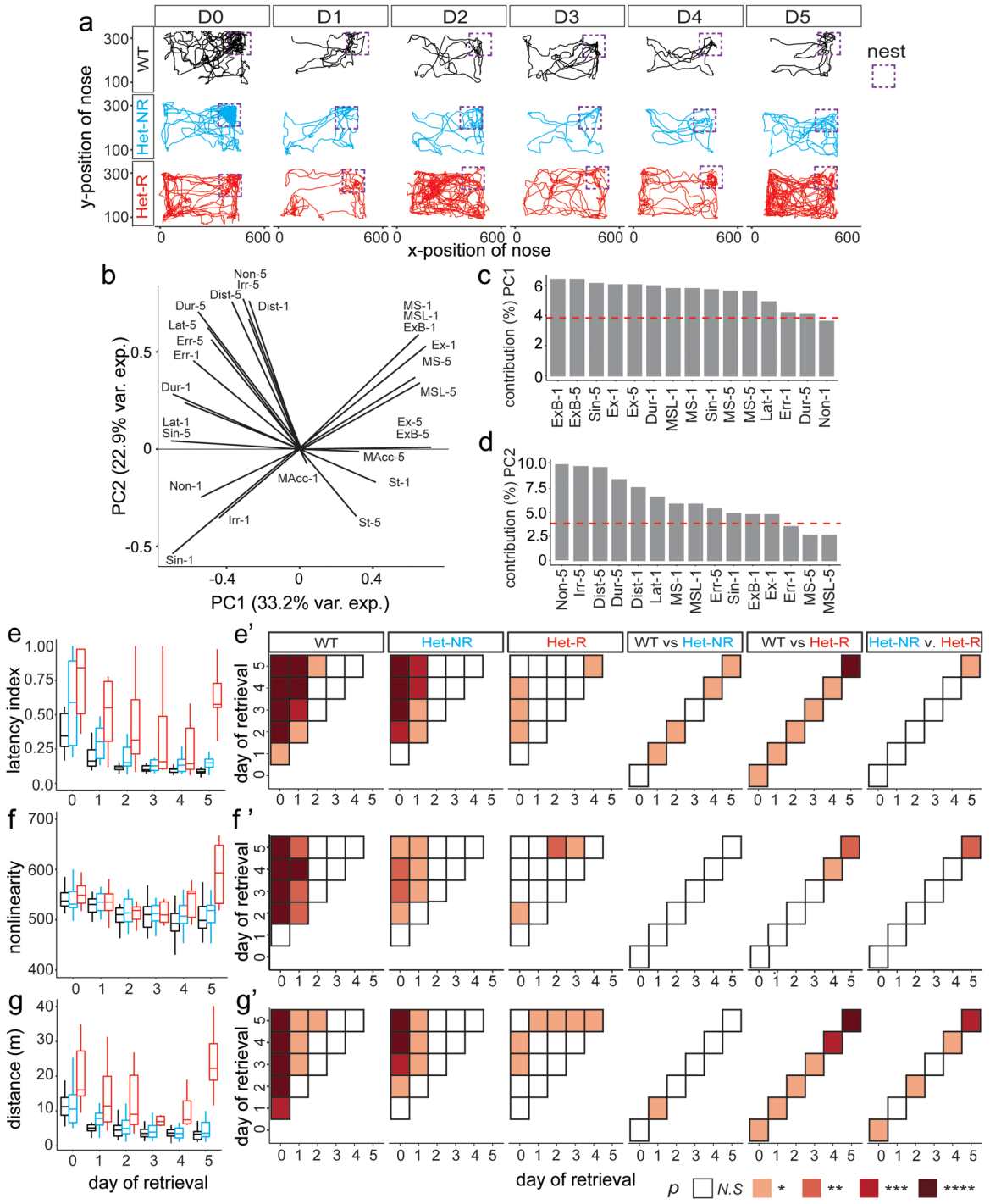


**Figure 4.** Multidimensional analysis distinguishes phenotypical heterogeneity in Het. ***a***, Example trajectories of WT (black), Het-NR (blue), and Het-R (red) *x–y* nose position during pup retrieval toward the nest (purple box) across 6 consecutive days of behavior. ***b***, Variable correlation plots of trajectory metrics (see key below) from D1 and D5 represented in PC space. Variable vectors grouped in the same direction are positively correlated, and vectors in the opposite direction are negatively correlated. Variance explained (var. exp. %) represents the percent of variance explained by each PC component. ***c***, ***d***, Percent contribution of top 15 (out of 26) metrics to Principal Component 1 (***c***) and Principal Component 2 (***d***). The red dashed line represents the average expected contribution of each metric. Everything above the red dashed line has a considerable contribution to the component. ***e***, Latency index, (***f***) median distance nonlinearity, and (***g***) median distance traveled during pup retrieval for each test day showed significant improvement in retrieval trajectories across test days for Het-NR. However, Het-R improved over days and regressed by D5 compared with D0 and were worse than WT and Het-NR. WT (black, $n = 21$), Het-NR (blue, $n = 14$), and Het-R (red, $n = 7$). ANOVA Type III analysis for multiple comparison with Šidák correction was performed and followed by post hoc Kruskal–Wallis uncorrected Dunn's (Extended Data Table 4-1) to test for statistical significance between genotypes and within phenotype across days (***e'***) latency index, (***f'***) nonlinearity, and (***g'***) distance. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ****$p < 0.0001$, N.S. = not significant. ***b–d***, Trajectory metric key: the number following the abbreviation indicates the day of retrieval. (Dist, distance; Dur, duration; Err, error; Ex, Emax; ExB, EmaxB; Irr, irregularity; Lat, latency; MAcc, mean acceleration; MS, mean speed; MSL, mean step length; Non, nonlinearity; Sin, sinuosity; St, straightness).

The total distance covered in a goal-directed trajectory decreases with efficiency and is thus indicative of the efficiency of the path taken. A three-way mixed-effects ANOVA revealed a significant effect of the day of retrieval ($F = 9.34$, $p < 0.0001$) and phenotype ($F = 12.23$, $p < 0.0001$) on distance and a significant interaction between those variables ($F = 2.73$, $p = 0.0034$; Extended Data Table 4-1). Post hoc testing revealed that WT and Het-NR significantly decreased their total distance covered over days compared with baseline within genotype (Fig. 4*g*,*g*', WT, Het-NR), suggesting an increased efficiency in determining the fastest path to retrieve pups back to the nest. Similar to WT and Het-NR, Het-R significantly decreased in total distance in early days compared with Het-R baseline (Fig. 4*g*,*g*', Het-NR). However, Het-R significantly increased the total distance traveled on D5 compared with D1–D4 and was similar to the total distance traveled at baseline (Fig. 4*g*,*g*', Het-R). Het-R covered more total distance than WT on all days (Fig. 4*g*,*g*', WT vs Het-R) and were significantly worse than Het-NR on D0, D2, D4, and D5 (Fig. 4*g*,*g*', Het-NR vs Het-R).

Overall, compared with WT, Het-NR displayed mild and day-specific atypical metrics in latency index and distance traveled, while still displaying improvement from their own baseline performance in all metrics. However, compared with WT and Het-NR, Het-R had a larger latency index, made more directional changes, and covered more distance during retrieval. Compared with the baseline, Het-R showed some improvement but was unable to consolidate the improvement, suggesting regression in this social motor task by D5.

### Features of Het regression are context-dependent

Given the dynamic changes in trajectory metrics during pup retrieval over days in the Het-R, we wondered if these features were specific to the act of pup retrieval or other generalized motor features in this subpopulation. Thus, we performed a PCA of trajectories during the habituation and isolation phases of behavior (Fig. 5). As latency and error metrics are specific to adult and pup interactions during retrieval, and duration is constant, we excluded them in the multidimensional analysis. During habituation, the pups huddle together in the nest, and the adult can choose whether to interact with the pups or not (Fig. 1*a*). This phase serves as a control to determine if the context and action of retrieving pups distinguish Het-NR and Het-R. We plotted the habituation trajectories [example plots for WT (black), Het-NR (blue), and Het-R (red)] and observed no differences in the patterns of their trajectories between D0 and D5 (Fig. 5*a*). Both genotypes tend to move around the entire cage during the 5 min habituation phase. No distinct groups emerged between WT, Het-NR, and Het-R from the differential PCA analyses (Fig. 5*b*,*c*). A three-way mixed-effects ANOVA revealed no significant effect of the day of habituation ($F = 0.09$, $p = 0.99$) or phenotype ($F = 0.90$, $p = 0.41$) on nonlinearity and no significant interaction between those variables ($F = 0.23$, $p = 0.99$; Extended Data Table 5-1). Post hoc testing revealed no significant difference in the average nonlinearity during habituation regardless of phenotype (Fig. 5*d*; Extended Data Table 5-1). A three-way mixed-effects ANOVA revealed no significant effect of the day of habituation ($F = 1.37$, $p = 0.24$) or phenotype ($F = 1.60$, $p = 0.20$) on distance and no significant interaction between those variables ($F = 0.35$, $p = 0.96$; Extended Data Table 5-1). Post hoc testing revealed no significant difference in the average distance during habituation regardless of phenotype (Fig. 5*e*).

During the isolation phase, the pups are removed from the home cage, and the adult is left alone for 2 min (Fig. 1*a*). We plotted the isolation trajectories of the three groups and observed no differences in the patterns of their trajectories between D0 and D5 (Fig. 5*f*). Similar to the habituation phase, both genotypes moved about the cage with no pups present, and no distinct groups emerged from the isolation-specific PCA analyses (Fig. 5*f*–*h*). A three-way mixed-effects ANOVA revealed no significant effect of the day of isolation ($F = 0.71$, $p = 0.16$) or phenotype ($F = 3.055$, $p = 0.21$) on nonlinearity and no significant interaction between those variables ($F = 1.077$, $p = 0.38$; Extended Data Table 5-1). Post hoc testing revealed no significant difference in average nonlinearity during isolation regardless of phenotype (Fig. 5*i*). A three-way mixed-effects ANOVA revealed no significant effect on the day of isolation ($F = 0.64$, $p = 0.67$) or phenotype ($F = 2.032$, $p = 0.13$) on distance and no significant interaction between those variables ($F = 0.60$, $p = 0.81$; Extended Data Table 5-1). Post hoc testing revealed no significant difference in average distance during isolation regardless of phenotype (Fig. 5*j*). These findings are consistent with the manual frame-by-frame analysis of maternal behaviors, which demonstrated that WT and Het behave similarly during the habituation and isolation phases of pup retrieval (Stevenson et al., 2021). Together, these results suggest context-specific, atypical behavioral features of Het-R and Het-NR and not generalized motor phenotype in genotypic adult Het. This is a critical finding that points to the context-specific processing of sensorimotor information in genotypic Het and the subsequent execution of specific motor sequences in a social context.

### Analysis of early days identifies individual variation in later adult Het-regressors

In order to determine individual variations in Het, we plotted individual trajectories of Het-NR and Het-R over all days. As expected, we observed that Het-NR, such as animals 8, 20, and 21, displayed some variability in adapting over time compared with the other Het-NR animals that plateaued and consolidated their efficiency in later days (Fig. 6*a*). In contrast, we observed variability in the individual trends in Het-R. Animal 10 (x symbol) did not improve in pup retrieval efficiency over days, animals 3, 4, 6, 9, and 19 improved over days but regressed by D5, and animal 2 (circle symbol) improved over days but showed instability in D4 (Fig. 6*a*'), suggestive of different neural mechanisms being disrupted in individual Het over days.

While it has been difficult to phenotype regression in Het mouse models of RTT, it is an even greater challenge to determine the time, frequency, and duration to apply therapeutic intervention to mitigate regression before its onset. We thus asked if it was possible to identify the Het-R population (Figs. 3, 6*a*') from the initial days of pup retrieval performance. Using D0 and D1 metrics (Fig. 6*b*), we obtained a weak informative power for identifying Het-NR (11/14 = 78.6% accuracy) and Het-R (3/7 = 43% accuracy). However, combinations of D0–D2 (Fig. 6*c*) and D0–D3 (Fig. 6*d*) metrics elicited stronger informative power for identifying Het-NR and a majority of Het-R early on (Het-NR: 13/14 = 93%, Het-R: 5/7 = 71%; Het-NR: 13/14 = 93%, Het-R: 6/7 = 86%, respectively). Despite the observations from select metrics that show Het-R trends similar to WT and Het-NR in early days (Fig. 4*e*–*g*), regression in social behavior can still be identified using all metrics (Fig. 6*c*,*d*). Together, these results suggest that we can identify potential regressors with a smaller number of trials, which has implications for designing preclinical therapeutic studies in the future.

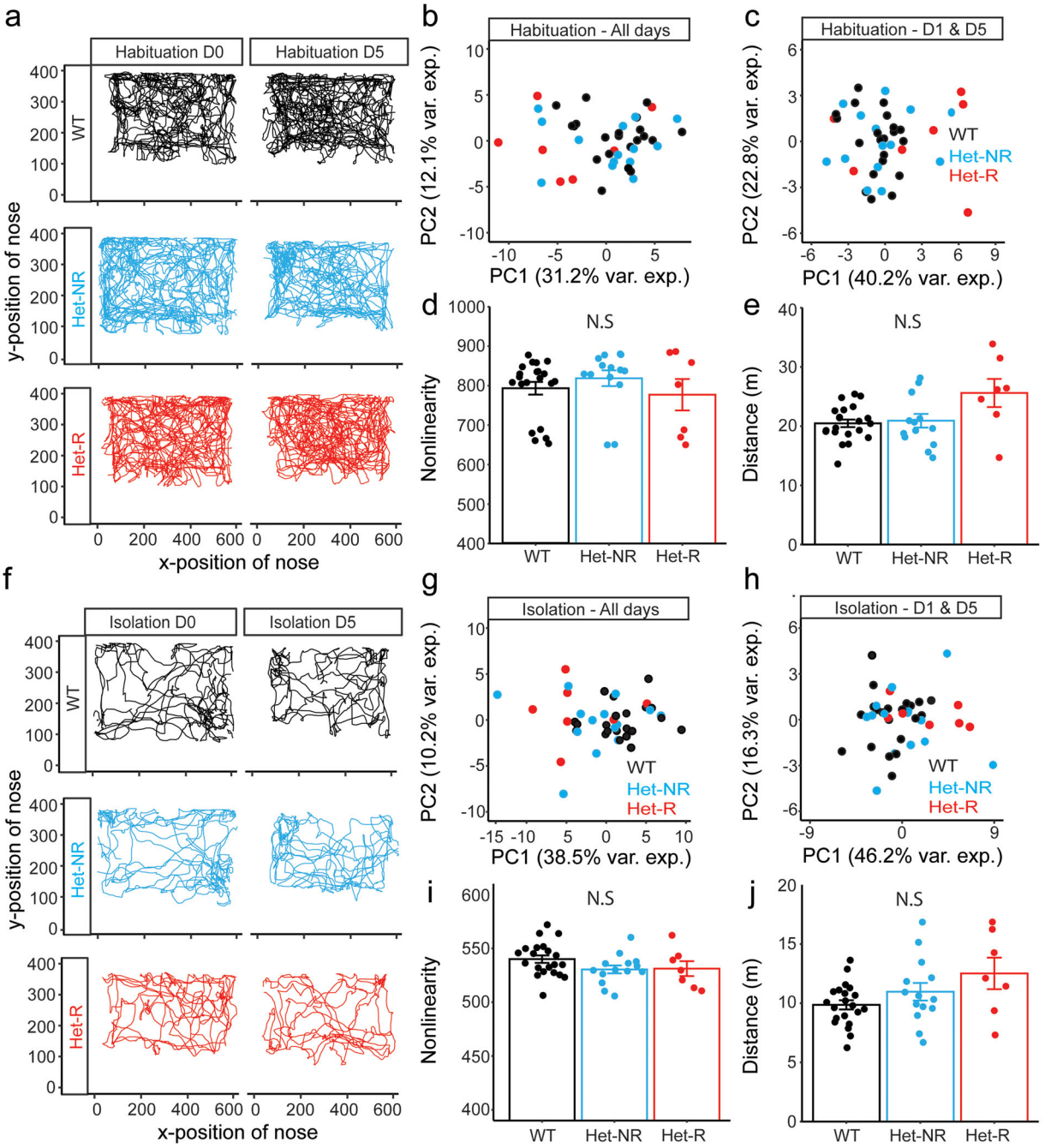


**Figure 5.** Regression is context-dependent. ***a***, Example trajectories of WT (black), Het-NR (blue), and Het-R (red) *x–y* nose position during habituation on D0 and D5. ***b***, ***c***, PCA for all 6 d (***b***) and combined D1 and D5 (***c***) of habituation revealed no difference between identified groups during habituation (WT = black, Het-NR = light-blue, and Het-R = red). Consistent with this finding, there is no significant difference in the average (***d***) nonlinearity and (***e***) distance traveled across the 6 d of the habituation phase of the pup retrieval paradigm. ***f***, Example trajectories of WT (black), Het-NR (blue), and Het-R (red) *x–y* nose position during isolation for D0 and D5. ***g–j***, PCA for all days (***g***) and combined D1 and D5 (***h***) of isolation revealed no difference between the identified groups in the absence of pups. Consistent with this finding, there is no significant difference in the average (***i***) nonlinearity and (***j***) distance traveled across the 6 d of the isolation phase of the pup retrieval paradigm. For ***b***, ***c***, ***g***, and ***h***, each dot represents an animal, color-coded by genotype and phenotype. For ***d***, ***e***, ***i***, and ***j***, the mean ± SEM across each phenotype is represented. Each dot represents the average nonlinearity or distance of each animal across 6 d of habituation (***d***, ***e***) and isolation (***i***, ***j***), respectively. ANOVA Type III analysis for multiple comparison with Šidák correction was performed and followed by post hoc Kruskal–Wallis uncorrected Dunn's (Extended Data Table 5-1) to test for statistical significance between phenotypes and within phenotype across: N.S = not significant. WT (black, $n = 21$), Het-NR (blue, $n = 14$), and Het-R (red, $n = 7$).

### Experience-dependent regression in Het is specific to age

We previously reported that 6-week-old (adolescent) Het performed pup retrieval comparable with adolescent WT, using endpoint metrics of latency index and error (Mykins et al., 2023). This is a particularly important result that shows that genotypic Het can perform as efficiently as the WT at an earlier age, which then results in regression features in adulthood. To determine if dynamic pose estimation metrics can identify early signs of regression in adolescent Het, we generated another deep learning model and pose for adolescent WT and Het animals and quantified trajectory metrics. We did not observe any distinct genotype groups through the PCA analysis (Fig. 7*a*,*b*). In support of this finding,

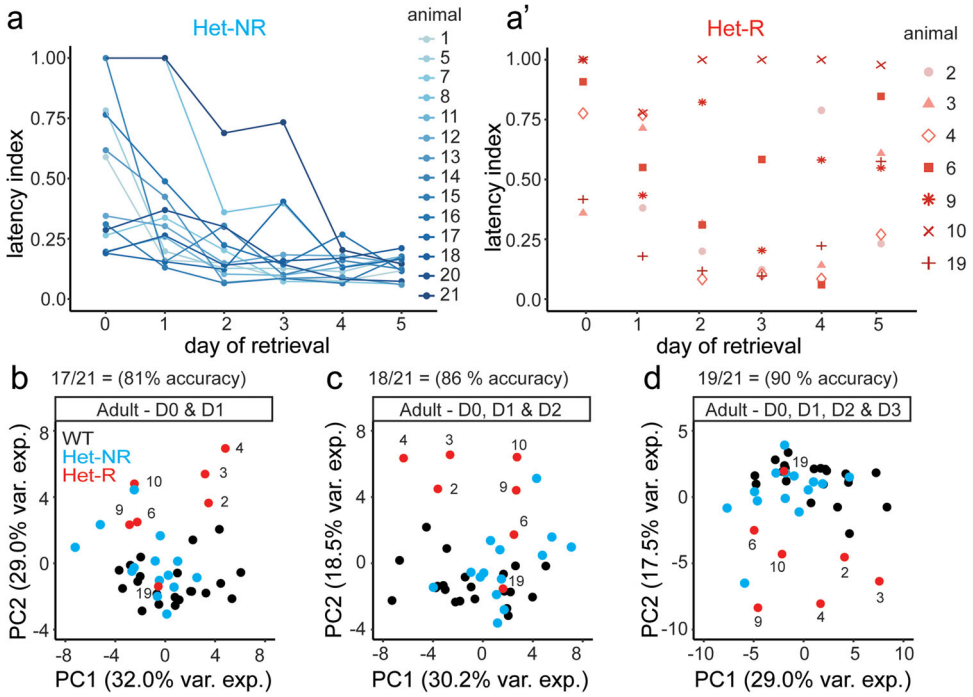


**Figure 6.** Performance in early days can identify future adult Het-regressors. ***a–a'***, Individual animal trends of latency index for the identified (***a***) Het-nonregressors (Het-NR) and (***a'***) Het-regressors (Het-R) from Figures 2 and 3. Trend lines are color-coded by animal number and shape to highlight individual variation in latency index over 6 d for Het-NR (blue-gradient) and Het-R (red-gradient + shape to easily identify animals). ***b–d***, PCA analysis for D0 and D1 (***b***), D0, D1, and D2 (***c***), and D0, D1, D2, and D3 (***d***) of retrieval. Analysis revealed that the first 3 (***c***) or 4 d (***d***) of pup retrieval accurately identified Het-NR and Het-R, while the first 2 d (***b***) had a lower accuracy for identifying Het-NR and Het-R. Accuracy was determined by correctly identifying the 21 Het into their established groups (Figs. 3, 4). The dots represent each animal [WT (black, $n = 21$), Het-NR (blue, $n = 14$), and Het-R (red, $n = 7$)], and Het-R are numbered and shaped to easily identify regressors over time.

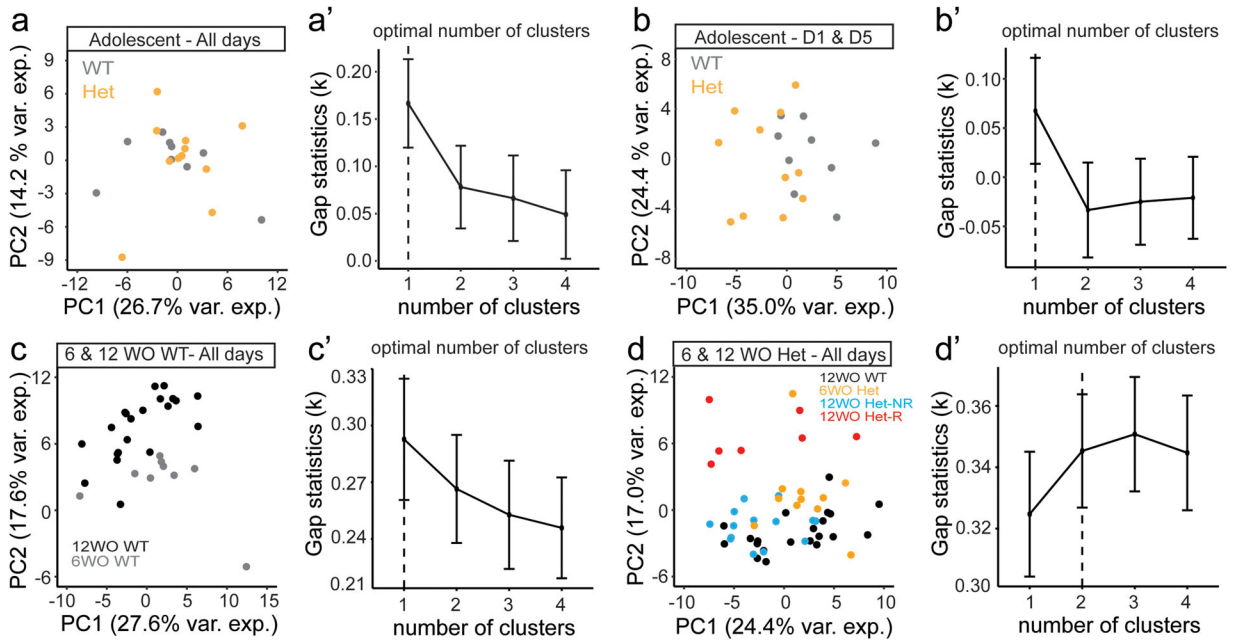


**Figure 7.** Het regression is age-specific. ***a***, PCA for all days of retrieval, and (***b***) only D1 and D5 of retrieval reveals no difference in PC representation between 6-week-old (adolescent) WT and Het [WT (9) = gray, Het (11) = orange]. ***a'***, ***b'***, Gap statistics test confirms that the optimal number of clusters is one (dashed line) for (***a'***) all days of retrieval and (***b'***) D1 and D5 of retrieval. ***c***, PCA for all days of retrieval reveals no difference in PC representation between adolescent WT ($n = 9$, gray) and adult WT ($n = 21$, black) in agreement with gap statistics (***c'***). ***d***, PCA for all days of retrieval shows adolescent Het clusters with adult Het-NR and adult WT, while Het-R cluster separately in agreement with an optimal number of two clusters (***d'***).

we performed the gap statistic test and determined that the optimal number of clusters between both genotypes was one (Fig. 7*a'*,*b'*).

To determine if there are developmental differences in retrieval trajectories between adolescence and adulthood within each genotype, we performed PCA on all metrics for all days (Fig. 7*c*) between adolescent and adult WT. We observed no distinct clusters between adolescent WT (gray) and adult WT (black; Fig. 7*c*) in agreement with gap statistics (Fig. 7*c'*). We

performed PCA on 13 metrics for all days (Fig. 7*d*) between adolescent and adult Het and adult WT. The majority of adolescent Het grouped with adult WT and adult Het-NR (Fig. 7*d*). However, one adolescent Het grouped with adult Het-R (Fig. 7*d*). These collective findings were supported by a gap statistic test that indicated there are an optimal number of two clusters (Fig. 7*d'*). The one adolescent Het clustered with Het-R and may indicate an adolescent Het at risk for regression over time. Future large-scale and longitudinal studies are needed to address these possibilities.

## Discussion

One of the major hurdles in diagnosing, understanding, and treating heterogeneous rare neuropsychiatric diseases is the lack of appropriate tools to robustly identify and classify complex morbidity over time and trials with individual resolution. While males are typically severely affected by *MECP2* dosage, females typically exhibit less severe and more heterogeneous phenotypes due to random X-chromosome inactivation and mosaic expression of the wild-type protein (Mykins et al., 2023). Thus, disease pathobiology caused by X-linked mutations, and by *MECP2*, is fundamentally different in females than males. Currently, there is minimal consensus on behavioral phenotypes in female Het mice, due to strain variations and lack of systematic analysis over different ages (Stearns et al., 2007; Garg et al., 2013; Samaco et al., 2013; Krishnan et al., 2017; Fagiolini et al., 2020; Ribeiro and MacDonald, 2020; Stevenson et al., 2021; Mykins et al., 2023). This is particularly important for designing better therapeutic interventions, as symptoms change over time in individual patients. Thus, developing and utilizing appropriate preclinical models are a necessary step. Here, we utilize a pose estimation algorithm and multidimensional analysis to characterize complex social cognition phenotypes in the heterogeneous population of adolescent and adult female *Mecp2*-heterozygous mice. We found a context- and age-specific regression in a subset of genotypic Het (Het-regressors, Het-R) and an age-specific delay in efficient behavior in another subset of genotypic Het (Het-non-regressors, Het-NR). These are novel results that finally achieve categorization of known, yet elusive, phenotypic heterogeneity in female X-linked disorders. Additionally, this is the first report to faithfully recapitulate regression and mild developmental delay in this apt preclinical model. This robust phenotyping in individual animals will be critical in determining the mechanisms connecting the known *MECP2* mosaicism, etiology, and disease progression in RTT over time, contexts, and physiological states. Future work will be needed to correlate heterogeneity in disease severity to known models of *MECP2* mutations and in different relevant behavioral models (Neul et al., 2008; Vashi and Justice, 2019; Ehrhart et al., 2021). The implications of such approaches will be in establishing robust and reproducible phenotyping of female mice across labs and providing biomarkers through longitudinal progression, physiological ages, and symptom severity with individual resolution, for better design for clinical interventions (Leonard et al., 2022).

### Pup retrieval task as an ethologically relevant task to study complex social cognition

Maternal behavior and pup retrieval tasks are well-established paradigms to study complex dyadic interactions focusing on neurological processes such as motivation, goal-directed movements, sensory processing and perception, social communication, and cognition (Wiesner and Sheard, 1933; Beach and Jaynes, 1956; Rosenblatt, 1967; Stern and Kolunie, 1989; Morgan et al., 1992; Stern, 1996; Champagne et al., 2007; Alsina-Llanes et al., 2015; Lonstein et al., 2015; Champagne and Curley, 2016; Stevenson et al., 2021). In our assay, naïve virgin female mice are sensitized to perform efficient pup retrieval. This scenario is akin to a person moving to a new place with roommates, and then motivated (or induced) to help/interact and care for newborn babies, without any prior experience, for short periods of time every day. It is a challenging task to navigate this complex social environment, with demands on cognitive flexibility, sensory processing and perception, and skilled goal-directed motor sequences. The purpose of using a complex social task, which taps into a combination of innate and learned neural circuitry, in this heterogeneous population is based on the premise that brains are evolved to engage and process complex environmental and social information; thus, the challenges will reveal dynamic and nuanced phenotypes in behaviors and neural processing over time and contexts, with reproducible variations in individual animals. Thus, such complexity is critical for a better utilization of preclinical models of human disorders.

At the neurobiology level, much is known about the contributions of specific brain regions and neuromodulatory systems to efficient maternal care (Hemel, 1973; Stern and Mackinnon, 1978; Ehret et al., 1987; Morgan et al., 1992; Keer and Stern, 1999; Cohen et al., 2011; Marlin et al., 2015; Kohl et al., 2017, 2018; Krishnan et al., 2017; Stevenson et al., 2021). The neural circuit mechanism by which naïve female mice adapt repetitive and stereotyped motions over days to efficiently retrieve scattered pups back to the nest is still under investigation. Perturbation in the dopaminergic mesolimbic pathway, ventral tegmental area, and medial prefrontal cortex serotonergic neurons impairs sequential pup retrieval and maternal behavior, suggesting a role for neuromodulation in motivation and reward of the pup (Hansen et al., 1991a,b, 1993; Hansen, 1994; Keer and Stern, 1999; Lonstein and De Vries, 2000; Wu et al., 2016; Gao et al., 2018, 2020; Holschbach et al., 2018; Xie et al., 2023). Dopaminergic activity increases in the ventral tegmental area and nucleus accumbens before pup contact and is required for reinforcement learning of pup retrieval over days in nulliparous female mice (Dai et al., 2022; Xie et al., 2023). Many of these established circuits have strong projections into higher cortical regions important for sensory motor integration. In this study, we have established metrics indicating that naive wild-type mice are adapting sensorimotor strategies to increase retrieval efficiency and potentially maximize the perceived social reward from pups. We hypothesize that mice prioritize sensory-dependent social reward of the pups in early days of retrieval and utilize higher cortical regions to consolidate motor sequences in later days. Future work using in vivo optogenetic manipulation combined with real-time pup retrieval dynamics to investigate the role of sensorimotor neural circuits is important to elucidate the underlying neural mechanisms of this relevant task.

### Neurobiology underlying sensory and social communication deficits in Rett syndrome

Long-standing work in the field RTT has identified disruptions in cortical and subcortical regions to be major factors for sensory, motor, and social communication deficits in predominantly *Mecp2-null male* rodent models. In *Mecp2*-heterozygous female mice, it is clear that tactile, auditory, and prefrontal cortices are dysregulated, likely giving rise to sensory and social phenotypes (Krishnan et al., 2017; Lau et al., 2020a; Achilly et al., 2021; Xu et al., 2022; Zhao et al., 2022; Rupert et al., 2023).

While previous studies have reported regression in motor skills and respiratory function over weeks in *Mecp2* rodent models, it remains unclear when and how regression in social skills arises over time within Het (Huang et al., 2016; Veeraragavan et al., 2016; Achilly et al., 2021). Here, using PCA and multidimensional analysis of a social cognition pup retrieval task, we observed adult Het-R are capable of adapting goal-directed trajectories toward pups in early days and regress as they are unable to consolidate or maintain these efficient trajectories in later days (Fig. 4). Compared with WT, Het-NR have a milder phenotype but are capable of consolidating efficient trajectories over time. Thus, Het-NR are able to compensate behaviorally, while Het-R are unable to consolidate social and sensory cues to execute efficient retrieval. Het-R regression is specific to age as we do not observe regression in adolescent Het (Mykins et al., 2023). We observe adult Het have increased interactions with novel pup stimuli and increased interactions with textures compared with WT controls, supporting that Het may have issues in cortical circuits for discriminating social and nonsocial cues (Stevenson et al., 2021; Xu et al., 2022; Zhao et al., 2022; Mykins et al., 2023). Severe cortical dysregulation in these circuits may result in the inability of Het-R to consolidate changes in sensory and social cues over time or prevent reinforcement learning of the reward of the pup, ultimately leading to regression. Furthermore, we have previously shown that in adult Het, pyramidal and PV neurons in the auditory cortex respond differentially to broad tones and pup calls, with lesser synaptic plasticity in the PV neurons (Lau et al., 2020a; Rupert et al., 2023). The mPFC and auditory cortex reciprocally provide feedback for auditory tone discrimination; thus, optogenetic manipulation studies between these regions in the context of pup retrieval are needed to further dissect the sensory perception and motor activation pathology in Het (Rodgers and DeWeese, 2014).

An alternative hypothesis is that Het-R may lose the ability to transfer to the next task of bringing the pups back to the nest, suggesting issues in brain regions important for motor planning and execution. In support of this hypothesis, we reported that adult Het exhibit abnormal maternal behavior task transition probabilities during pup retrieval (Stevenson et al., 2021). The locus ceruleus is important for mediating maternal behavior in mice via the global release of noradrenaline between transitioning of behavioral states, suggesting that dysfunction in these circuits could contribute to inefficient pup retrieval in Het. *Mecp2* mutant mice have profound physiological deficits in the locus ceruleus (Thomas and Palmiter, 1997; Taneja et al., 2009; Huang et al., 2016; Howell et al., 2017; Dvorkin and Shea, 2022).

### Molecular and cellular mechanisms driving phenotypic variations in Het

The random X-chromosome inactivation patterns across the body and brain and the type of *MECP2* mutations are major contributing factors to the behavioral, cellular, and molecular heterogeneity in RTT. *MECP2* expression increases in the postnatal brain and correlates with the timing of critical period plasticity of sensory cortices (LaSalle, 2001; Braunschweig et al., 2004; Durand et al., 2012; Krishnan et al., 2015; Picard and Fagiolini, 2019). Recently, we reported an intriguing observation of a precocious increase in *MECP2* expression within specific cell types in the primary somatosensory barrel cortex of adolescent Het that coincides with typical behavioral performance during sensory-relevant tasks (Mykins et al., 2023). We speculated that this may provide compensatory phenotypic benefits at an earlier age, while the inability to further increase *MECP2* levels in adulthood leads to regressive phenotypes (Mykins et al., 2023). Thus, mosaic *MECP2* expression or the inability to maintain cell-type-specific *MECP2* expression may underlie behavioral regression in our Het-R phenotype. Further in-depth analysis of cell-type-specific *MECP2* expression within our adult Het-NR and Het-R phenotypes is needed.

### Application of computational neuroethology tools to understand complex endophenotypes in neuropsychiatric disorders

Basic science research using behavioral endpoints in rodent models has identified key molecular therapeutic targets for treating core symptoms of RTT patients (Gogliotto et al., 2016, 2017; Tai et al., 2016; Li et al., 2017; Smith-Hicks et al., 2017; O'Leary et al., 2018; Garré et al., 2020; Gualniera et al., 2021). Recently, the FDA approved Trofinetide as the first drug for treating disease pathology in RTT patients (Neul et al., 2022). However, the challenge of predictive power to determine which therapeutic interventions will work when and how for individual patients is the holy grail of personalized medicine. In the preclinical stages, systematic characterization of etiologically relevant behaviors in female mouse models of X-linked disorders is critical to move forward in understanding the mechanisms of complex endophenotypes in patients and identifying the efficacy of clinical therapeutics.

The emergence of computational neuroethology tools now provides a unique opportunity to shift the paradigm from single constrained reductionist behaviors to studying complex behavior in free-moving animals (Shemesh and Chen, 2023). The free availability of open-source marker-less deep learning and machine learning software for behavioral segmentation using pose makes it attainable and affordable to analyze complex free-moving behaviors (Berman et al., 2014; Wiltschko et al., 2015; Mathis et al., 2018; Nath et al., 2019; Pereira et al., 2019, 2022; Nilsson et al., 2020; Hsu and Yttri, 2021; Segalin et al., 2021; Sun et al., 2021; Goodwin et al., 2022; Lauer et al., 2022; Luxem et al., 2022; Winters et al., 2022). Using unsupervised machine learning, multiple groups have identified subtle, novel behavioral differences in wild-type mice and mouse models for Alzheimer's disease, stroke, and autism, highlighting the importance of these approaches for modeling complex endophenotypes of neurological disorders in preclinical animal studies (Wiltschko et al., 2020; Huang et al., 2021; Segalin et al., 2021; Klibaite et al., 2022; Luxem et al., 2022; Weber et al., 2022).

As the field is broadening to include studies in female *Mecp2*-heterozygous mice, there is a need for efficient methods to assess phenotypic severity. Guy et al. (2007) established a severity scoring system. However, this system is specific to rapid phenotypic onset in null males and does not apply to female Het, which have a slow and progressive phenotypic onset and display considerable heterogeneity in behaviors. Additionally, this system requires extensive monitoring over time and introduces observer bias due to its qualitative nature. We highlight the advantage of using multidimensional analysis of behavior (Fig. 3) for phenotypic characterization (Fig. 4) and provide a means to score and evaluate Het phenotypic severity. The brevity and minimal resources for this approach make it an attractive approach to evaluate the efficacy of therapeutics in ongoing preclinical studies in the apt female Het mouse model.

## References

Achilly NP, Wang W, Zoghbi HY (2021) Presymptomatic training mitigates functional deficits in a mouse model of Rett syndrome. Nature 592:596–600.

Alsina-Llanes M, De Brun V, Olazábal DE (2015) Development and expression of maternal behavior in naïve female C57BL/6 mice: ontogeny of maternal behavior in mice. Dev Psychobiol 57:189–200.
Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet 23:185–188.
Batschelet E (1981) *Circular statistics in biology*. New York, NY: Academic Press.
Beach FA, Jaynes J (1956) Studies of maternal retrieving in rats. III. Sensory cues involved in the lactating female's response to her young. Behaviour 10:104–124.
Benhamou S (2004) How to reliably estimate the tortuosity of an animal's path. J Theor Biol 229:209–220.
Benhamou S (2014) Of scales and stationarity in animal movements. Ecol Lett 17:261–272.
Berman GJ, Choi DM, Bialek W, Shaevitz JW (2014) Mapping the stereotyped behaviour of freely moving fruit flies. J R Soc Interface 11:20140672.
Bovet P, Benhamou S (1988) Spatial analysis of animals' movements using a correlated random walk model. J Theor Biol 131:419–433.
Braunschweig D, Simcox T, Samaco RC, LaSalle JM (2004) X-chromosome inactivation ratios affect wild-type MeCP2 expression within mosaic Rett syndrome and Mecp2-/+ mouse brain. Hum Mol Genet 13:1275–1286.
Buchanan CB, Stallworth JL, Scott AE, Glaze DG, Lane JB, Skinner SA, Tierney AE, Percy AK, Neul JL, Kaufmann WE (2019) Behavioral profiles in Rett syndrome: data from the natural history study. Brain Dev 41:123–134.
Champagne FA, Curley JP (2016) Plasticity of the maternal brain across the lifespan: plasticity of the maternal brain across the lifespan. New Dir Child Adolesc Dev 2016:9–21.
Champagne FA, Curley JP, Keverne EB, Bateson PPG (2007) Natural variations in postpartum maternal care in inbred and outbred mice. Physiol Behav 91:325–334.
Charman T, Cass H, Owen L, Wigram T, Slonims V, Weeks L, Wisbeach A, Reilly S (2002) Regression in individuals with Rett syndrome. Brain Dev 24:281–283.
Cheung A, Zhang S, Stricker C, Srinivasan MV (2007) Animal navigation: the difficulty of moving in a straight line. Biological Cybernetics 97:47–61.
Cohen L, Rothschild G, Mizrahi A (2011) Multisensory integration of natural odors and sounds in the auditory cortex. Neuron 72:357–369.
Cosentino L, Vigli D, Franchi F, Laviola G, De Filippis B (2019) Rett syndrome before regression: a time window of overlooked opportunities for diagnosis and intervention. Neurosci Biobehav Rev 107:115–135.
Dai B, Sun F, Tong X, Ding Y, Kuang A, Osakada T, Li Y, Lin D (2022) Responses and functions of dopamine in nucleus accumbens core during social behaviors. Cell Rep 40:111246.
Djukic A, Valicenti McDermott M (2012) Social preferences in Rett syndrome. Pediatr Neurol 46:240–242.
Djukic A, Valicenti McDermott M, Mavrommatis K, Martins CL (2012) Rett syndrome: basic features of visual processing—a pilot study of eye-tracking. Pediatr Neurol 47:25–29.
Dunlap AG, Besosa C, Pascual LM, Chong KK, Walum H, Kacsoh DB, Tankeu BB, Lu K, Liu RC (2020) Becoming a better parent: mice learn sounds that improve a stereotyped maternal behavior. Horm Behav 124:104779.
Durand S, Patrizi A, Quast KB, Hachigian L, Pavlyuk R, Saxena A, Carninci P, Hensch TK, Fagiolini M (2012) NMDA receptor regulation prevents regression of visual cortical function in the absence of Mecp2. Neuron 76:1078–1090.
Dvorkin R, Shea SD (2022) Precise and pervasive phasic bursting in locus coeruleus during maternal behavior in mice. J Neurosci 42:2986–2999.
Ehret G, Koch M, Haack B, Markl H (1987) Sex and parental experience determine the onset of an instinctive behavior in mice. Naturwissenschaften 74:47–47.
Ehrhart F, et al. (2021) A catalogue of 863 Rett-syndrome-causing MECP2 mutations and lessons learned from data integration. Sci Data 8:10.
Einspieler C, Marschik PB (2019) Regression in Rett syndrome: developmental pathways to its onset. Neurosci Biobehav Rev 98:320–332.
Fagiolini M, Patrizi A, LeBlanc J, Jin L-W, Maezawa I, Sinnett S, Gray SJ, Molholm S, Foxe JJ (2020) Intellectual and developmental disabilities research centers: a multidisciplinary approach to understand the pathogenesis of methyl-CpG binding protein 2-related disorders. Neuroscience 445:190–206.
Gao J, Nie L, Li Y, Li M (2020) Serotonin 5-HT2A and 5-HT2C receptors regulate rat maternal behavior through distinct behavioral and neural mechanisms. Neuropharmacology 162:107848.
Gao J, Wu R, Davis C, Li M (2018) Activation of 5-HT2A receptor disrupts rat maternal behavior. Neuropharmacology 128:96–105.
Garg SK, Lioy DT, Cheval H, McGann JC, Bissonnette JM, Murtha MJ, Foust KD, Kaspar BK, Bird A, Mandel G (2013) Systemic delivery of MeCP2 rescues behavioral and cellular deficits in female mouse models of Rett syndrome. J Neurosci 33:13612–13620.
Garré JM, Silva HM, Lafaille JJ, Yang G (2020) P2x7 receptor inhibition ameliorates dendritic spine pathology and social behavioral deficits in Rett syndrome mice. Nat Commun 11:1784.
Goffin D, et al. (2012) Rett syndrome mutation MeCP2 T158A disrupts DNA binding, protein stability and ERP responses. Nat Neurosci 15:274–283.
Gogliotti RG, et al. (2017) mGlu $_7$ potentiation rescues cognitive, social, and respiratory phenotypes in a mouse model of Rett syndrome. Sci Transl Med 9:eaai7459.
Gogliotti RG, et al. (2016) mGlu $_5$ positive allosteric modulation normalizes synaptic plasticity defects and motor phenotypes in a mouse model of Rett syndrome. Hum Mol Genet 25:1990–2004.
Goin-Kochel RP, Esler AN, Kanne SM, Hus V (2014) Developmental regression among children with autism spectrum disorder: onset, duration, and effects on functional outcomes. Res Autism Spectr Disord 8:890–898.
Goin-Kochel RP, Trinh S, Barber S, Bernier R (2017) Gene disrupting mutations associated with regression in autism spectrum disorder. J Autism Dev Disord 47:3600–3607.
Goldberg WA, Thorsen KL, Osann K, Spence MA (2008) Use of home videotapes to confirm parental reports of regression in autism. J Autism Dev Disord 38:1136–1146.
Gonzales ML, LaSalle JM (2010) The role of MeCP2 in brain development and neurodevelopmental disorders. Curr Psychiatry Rep 12:127–134.
Goodwin NL, Nilsson SRO, Choong JJ, Golden SA (2022) Toward the explainability, transparency, and universality of machine learning for behavioral classification in neuroscience. Curr Opin Neurobiol 73:102544.
Gualniera L, Singh J, Fiori F, Santosh P (2021) Emotional behavioural and autonomic dysregulation (EBAD) in Rett syndrome – EDA and HRV monitoring using wearable sensor technology. J Psychiatr Res 138:186–193.
Guy J, Gan J, Selfridge J, Cobb S, Bird A (2007) Reversal of neurological defects in a mouse model of Rett syndrome. Science 315:1143–1147.
Guy J, Hendrich B, Holmes M, Martin JE, Bird A (2001) A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nat Genet 27:322–326.
Hagberg B, Aicardi J, Dias K, Ramos O (1983) A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. Ann Neurol 14:471–479.
Han Z-A, Jeon HR, Kim SW, Park JY, Chung HJ (2012) Clinical characteristics of children with Rett syndrome. Ann Rehabil Med 36:334.
Hansen S (1994) Maternal behavior of female rats with 6-OHDA lesions in the ventral striatum: characterization of the pup retrieval deficit. Physiol Behav 55:615–620.
Hansen S, Bergvall ÅH, Nyiredi S (1993) Interaction with pups enhances dopamine release in the ventral striatum of maternal rats: a microdialysis study. Pharmacol Biochem Behav 45:673–676.
Hansen S, Harthon C, Wallin E, Löfberg L, Svensson K (1991a) Mesotelencephalic dopamine system and reproductive behavior in the female rat: effects of ventral tegmental 6-hydroxydopamine lesions on maternal and sexual responsiveness. Behav Neurosci 105:588–598.
Hansen S, Harthon C, Wallin E, Löfberg L, Svensson K (1991b) The effects of 6-OHDA-induced dopamine depletions in the ventral or dorsal striatum on maternal and sexual behavior in the female rat. Pharmacol Biochem and Behav 39:71–77.
Hemel SBV (1973) Pup retrieving as a reinforcer in nulliparous mice [1]. J Exp Anal Behav 19:233–238.
Holschbach MA, Vitale EM, Lonstein JS (2018) Serotonin-specific lesions of the dorsal raphe disrupt maternal aggression and caregiving in postpartum rats. Behav Brain Res 348:53–64.
Howell CJ, Sceniak MP, Lang M, Krakowiecki W, Abouelsoud FE, Lad SU, Yu H, Katz DM (2017) Activation of the medial prefrontal cortex reverses cognitive and respiratory symptoms in a mouse model of Rett syndrome. Eneuro 4:ENEURO.0277-17.2017.

Hsu AI, Yttri EA (2021) B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors. Nat Commun 12:5188.
Huang R, et al. (2021) Machine learning classifies predictive kinematic features in a mouse model of neurodegeneration. Sci Rep 11:3950.
Huang T-W, Kochukov MY, Ward CS, Merritt J, Thomas K, Nguyen T, Arenkiel BR, Neul JL (2016) Progressive changes in a distributed neural circuit underlie breathing abnormalities in mice lacking MeCP2. J Neurosci 36:5572–5586.
Insafutdinov E, Pishchulin L, Andres B, Andriluka M, Schiele B (2016) Deepercut: a deeper, stronger, and faster multi-person pose estimation model. In: *Computer vision – ECCV 2016*, Vol. 9910 (Leibe B, Matas J, Sebe N, Welling M, eds), pp 34–50. Springer International Publishing.
Ito-Ishida A, Ure K, Chen H, Swann JW, Zoghbi HY (2015) Loss of MeCP2 in parvalbumin-and somatostatin-expressing neurons in mice leads to distinct Rett syndrome-like phenotypes. Neuron 88:651–658.
Keer SE, Stern JM (1999) Dopamine receptor blockade in the nucleus accumbens inhibits maternal retrieval and licking, but enhances nursing behavior in lactating rats. Physiol Behav 67:659–669.
Kerr AM (1987) Report on the Rett syndrome workshop: Glasgow, Scotland, 24–25 May 1986. J Intellect Disabil Res 31:93–113.
Kerr A (1995) Early clinical signs in the Rett disorder. Neuropediatrics 26: 67–71.
Key AP, Jones D, Peters S (2019) Spoken word processing in Rett syndrome: evidence from event-related potentials. Int J Dev Neurosci 73:26–31.
Kirby RS, Lane JB, Childers J, Skinner SA, Annese F, Barrish JO, Glaze DG, MacLeod P, Percy AK (2010) Longevity in Rett syndrome: analysis of the North American Database. J Pediatr 156:135–138.e1.
Kitamura T, Imafuku M (2015) Behavioural mimicry in flight path of Batesian intraspecific polymorphic butterfly *Papilio polytes*. Proc R Soc B Biol Sci 282:20150483.
Klibaite U, Kislin M, Verpeut JL, Bergeler S, Sun X, Shaevitz JW, Wang SS-H (2022) Deep phenotyping reveals movement phenotypes in mouse neurodevelopmental models. Mol Autism 13:12.
Koch M, Ehret G (1989) Estradiol and parental experience, but not prolactin are necessary for ultrasound recognition and pup-retrieving in the mouse. Physiol Behav 45:771–776.
Kohl J, Autry AE, Dulac C (2017) The neurobiology of parenting: a neural circuit perspective. BioEssays 39:1–11.
Kohl J, et al. (2018) Functional circuit architecture underlying parental behaviour. Nature 556:326–331.
Krishnan K, Lau BYB, Ewall G, Huang ZJ, Shea SD (2017) MECP2 regulates cortical plasticity underlying a learned behaviour in adult female mice. Nat Commun 8:14077.
Krishnan K, Wang B-S, Lu J, Wang L, Maffei A, Cang J, Huang ZJ (2015) MeCP2 regulates the timing of critical period plasticity that shapes functional connectivity in primary visual cortex. Proc Natl Acad Sci U S A 112:E4782–E4791.
Lambert S, Maystadt I, Boulanger S, Vrielynck P, Destrée A, Lederer D, Moortgat S (2016) Expanding phenotype of p.Ala140Val mutation in MECP2 in a 4 generation family with X-linked intellectual disability and spasticity. Eur J Med Genet 59:522–525.
LaSalle JM (2001) Quantitative localization of heterogeneous methyl-CpG-binding protein 2 (MeCP2) expression phenotypes in normal and Rett syndrome brain by laser scanning cytometry. Hum Mol Genet 10:1729–1740.
Lau BYB, Krishnan K, Huang ZJ, Shea SD (2020a) Maternal experience-dependent cortical plasticity in mice is circuit- and stimulus-specific and requires MECP2. J Neurosci 40:1514–1526.
Lau BYB, Layo DE, Emery B, Everett M, Kumar A, Stevenson P, Reynolds KG, Cherosky A, Bowyer SAHet al. (2020b) Lateralized expression of cortical perineuronal nets during maternal experience is dependent on MECP2. Eneuro 7:ENEURO.0500-19.2020.
Lauer J, et al. (2022) Multi-animal pose estimation, identification and tracking with DeepLabCut. Nat Methods 19:496–504.
LeBlanc JJ, DeGregorio G, Centofante E, Vogel-Farley VK, Barnes K, Kaufmann WE, Fagiolini M, Nelson CA (2015) Visual evoked potentials detect cortical processing deficits in Rett syndrome: VEP in Rett syndrome. Ann Neurol 78:775–786.
Lee L-J, Tsytsarev V, Erzurumlu RS (2017) Structural and functional differences in the barrel cortex of *Mecp2* null mice. J Comp Neurol 525:3951–3961.
Leonard H, Bower C (1998) Is the girl with Rett syndrome normal at birth? Dev Med Child Neurol 40:115–121.
Leonard H, Gold W, Samaco R, Sahin M, Benke T, Downs J (2022) Improving clinical trial readiness to accelerate development of new therapeutics for Rett syndrome. Orphanet J Rare Dis 17:108.
Li W, Bellot-Saez A, Phillips ML, Yang T, Longo FM, Pozzo-Miller L (2017) A small-molecule TrkB ligand restores hippocampal synaptic plasticity and object location memory in Rett syndrome mice. Dis Model Mech 10:837–845.
Lo F-S, Blue ME, Erzurumlu RS (2016) Enhancement of postsynaptic $GABA_A$ and extrasynaptic NMDA receptor-mediated responses in the barrel cortex of *Mecp2* -null mice. J Neurophysiol 115:1298–1306.
Lonstein JS, De Vries GJ (2000) Sex differences in the parental behavior of rodents. Neurosci Biobehav Rev 24:669–686.
Lonstein JS, Lévy F, Fleming AS (2015) Common and divergent psychobiological mechanisms underlying maternal behaviors in non-human and human mammals. Horm Behav 73:156–185.
Luxem K, Mocellin P, Fuhrmann F, Kürsch J, Miller SR, Palop JJ, Remy S, Bauer P (2022) Identifying behavioral structure from deep variational embeddings of animal motion. Commun Biol 5:1267.
Luxem K, Sun JJ, Bradley SP, Krishnan K, Yttri E, Zimmermann J, Pereira TD, Laubach M (2023) Open-source tools for behavioral video analysis: setup, methods, and best practices. eLife 12:e79305.
Marlin BJ, Mitre M, D'amour JA, Chao MV, Froemke RC (2015) Oxytocin enables maternal behaviour by balancing cortical inhibition. Nature 520: 499–504.
Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M (2018) DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci 21:1281–1289.
McLean DJ, Skowron Volponi MA (2018) trajr: an R package for characterisation of animal trajectories. Ethology 124:440–448.
McVicar KA, Shinnar S (2004) Landau–Kleffner syndrome, electrical status epilepticus in slow wave sleep, and language regression in children. Ment Retard Dev Disabil Res Rev 10:144–149.
Merbler AM, Byiers BJ, Hoch J, Dimian AC, Barney CC, Feyma TJ, Beisang AA, Bartolomucci A, Symons FJ (2020) Preliminary evidence that resting state heart rate variability predicts reactivity to tactile stimuli in Rett syndrome. J Child Neurol 35:42–48.
Morgan HD, Fleming AS, Stern JM (1992) Somatosensory control of the onset and retention of maternal responsiveness in primiparous Sprague–Dawley rats. Physiol Behav 51:549–555.
Mykins M, et al. (2023) Wild-type MECP2 expression coincides with age-dependent sensory phenotypes in a female mouse model for Rett syndrome. J Neurosci Res 101:1236–1258.
Nath T, Mathis A, Chen AC, Patel A, Bethge M, Mathis MW (2019) Using DeepLabCut for 3D markerless pose estimation across species and behaviors. Nat Protoc 14:2152–2176.
Neul JL, Benke TA, Marsh ED, Suter B, Silveira L, Fu C, Peters SU, Percy AK, Group RSNHS (2023) *Top Caregiver Concerns in Rett syndrome and related disorders: Data from the US Natural History Study* [Preprint]. In Review.
Neul JL, Fang P, Barrish J, Lane J, Caeg EB, Smith EO, Zoghbi H, Percy A, Glaze DG (2008) Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome. Neurology 70:1313–1321.
Neul JL, et al. (2010) Rett syndrome: revised diagnostic criteria and nomenclature. Ann Neurol 68:944–950.
Neul JL, Percy AK, Benke TA, Berry-Kravis EM, Glaze DG, Peters SU, Jones NE, Youakim JM (2022) Design and outcome measures of LAVENDER, a phase 3 study of trofinetide for Rett syndrome. Contemp Clin Trials 114:106704.
Nilsson SR, et al. (2020) Simple Behavioral Analysis (SimBA) – an open source toolkit for computer classification of complex social behaviors in experimental animals. *bioRxiv*, 2020.04.19.049452.
Nomura Y (2005) Early behavior characteristics and sleep disturbance in Rett syndrome. Brain Dev 27:S35–S42.
Nomura Y, Segawa M (1990) Clinical features of the early stage of the Rett syndrome. Brain Dev 12:16–19.
O'Leary HM, et al. (2018) Placebo-controlled crossover assessment of mecasermin for the treatment of Rett syndrome. Ann Clin Transl Neurol 5: 323–332.
Orefice LL, Zimmerman AL, Chirila AM, Sleboda SJ, Head JP, Ginty DD (2016) Peripheral mechanosensory neuron dysfunction underlies tactile and behavioral deficits in mouse models of ASDs. Cell 166:299–313.
Ozonoff S, Gangi D, Hanzel EP, Hill A, Hill MM, Miller M, Schwichtenberg AJ, Steinfeld MB, Parikh C, Iosif A-M (2018) Onset patterns in autism:

variation across informants, methods, and timing: measuring onset patterns in ASD. Autism Res 11:788–797.
Ozonoff S, Iosif A-M (2019) Changing conceptualizations of regression: what prospective studies reveal about the onset of autism spectrum disorder. Neurosci Biobehav Rev 100:296–304.
Pereira TD, Aldarondo DE, Willmore L, Kislin M, Wang SSH, Murthy M, Shaevitz JW (2019) Fast animal pose estimation using deep neural networks. Nat Methods 16:117–125.
Pereira TD, et al. (2022) SLEAP: a deep learning system for multi-animal pose tracking. Nat Methods 19:486–495.
Peters SU, Gordon RL, Key AP (2015) Induced gamma oscillations differentiate familiar and novel voices in children with *MECP2* duplication and Rett syndromes. J Child Neurol 30:145–152.
Picard N, Fagiolini M (2019) MeCP2: an epigenetic regulator of critical periods. Curr Opin Neurobiol 59:95–101.
Popovitz J, Mysore SP, Adwanikar H (2021) Neural markers of vulnerability to anxiety outcomes after traumatic brain injury. J Neurotrauma 38: 1006–1022.
Rett A (1966) Ber ein eigenartiges hirnatrophisches syndrom bei hyperammonemia in kindersalter. (On an unusual brain atrophy syndrome with hyperammonemia in childhood.). Wein Med Wochensch 116:26.
Ribeiro MC, MacDonald JL (2020) Sex differences in *Mecp2*-mutant Rett syndrome model mice and the impact of cellular mosaicism in phenotype development. Brain Res 1729:146644.
Rodgers CC, DeWeese MR (2014) Neural correlates of task switching in prefrontal cortex and primary auditory cortex in a novel stimulus selection task for rodents. Neuron 82:1157–1170.
Rosenblatt JS (1967) Nonhormonal basis of maternal behavior in the rat. Science 156:1512–1513.
Rupert DD, Pagliaro AH, Choe J, Shea SD (2023) Selective deletion of methyl CpG binding protein 2 from parvalbumin interneurons in the auditory cortex delays the onset of maternal retrieval in mice. J Neurosci 43:6745–6759.
Samaco RC, McGraw CM, Ward CS, Sun Y, Neul JL, Zoghbi HY (2013) Female Mecp2+/− mice display robust behavioral deficits on two different genetic backgrounds providing a framework for pre-clinical studies. Hum Mol Genet 22:96–109.
Schüle B, Armstrong D, Vogel H, Oviedo A, Francke U (2008) Severe congenital encephalopathy caused by MECP2 null mutations in males: central hypoxia and reduced neuronal dendritic structure. Clin Genet 74:116–126.
Segalin C, Williams J, Karigo T, Hui M, Zelikowsky M, Sun JJ, Perona P, Anderson DJ, Kennedy A (2021) The mouse action recognition system (MARS) software pipeline for automated analysis of social behaviors in mice. eLife 10:e63720.
Shemesh Y, Chen A (2023) A paradigm shift in translational psychiatry through rodent neuroethology. Mol Psychiatry 28:993–1003.
Smith-Hicks CL, et al. (2017) Randomized open-label trial of dextromethorphan in Rett syndrome. Neurology 89:1684–1690.
Stallworth JL, et al. (2019) Hand stereotypies: lessons from the Rett syndrome natural history study. Neurology 92:e2594–e2603.
Stearns NA, Schaevitz LR, Bowling H, Nag N, Berger UV, Berger-Sweeney J (2007) Behavioral and anatomical abnormalities in Mecp2 mutant mice: a model for Rett syndrome. Neuroscience 146:907–921.
Stern JM (1996) Somatosensation and maternal care in Norway rats. In: *Advances in the study of behavior* (Rosenblatt JS, Snowdon CT, eds), Vol. 25, pp 243–294. Elsevier.
Stern JM, Kolunie JM (1989) Perioral anesthesia disrupts maternal behavior during early lactation in Long–Evans rats. Behav Neural Biol 52:20–38.
Stern JM, Mackinnon DA (1978) Sensory regulation of maternal behavior in rats: effects of pup age. Dev Psychobiol 11:579–586.
Stevenson PK, Casenhiser DM, Lau BYB, Krishnan K (2021) Systematic analysis of goal-related movement sequences during maternal behaviour in a female mouse model for Rett syndrome. Eur J Neurosci 54: 4528–4549.
Su S-H, Kao F-C, Huang Y-B, Liao W (2015) MeCP2 in the rostral striatum maintains local dopamine content critical for psychomotor control. J Neurosci 35:6209–6220.
Sun JJ, Kennedy A, Zhan E, Anderson DJ, Yue Y, Perona P (2021) Task Programming: Learning Data Efficient Behavior Representations. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2875–2884.
Symons FJ, Barney CC, Byiers BJ, McAdams BD, Foster SXYL, Feyma TJ, Wendelschafer-Crabb G, Kennedy WR (2019) A clinical case–control comparison of epidermal innervation density in Rett syndrome. Brain Behav 9:e01285.
Tai DJC, Liu YC, Hsu WL, Ma YL, Cheng SJ, Liu SY, Lee EHY (2016) MeCP2 SUMOylation rescues Mecp2-mutant-induced behavioural deficits in a mouse model of Rett syndrome. Nat Commun 7:10552.
Tammimies K (2019) Genetic mechanisms of regression in autism spectrum disorder. Neurosci Biobehav Rev 102:208–220.
Tanas JK, Kerr DD, Wang L, Rai A, Wallaard I, Elgersma Y, Sidorov MS (2022) Multidimensional analysis of behavior predicts genotype with high accuracy in a mouse model of Angelman syndrome. Transl Psychiatry 12:426.
Taneja P, Ogier M, Brooks-Harris G, Schmid DA, Katz DM, Nelson SB (2009) Pathophysiology of locus coeruleus neurons in a mouse model of Rett syndrome. J Neurosci 29:12187–12195.
Thomas SA, Palmiter RD (1997) Impaired maternal behavior in mice lacking norepinephrine and epinephrine. Cell 91:583–592.
Thurm A, Powell EM, Neul JL, Wagner A, Zwaigenbaum L (2018) Loss of skills and onset patterns in neurodevelopmental disorders: understanding the neurobiological mechanisms: onset patterns in neurodevelopmental disorders. Autism Res 11:212–222.
Vashi N, Justice MJ (2019) Treating Rett syndrome: from mouse models to human therapies. Mamm Genome 30:90–110.
Veeraragavan S, et al. (2016) Loss of MeCP2 in the rat models regression, impaired sociability and transcriptional deficits of Rett syndrome. Hum Mol Genet 25:3284–3302.
Weber RZ, Mulders G, Kaiser J, Tackenberg C, Rust R (2022) Deep learning-based behavioral profiling of rodent stroke recovery. BMC Biol 20:232.
Wickham H (2016) *Ggplot2*. New York: Springer-Verlag.
Wiesner BP, Sheard NM (1933) Maternal behavior in the rat. *Oliver & Boyd*.
Wiltschko AB, Johnson MJ, Iurilli G, Peterson RE, Katon JM, Pashkovski SL, Abraira VE, Adams RP, Datta SR (2015) Mapping sub-second structure in mouse behavior. Neuron 88:1121–1135.
Wiltschko AB, Tsukahara T, Zeine A, Anyoha R, Gillis WF, Markowitz JE, Peterson RE, Katon J, Johnson MJ, Datta SR (2020) Revealing the structure of pharmacobehavioral space through motion sequencing. Nat Neurosci 23:1433–1443.
Winters C, Gorssen W, Ossorio-Salazar VA, Nilsson S, Golden S, D'Hooge R (2022) Automated procedure to assess pup retrieval in laboratory mice. Sci Rep 12:1663.
Wu R, Gao J, Chou S, Davis C, Li M (2016) Behavioral, pharmacological and neuroanatomical analysis of serotonin 2C receptor agonism on maternal behavior in rats. Psychoneuroendocrinology 73:252–262.
Xie Y, Huang L, Corona A, Pagliaro AH, Shea SD (2023) A dopaminergic reward prediction error signal shapes maternal behavior in mice. Neuron 111:557–570.e7.
Xu P, Yue Y, Su J, Sun X, Du H, Liu Z, Simha R, Zhou J, Zeng C, Lu H (2022) Pattern decorrelation in the mouse medial prefrontal cortex enables social preference and requires MeCP2. Nat Commun 13:3899.
Yamamoto K, Gris KV, Sotelo Fonseca JE, Gharagozloo M, Mahmoud S, Simard C, Houle-Martel D, Cloutier T, Gris P, Gris D (2018) Exhaustive multi-parametric assessment of the behavioral array of daily activities of mice using cluster and factor analysis. Frontiers in Behav Neurosci 12:187.
Zhao Z, et al. (2022) Encoding of social novelty by sparse GABAergic neural ensembles in the prelimbic cortex. Sci Adv 8:eabo4884.