# Does a 60-Second Breathing Exercise Improve Fine-Motor Precision?

Hiro Naito, Catalina Villouta, Odysseus Kaloidas, Nathaniel del Rosario

## 1 Introduction

Fine-motor precision underlies a wide range of everyday tasks, from surgical procedures to handwriting to operating a computer. While interventions aimed at improving motor performance have attracted substantial research interest, the question of whether a brief, scalable breathing exercise can produce a measurable effect in healthy adults remains open. Prior work has demonstrated motor benefits from breathing and mindfulness protocols lasting five to thirty minutes (Yadav & Mutha, 2016; Tang et al., 2007; Buchanan & Janelle, 2022), but the feasibility of a 60-second protocol has not been established. We designed a randomized controlled trial to test precisely this question: does a 60-second guided breathing exercise improve fine-motor cursor precision in a cursor line-drawing task?

We hypothesized that the intervention would reduce drawing error, operating through three plausible mechanisms: autonomic regulation via vagal tone enhancement (Laborde et al., 2022; Magnon et al., 2021; Zaccaro et al., 2018), improved attentional focus through rhythmic concentration (Larson et al., 2013), and reduced performance anxiety. Under the null hypothesis, we expected no difference in mean squared error (MSE) between participants randomly assigned to the breathing condition and those who proceeded directly to the task.

## 2 Study Design

### 2.1 Intervention and Outcomes

Participants were randomly assigned to one of two conditions. The **treatment** group completed a paced breathing protocol—a 4-second inhale, 2-second hold, and 6-second exhale, repeated for 60 seconds—immediately before performing a cursor line-drawing task. The **control** group proceeded directly to the drawing task without any pre-task intervention. The potential outcomes we compared were the mean squared deviation of drawn cursor trajectories from an ideal straight line (MSE), averaged across six rounds of the task, for the same participant under treatment versus under control.

## 2.2 Randomization

Randomization was implemented at the point of page load via a custom web application (hironai.to/straight). Each visitor was assigned automatically to treatment or control with equal probability, eliminating any opportunity for self-selection. Among the 224 participants who completed the study, 109 were assigned to control and 115 to treatment; a near-even split consistent with the intended 50/50 allocation and supporting unbiased randomization.

We verified covariate balance by comparing the distributions of gender, age group, mouse type, handedness, and self-reported tiredness between conditions. The two groups were broadly similar across all dimensions; the largest observed difference was a slightly higher proportion of female participants in the treatment group (approximately 29% vs. 22% in control). We include all covariates in our primary model to account for any residual imbalance.

## 2.3 Outcome Measures

The outcome was the mean squared error (MSE) between each participant's drawn cursor path and the ideal straight line connecting two endpoints, computed deterministically from logged cursor coordinates with timestamps. Participants completed six rounds of the task; we averaged MSE across rounds to reduce within-person noise. Secondary measures included mean absolute error (MAE) and task completion time, which serve as robustness checks.

Baseline questionnaire covariates included gender, age group, mouse type, dominant hand, and self-reported tiredness. The primary outcome was entirely behavioral. We do not observe a direct measure of whether treated participants followed the breathing pace on screen, so the analysis should be interpreted as an intent-to-treat comparison based on assignment rather than verified compliance.

## 2.4 Participant Flow

Our main analysis uses one row per completed participant session with exactly one clean occurrence of rounds 1 through 6. We began with 248 completion-like records in the backend logs. We then removed 19 records with malformed round structures, such as repeated round numbers or an incorrect number of stored attempts. Finally, we removed 5 repeated saves of the same session by keeping the latest saved version of each session key. This leaves an analytic sample of 224 participants (109 control, 115 treatment).

## 2.5 Power Analysis

A simulation-based power analysis, conducted prior to data collection, estimated power by simulating 1,000 covariate-adjusted regressions (treatment + age + fatigue + device type) with HC1 robust standard errors per sample size. The covariates used in the power simulation are conceptually aligned with those in the final analysis (gender, age group, mouse type, hand, tiredness), though operationalized differently given that the simulation used synthetic data.

Table 1: Sample flow for the primary analysis. The main analysis keeps completion-like records with exactly one occurrence of rounds 1 through 6 and removes repeated saves of the same session.

| Stage | N |
|---|---|
| Backend log rows | 1201 |
| Completion-like records retained | 248 |
| Removed for malformed round structure | 19 |
| Removed as repeated saves of the same session | 5 |
| Final analytic sample | 224 |

Table 2: Descriptive statistics by experimental condition for the primary analytic sample. MSE is mean squared error averaged across six rounds, MAE is mean absolute error, and log(MSE) is the natural-log transformation used in the primary regressions.

| Group | N | Mean MSE | Median MSE | Median MAE | Median Accuracy | Mean log(MSE) |
|---|---|---|---|---|---|---|
| Control | 109 | 415.4 | 138.3 | 8.21 | 83.6 | 5.099 |
| Treatment | 115 | 534.6 | 119.1 | 7.31 | 85.4 | 4.800 |

Three effect size scenarios were considered: conservative (5% MSE improvement, $d = 0.17$), expected (8%, $d = 0.27$), and optimistic (12%, $d = 0.40$), all assuming a baseline MSE of 100 (SD = 30). Under the expected scenario, 80% power is achieved at approximately 225 participants per group (450 total); the optimistic scenario requires only approximately 100 per group, while the conservative scenario would require over 400 per group and was not feasibly achievable. We targeted the expected scenario and recommended recruiting 225 per group. Our achieved analytic sample of 224 participants falls below this target, and the study should be treated as potentially underpowered for detecting small effects.

# 3 Data

The primary analytic sample contains 224 participant sessions. Approximately 89% of analyzed records carry MTurk-linked identifiers, with the remainder coming from family, friends, and the MIDS Slack channel. Mean task accuracy was 81.7% in the control group and 82.8% in the treatment group. At the center of the distribution, treatment sessions look somewhat better: median MSE is 119.1 in treatment versus 138.3 in control, and median MAE is 7.31 versus 8.21. This is useful context because raw mean MSE is more sensitive to a small number of very poor sessions than the median-based summaries are.

The raw MSE outcome is substantially right-skewed (skewness $\approx 3.6$ in the full sample), driven by a small number of participants with very poor drawing performance. We address this by applying a natural log transformation to MSE as our primary outcome variable,

which produces an approximately normal distribution and allows regression coefficients to be interpreted as multiplicative changes in geometric mean MSE.

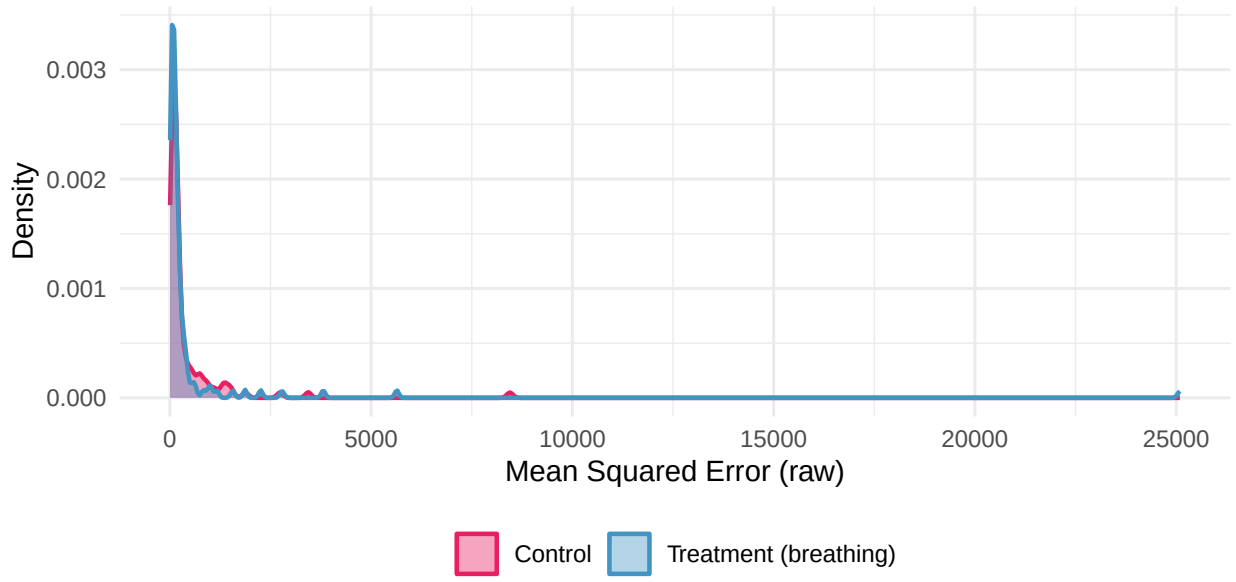


Figure 1: Distribution of raw mean squared error (MSE) by experimental condition in the primary analytic sample. Both groups show strong right skew, with a long tail of high-error observations. This motivates the log transformation used in the primary models.

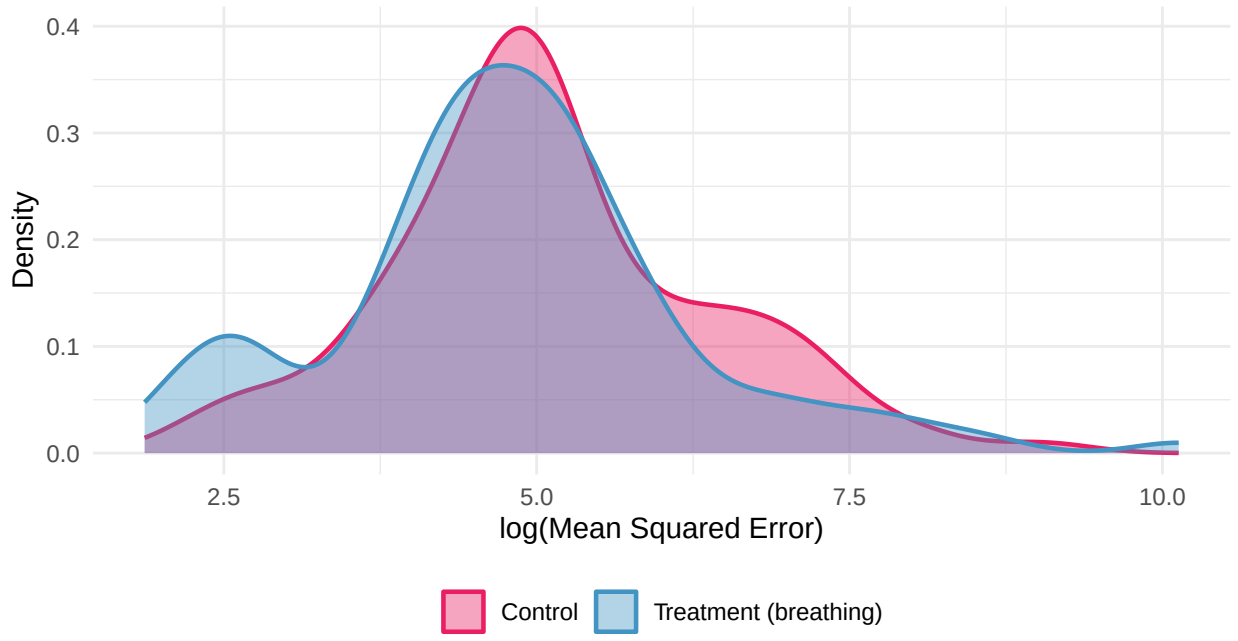


Figure 2: Distribution of log-transformed mean squared error (MSE) by experimental condition in the primary analytic sample. Treatment participants assigned to the 60-second guided breathing protocol are shown in teal; control participants who proceeded directly to the task are shown in pink. The distributions overlap substantially, with the treatment group showing a slight leftward shift indicating lower (better) error.

# 4 Analysis

## 4.1 Estimation Strategy

We estimate treatment effects using ordinary least squares regression with heteroscedasticity-robust standard errors (HC1). Models progress from a simple unadjusted treatment contrast to a fully adjusted specification including pre-registered covariates: gender, age group, mouse type, dominant hand, and self-reported tiredness. All models use log-transformed MSE as the primary outcome.

The log transformation is motivated by the raw outcome's distributional properties. MSE is strongly right-skewed in our sample (skewness $\approx 3.6$), which produces unstable inference on the raw scale and gives a disproportionate role to a small number of very poor sessions. Log-transforming reduces that skew and allows coefficients to be interpreted as approximate proportional changes in geometric mean MSE. Raw-MSE results are reported in the Appendix as a sensitivity check.

The unadjusted model is:

$$\log(\text{MSE}_i) = \beta_0 + \beta_1 \cdot \text{Treatment}_i + \varepsilon_i$$

Table 3: Regression estimates for the effect of the breathing intervention on log(MSE). The outcome is the natural log of average mean squared error across six drawing rounds. Model (1) is an unadjusted treatment-control contrast. Model (2) adds gender, age group, mouse type, dominant hand, and tiredness as precision covariates. Standard errors are heteroscedasticity-robust (HC1). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

| Model | Coefficient | SE | 95% CI | p-value | |
|---|---|---|---|---|---|
| (1) Unadjusted | -0.299 | 0.181 | [-0.653, 0.055] | 0.100 | * |
| (2) With covariates | -0.330 | 0.173 | [-0.668, 0.009] | 0.058 | * |

The primary adjusted model adds covariates to improve precision:

$$\log(\text{MSE}_i) = \beta_0 + \beta_1 \cdot \text{Treatment}_i + \mathbf{X}_i'\gamma + \varepsilon_i$$

where $\mathbf{X}_i$ includes gender, age group, mouse type, hand, and tiredness. Since treatment is randomly assigned, adding covariates improves precision without introducing bias.

## 4.2 Results

Table 3 presents estimates of the treatment effect under both model specifications.

In the unadjusted model, the treatment coefficient is -0.299 (SE = 0.181, p = 0.1). In the primary adjusted model, the treatment effect is -0.33 (SE = 0.173, 95% CI [-0.668, 0.009], p = 0.058). Back-transformed to the original scale, the adjusted estimate corresponds to about a 28.1% lower geometric mean MSE in treatment than in control. We view this pattern as suggestive rather than definitive: the point estimates are consistently favorable to treatment, but the interval is still fairly wide.

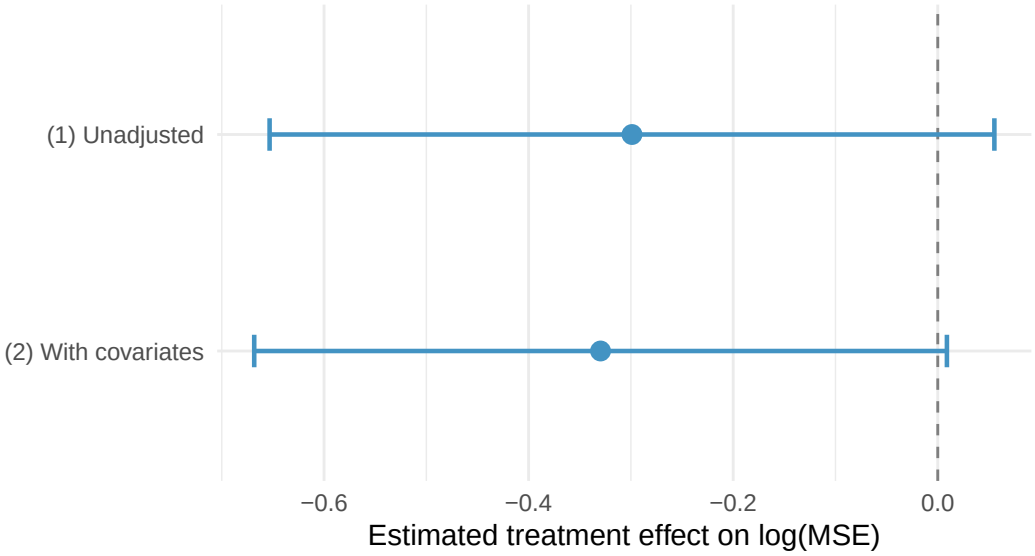


Figure 3: Point estimates and 95% confidence intervals for the treatment effect on log(MSE) under two model specifications. Negative values indicate lower (better) error in the treatment group. The dashed vertical line denotes the null hypothesis of no effect.

The addition of covariates modestly sharpens the estimate, consistent with the covariates reducing residual variance. Figure 3 shows that both point estimates are negative, with the adjusted specification somewhat more precise than the unadjusted one. A non-parametric bootstrap (B = 2,000 resamples) on the adjusted model yields a 95% confidence interval of [-0.664, 0.029], which is directionally consistent with the main regression.

## 4.3 Sensitivity Analyses

We conducted three sensitivity checks. First, using raw (untransformed) MSE as the outcome yields a treatment coefficient of -17.7 ($p = 0.921$), which is not statistically significant. This is consistent with the descriptive evidence that the raw scale is heavily influenced by a small number of poor sessions.

Second, we examined whether including gender affects the estimate. Because treatment was randomized, gender functions here as a pre-treatment precision covariate rather than as a remedy for omitted-variable bias in the usual observational sense. Excluding gender from the model yields a treatment coefficient of -0.334 ($p = 0.052$), which is slightly stronger than the primary estimate. The sign and general magnitude of the effect therefore do not depend on this modeling choice.

Third, we trimmed 3 high-MSE observations using the same z-score rule that appeared in earlier drafts. In that trimmed sample, the treatment coefficient is -0.343 ($p = 0.036$). This makes the estimate somewhat more favorable to treatment, but because the trimming rule is discretionary, we report the full strict sample as the main analysis.

# 5 Discussion

A 60-second guided breathing exercise was associated with lower fine-motor cursor error relative to the no-intervention control. In our primary model, the treatment group showed about a 28.1% lower geometric mean MSE ($p = 0.058$). We interpret this as encouraging pilot evidence rather than a settled result. The estimated effect is smaller than those reported in studies using longer interventions, which is in line with what we would expect from a one-minute protocol.

That said, several caveats apply. We fell short of our recruitment target of 450 participants, so the result warrants caution. The sample skews heavily toward MTurk workers (approximately 89%), who may differ from the general population in both baseline motor precision and responsiveness to a brief pre-task intervention. We also cannot verify compliance directly from the logged data. The treatment effect should therefore be read as an intent-to-treat estimate of assignment to the breathing prompt, not as a per-protocol estimate of actually following the breathing pace.

The result also depends on how the outcome is summarized. On the raw MSE scale the treatment effect is not statistically significant, while the log-transformed outcome and the median-based descriptive statistics are more favorable to treatment. This pattern suggests that the treatment group may perform better at the center of the distribution even though a few very poor sessions still affect the raw mean. For that reason, we think log(MSE) is the better headline outcome, but the raw-scale null result should still be reported alongside it.

Future work could explore whether the effect holds under stricter compliance monitoring, whether it generalizes beyond cursor drawing, and whether it varies by individual differences in baseline arousal or breathing experience.

# 6 References


Buchanan, T. L., & Janelle, C. M. (2022). Emotions and ensuing motor performance are altered by regulating breathing frequency: Implications for emotion regulation and sport performance. *Frontiers in Psychology*, *13*, 963711. doi:10.3389/fpsyg.2022.963711

Laborde, S., Mosley, E., & Thayer, J. F. (2022). Effects of voluntary slow breathing on heart rate and heart rate variability: A systematic review and a meta-analysis. *Neuroscience & Biobehavioral Reviews*, *138*, 104711. doi:10.1016/j.neubiorev.2022.104711

Larson, M. J., Steffen, P. R., & Primosch, M. (2013). The impact of a brief mindfulness meditation intervention on cognitive control and error-related performance monitoring. *Frontiers in Human Neuroscience*, *7*, 308. doi:10.3389/fnhum.2013.00308

Magnon, V., Dutheil, F., & Vallet, G. T. (2021). Benefits from one session of deep and slow breathing on vagal tone and anxiety in young and older adults. *Scientific Reports*, *11*, 19267. doi:10.1038/s41598-021-98736-9

Tang, Y. Y., Ma, Y., Wang, J., Fan, Y., Feng, S., Lu, Q., Yu, Q., Sui, D., Rothbart, M. K., Fan, M., & Posner, M. I. (2007). Short-term meditation training improves attention and self-regulation. *Proceedings of the National Academy of Sciences*, *104*(43), 17152–17156. doi:10.1073/pnas.0707678104

Yadav, G., & Mutha, P. K. (2016). Deep breathing practice facilitates retention of newly learned motor skills. *Scientific Reports*, *6*, 37069. doi:10.1038/srep37069

Zaccaro, A., Piarulli, A., Laurino, M., Garbella, E., Menicucci, D., Neri, B., & Gemignani, A. (2018). How breath-control can change your life: A systematic review on psycho-physiological correlates of slow breathing. *Frontiers in Human Neuroscience*, *12*, 353.

# 7 Appendix

## 7.1 Appendix A: Regression Models on Raw MSE

Table 4 shows the full covariate-adjusted regression on raw (untransformed) MSE.

Table 4: OLS regression of raw (untransformed) MSE on treatment and covariates, with HC1 robust standard errors. The treatment coefficient is not statistically significant on this scale, reflecting the influence of the right-skewed outcome distribution.

| Term | Estimate | SE | p-value |
|---|---|---|---|
| Intercept | -532.19 | 918.49 | 0.563 |
| Treatment | -17.66 | 177.60 | 0.921 |
| Gender: Male | -274.52 | 333.01 | 0.411 |
| Gender: Prefer not to say | -2876.56 | 2652.75 | 0.279 |
| Age: 25-34 | 944.41 | 837.48 | 0.261 |
| Age: 35-44 | 473.19 | 651.19 | 0.468 |
| Age: 45-54 | 629.84 | 830.53 | 0.449 |
| Age: 55-64 | 594.78 | 822.10 | 0.470 |
| Age: 65+ | -536.24 | 1052.47 | 0.611 |
| Age: under 17 | 589.76 | 935.94 | 0.529 |
| Mouse: Other | -43.41 | 460.23 | 0.925 |
| Mouse: Standard | -202.64 | 323.68 | 0.532 |
| Mouse: Trackpad | 1956.84 | 2077.41 | 0.347 |
| Hand: Right | 269.42 | 392.25 | 0.493 |
| Tiredness | 97.04 | 96.86 | 0.318 |

## 7.2 Appendix B: Covariate Balance

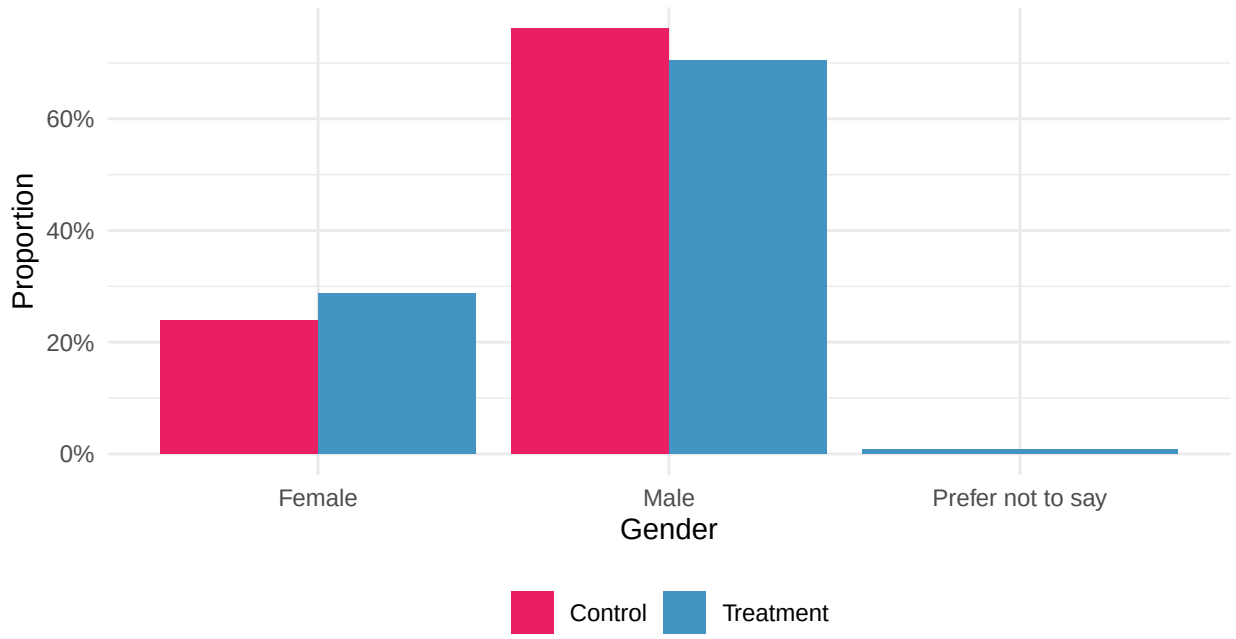


Figure 4: Gender distribution by experimental condition. The treatment group has a slightly higher proportion of female participants (29% vs. 22%), the largest observed imbalance across covariates.

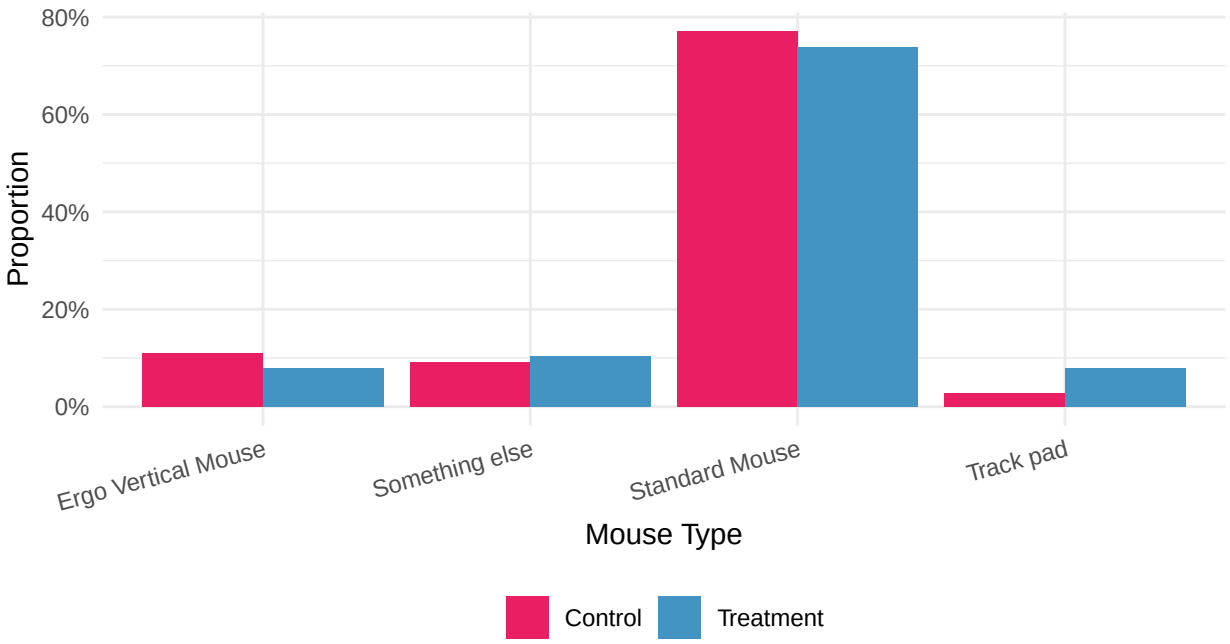


Figure 5: Mouse type distribution by experimental condition. Both groups are dominated by standard mouse users, with comparable proportions of trackpad users. The 'Something else' category is balanced across conditions.