## **Statistical methods**

Linear mixed-effects model (lmer function) was conducted to assess the effects of cage- and mouse-related variables on DAHT using the lme4 [9] package in R (Version 4.2.2) with the Nelder-Mead optimizer. Intraclass correlation coefficients (ICC) were obtained via the performance [10] package. The alpha level for all tests was *p* = .05.

First, we determined the inclusion or exclusion of independent variables using their ICCs from null models, i.e., a model of one independent variable against the DAHT. Variables with low ICCs (i.e., below .05) were excluded from the final model. Then, we progressively added effects and interactions of independent variables to the time-only model (i.e., model of only time and DAHT) in order from lowest (1) to highest level (3). Deviance testing via ANOVA was conducted following each addition of effects or interactions to assess whether its inclusion improved or worsened the time-only model. Since all independent variables were time-invariant, i.e., unchanging with the progression of time, they were set as fixed effects. DAHT was set as a random effect due to being time-variant.

# **Results and discussion**

A three-level hierarchical study was used to measure the daily average hold time (DAHT) over 62 days in 24 mice (14 WT, 10 HD) from 9 cages. The exact distribution of WT and HD mice across the cages is detailed in Table 2.

**Table 2. Distribution of WT and HD mice across cages.**

| **Cage** | **Number of WT mice** | **Number of HD mice** | **Total number of mice** |
| --- | --- | --- | --- |
| 1 | 2 | 0 | 2 |
| 2 | 2 | 1 | 3 |
| 3 | 1 | 1 | 2 |
| 4 | 1 | 2 | 3 |
| 5 | 1 | 2 | 3 |
| 6 | 1 | 1 | 2 |
| 7 | 2 | 1 | 3 |
| 8 | 1 | 2 | 3 |
| 9 | 3 | 0 | 3 |
| **Total** | **14** | **10** |  |

The ICC values from null models of all independent variables are presented in Table 3. Housing density was excluded from the final model due to its low ICC (ICC = .01). The greater number of three-mice groupings compared to two-mice groupings (6 > 3) may have contributed to the low ICC value. 16% and 19% of the variation in DAHT could be attributed to between-cage and between-genotype variance, respectively. Hence, they were retained for further analysis in the model.

**Table 3. ICC values from null models of all independent variables.**

| **Level** | **Variable** | **Adjusted ICC** |
| --- | --- | --- |
| Cage | Cage | .160 |
|  | Housing density | .010 |
| Mouse | Genotype | .190 |

The results from deviance testing are presented in Table 4. Including the fixed effect of mouse genotype and its interaction with time decreased deviance, significantly improving the model’s fit (*p* < .05). Including the fixed effect of cage appeared to decrease deviance, but the improvement was statistically insignificant. However, since its addition did not detriment the fit of the model, it was still included in the final model to examine its contribution to variation.

**Table 4. Deviance testing of addition of effects and interactions to time-only model.**

| **Model** | **Added effect/interaction** | **Deviance** | **Chi-squared ()** |
| --- | --- | --- | --- |
| Time-only |  | -1348.200 |  |
|  |  |  |  |
| Time, level 2 | Fixed effect of genotype, interaction between genotype and time | -1357.000 | 8.820\* |
| Time, levels 2-3 | Fixed effect of genotype, interaction between genotype and time, fixed effect of cage | -1369.000 | 12.030 |

**\*** *p* < .05

Table 5 displays the results of the final model. There was a significant relationship between mouse genotype and DAHT (*p* < .001). In particular, the significant interaction between time and mouse genotype (*p* < .05) suggests that over 62 days, the motor learning of WT mice was better than HD mice. These results support Hypothesis 1, agreeing with previous ANOVA results and existing literature involving water-reaching tasks [4,6] that HD mice exhibit motor impairment.

Additionally, with cage 8 as reference, the remaining did not have a significant relationship with DAHT (*p* > .05). These results support Hypothesis 2, suggesting that mice can complete behavioral tasks in their home environments without significantly compromising the collection of data. In turn, this also supports the design of automated home cage experiments in general, alongside its advantages of reduced stressors and more naturalistic pacing [4].

**Table 5. Final model examining effects of cage and mouse genotype on DAHT.**

|  | **Variable value** | **Estimated slope** | **Standard error** | ***t* values (*df* = 24)** |
| --- | --- | --- | --- | --- |
| (Intercept) |  | 0.249\*\* | 0.078 | 0.004 |
| Time |  | 0.002 | 0.002 | 1.326 |
| Genotype |  | 0.204\*\* | 0.062 | 3.273 |
| Time:genotype |  | 0.005\* | 0.002 | 2.264 |
| Cage | 8 (Reference) |  |  |  |
|  | 1 | 0.014 | 0.125 | 0.111 |
|  | 2 | 0.022 | 0.108 | 0.202 |
|  | 3 | -0.155 | 0.119 | -1.305 |
|  | 4 | 0.068 | 0.106 | 0.643 |
|  | 5 | 0.120 | 0.106 | 1.131 |
|  | 6 | -0.097 | 0.119 | -0.819 |
|  | 7 | -0.177 | 0.108 | -1.645 |
|  | 9 | -0.182 | 0.114 | -1.602 |

**\*** *p* < .05

**\*\*** *p* < .01