

# mouse\_decision\_model

Bohan Chen

2025-02-09

## 1. Abstract

This project analyzed neural and behavioral data from 18 experimental sessions involving mice performing perceptual decision-making tasks. This study explores how differences in visual contrast, decision bias, and neuronal firing rate affect task performance. A comprehensive descriptive analysis reveals uneven contrast distribution, potential bias, and that the greater the contrast difference, the higher the success rate. Reasoning analysis emphasizes the importance of late occurrence rate in predicting decision outcomes. The machine learning model (XGBoost) integrates cross session data to improve prediction accuracy, demonstrating the importance of neural and behavioral features in decision-making. The robustness of the research results was confirmed by excluding fatigue effects and comparing alternative models.

## 2. Introduction

- 1. Questions of Interest:** The project aims to understand the decision-making behavior of mice in response to visual stimuli, specifically focusing on the impact of contrast levels on success rates and neural activity.
- 2. Impact of Results:** The findings could provide insights into neural mechanisms underlying decision-making and potential biases in animal behavior.
- 3. Real-World Motivation:** Understanding these mechanisms can contribute to neuroscience research and potentially inform models of human decision-making.
- 4. Problem Setup:** The problem involves analyzing neural data from mice performing a visual decision task, with data sourced from multiple sessions.
- 5. Key Variables:** Feedback type (success/failure), contrast levels (left and right), neural firing rates, and brain areas.
- 6. Hypotheses:** Mice may show decision biases based on contrast levels, and neural activity may correlate with success rates.

### 3. Background

- 1. Source of Data:** Neural data from mice performing a visual decision task.
- 2. Target Population:** Mice used in neuroscience experiments.
- 3. Sampling Mechanism:** Data collected across 18 sessions, each with multiple trials.
- 4. Explanation of Variables:** Feedback type, contrast levels, neural firing rates, and brain areas are explained.
- 5. Relevant Research:** The project builds on existing neuroscience research on decision-making and neural activity.

### 4 Exploratory analysis

# 4.1 Count the number of trials, neurons, and brain areas for each session, which let the results are stored in a data frame for easy viewing.

##	Session	Trials	Neurons	Brain_Areas
## 1	1	114	734	8
## 2	2	251	1070	5
## 3	3	228	619	11
## 4	4	249	1769	11
## 5	5	254	1077	10
## 6	6	290	1169	5
## 7	7	252	584	8
## 8	8	250	1157	15
## 9	9	372	788	12
## 10	10	447	1172	13
## 11	11	342	857	6
## 12	12	340	698	12
## 13	13	300	983	15
## 14	14	268	756	10
## 15	15	404	743	8
## 16	16	280	474	6
## 17	17	224	565	6
## 18	18	216	1090	10

## 1. Trial Counts Across Sessions

1.Early Sessions (1–10): There is a steady increase in trials until Session 10, reaching the highest count (447). This suggests that early experimental phases may have involved learning or adaptation.

2.Later Sessions (11–18): Trial counts drop after Session 10, fluctuating between 216 (Session 18) and 404 (Session 15). The drop in later sessions might indicate task modifications, experimental adjustments, or cognitive fatigue in subjects. Sessions 15 and 11 have relatively higher trial counts, indicating possible secondary peaks in engagement.

## 2. Neuron Counts Across Sessions

1.High Neuron Count Sessions: Session 4 (1769 neurons) recorded the highest number of neurons, possibly due to specific experimental conditions. Sessions 6, 10, and 18 also have high neuron counts, suggesting key experimental milestones.

2.Low Neuron Count Sessions: Session 16 has the lowest neuron count (474), followed by Sessions 7 and 17. A general decline in neuron count is observed after Session 10, potentially reflecting changes in task complexity, experimental setup, or adaptation.

## 3. Brain Areas Across Sessions

1.Most Diverse Brain Activity: Session 8 and 13 recorded the highest number of brain areas (15), providing a broader representation of neural activity. Sessions 9, 10, 12, and 14 also have a relatively high number of brain areas.

2.Fewer Brain Areas: Sessions 2, 6, and 16 recorded the lowest number of brain areas (5–6). This may indicate more specialized neural engagement in those sessions.

## Summarize the distribution of feedback\_type, contrast\_left, and contrast\_right across sessions

```
## [1] "Feedback Type Distribution:"
```

```
## -1 1 -1 1 -1 1 -1 1 -1 1 -1 1 -1 1 -1 1 -1 1 -1 1 -1 1
## 45 69 92 159 77 151 83 166 86 168 75 215 83 169 89 161 117 255 170 277
## -1 1 -1 1 -1 1 -1 1 -1 1 -1 1 -1 1 -1 1 -1 1 -1 1 -1 1
## 70 272 89 251 61 239 82 186 95 309 79 201 38 186 42 174
```

```
## [1] "Left Contrast Distribution:"
```

```
## 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1
## 51 27 18 18 133 25 39 54 137 31 28 32 112 41 46 50
## 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1
## 112 46 43 53 122 52 53 63 120 46 47 39 94 40 31 85
## 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1
## 153 45 68 106 193 60 74 120 159 50 62 71 173 55 55 57
## 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1
## 135 43 52 70 130 38 41 59 191 56 68 89 124 42 43 71
## 103 41 35 45 103 36 34 43
```

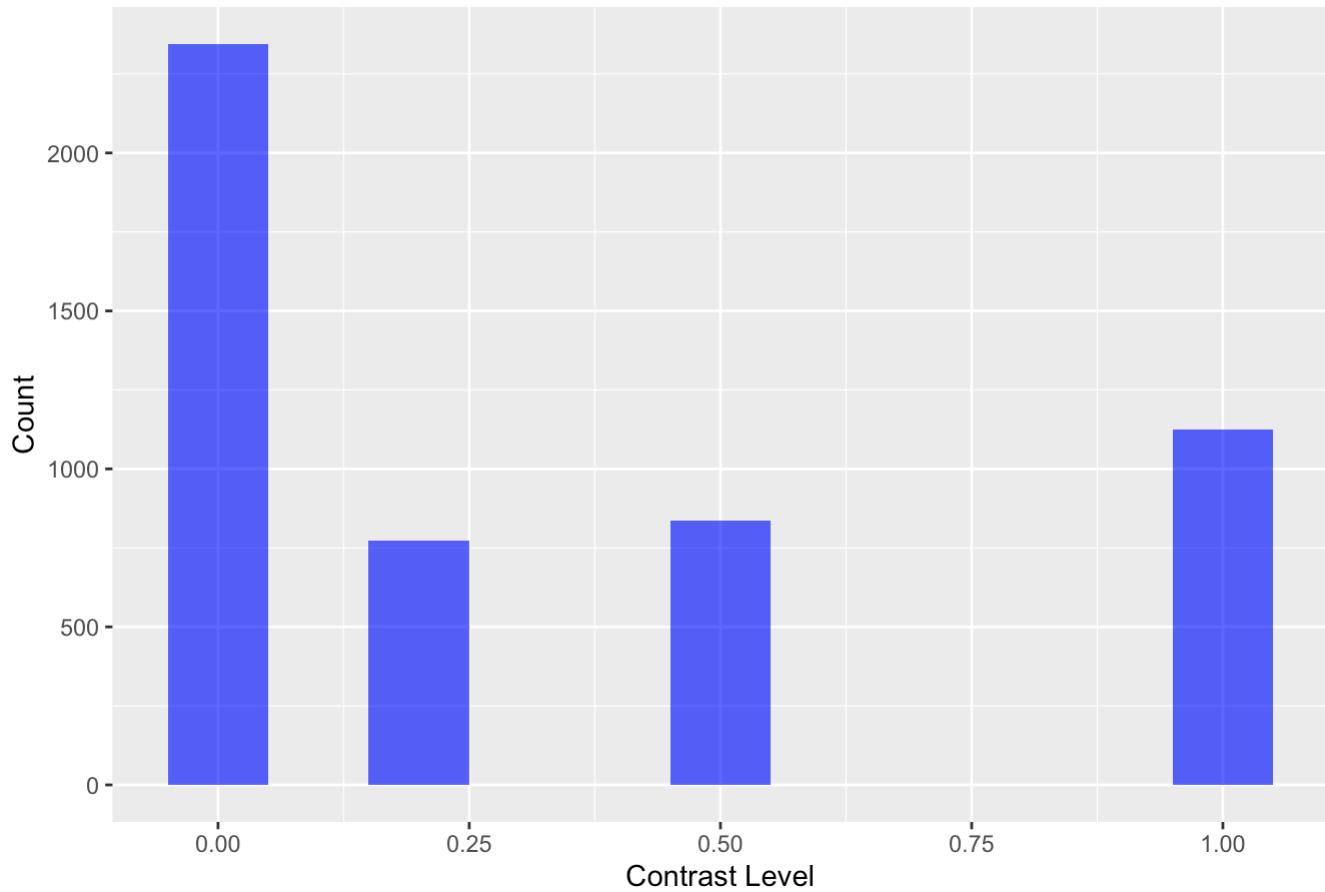
```
## [1] "Right Contrast Distribution:"
```

```
## 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1
## 43 14 23 34 115 41 34 61 109 26 31 62 107 55 41 46
## 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1
## 105 48 45 56 125 59 53 53 119 38 45 50 102 47 38 63
## 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1
## 193 74 48 57 245 73 53 76 173 47 53 69 167 50 50 73
## 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1 0 0.25 0.5 1
## 138 64 45 53 128 44 50 46 189 67 59 89 127 64 38 51
## 90 39 36 59 110 32 33 41
```

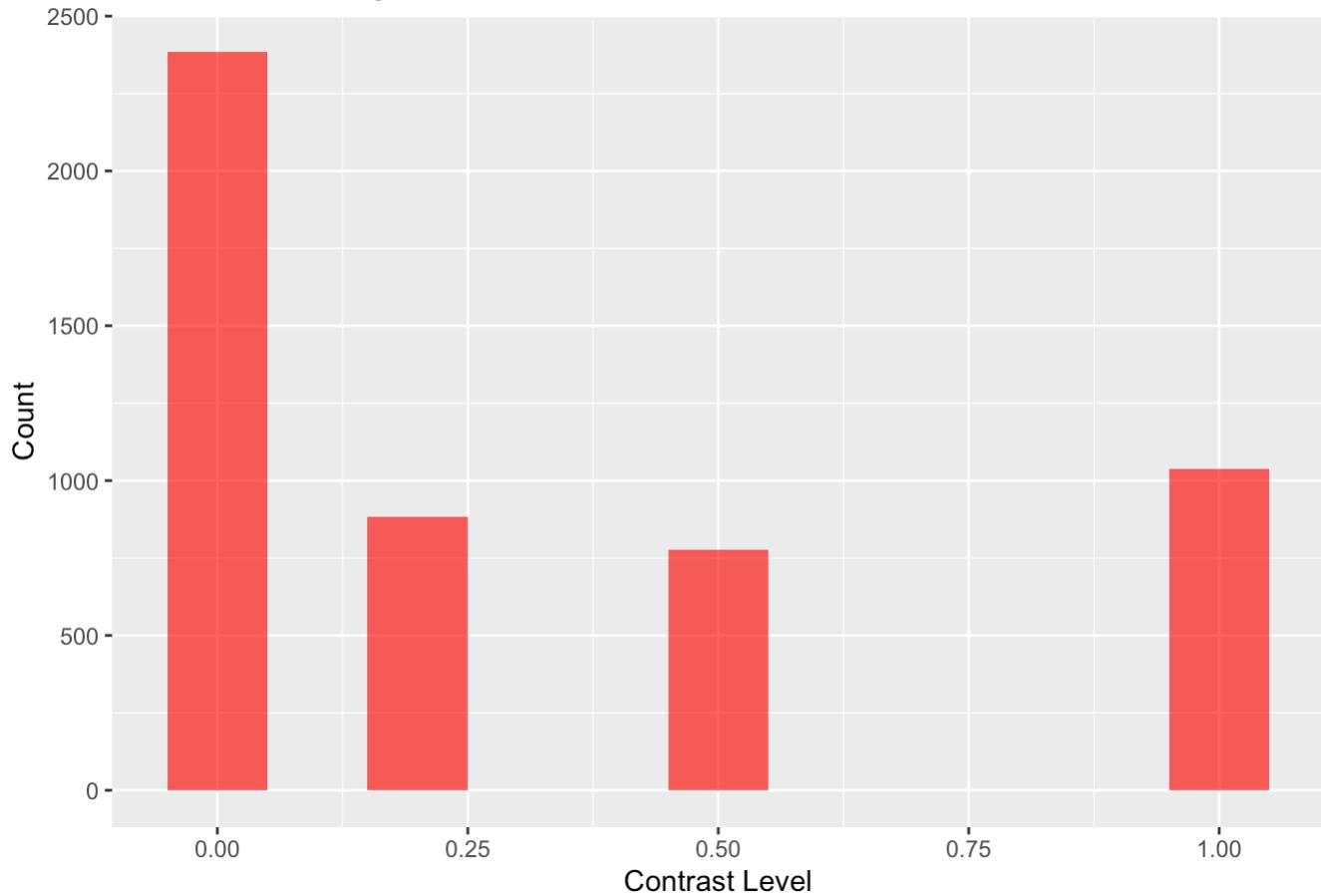
## 4.2 Visualize the distribution of contrast\_left and

# contrast\_right

Distribution of Left Contrast Levels



Distribution of Right Contrast Levels



## 1. Uneven Contrast Distributions

1. The majority of trials have contrast\_left = 0 and contrast\_right = 0, meaning that a large proportion of trials had no stimulus presented on either side. This suggests that a significant portion of trials may be “control” trials, where mice had to make a decision in the absence of visual cues.
2. There is a noticeable drop at contrast = 0.25, followed by a slight increase at contrast = 0.5 and contrast = 1.0, suggesting a non-uniform trial design. Left vs. Right Contrast Are Similar but Not Identical.
3. The overall shape of the left contrast and right contrast distributions appear similar. However, small differences in bar heights suggest that one side might have been slightly more biased toward certain contrast levels in some sessions. This could impact mouse decision-making behavior, as slight imbalances might introduce side preferences.

## 4.3 Success rate for each contrast level combination among different mice

```
##  
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':  
##  
##     filter, lag
```

```
## The following objects are masked from 'package:base':  
##  
##     intersect, setdiff, setequal, union
```

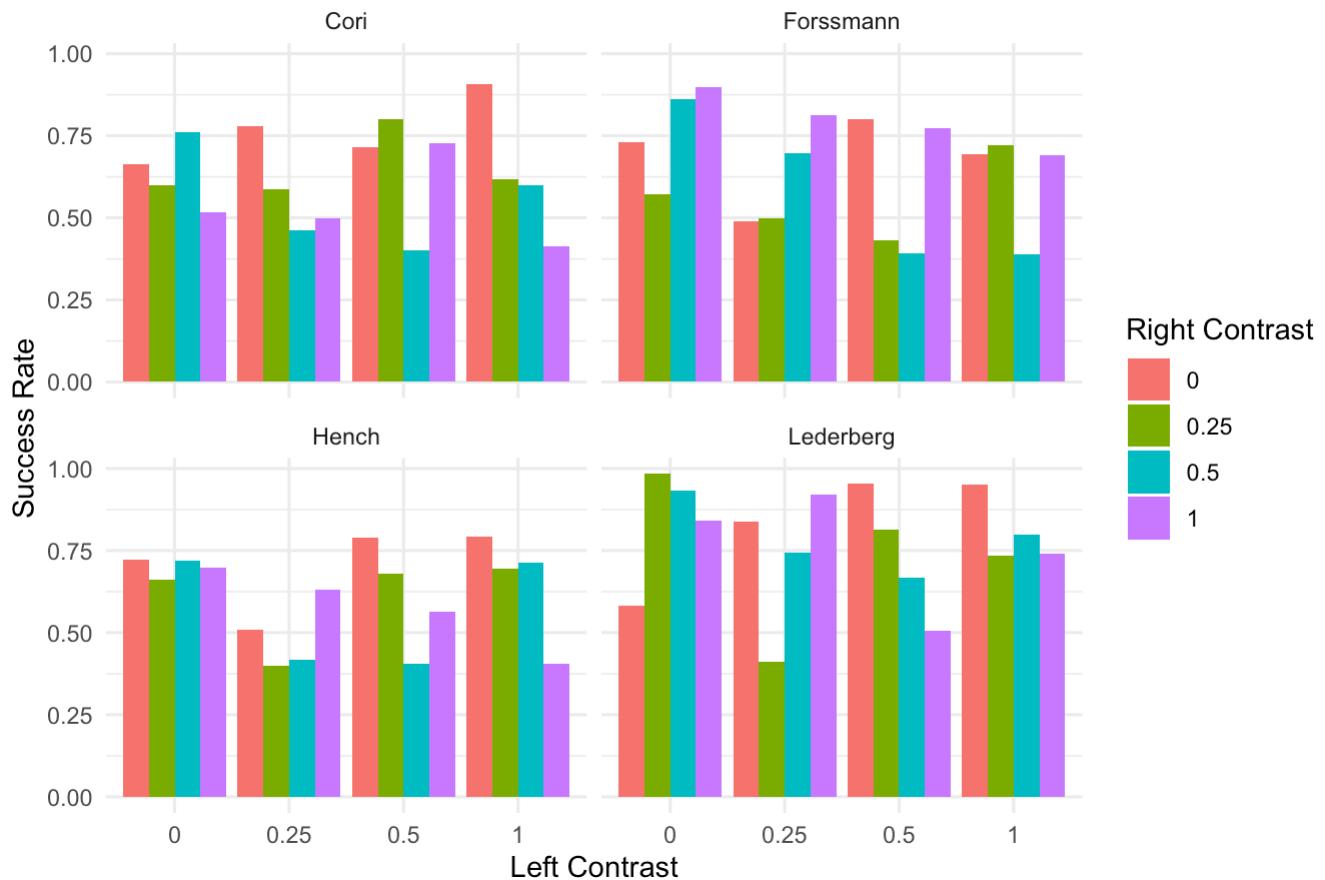
```

##   contrast_left contrast_right mouse_name success_rate
## 1          0.00          0.00    Cori    0.6626506
## 2          0.00          0.25    Cori    0.6000000
## 3          0.00          0.50    Cori    0.7600000
## 4          0.00          1.00    Cori    0.5176471
## 5          0.25          0.00    Cori    0.7777778
## 6          0.25          0.25    Cori    0.5882353
## 7          0.25          0.50    Cori    0.4615385
## 8          0.25          1.00    Cori    0.5000000
## 9          0.50          0.00    Cori    0.7142857
## 10         0.50          0.25    Cori    0.8000000
## 11         0.50          0.50    Cori    0.4000000
## 12         0.50          1.00    Cori    0.7272727
## 13         1.00          0.00    Cori    0.9069767
## 14         1.00          0.25    Cori    0.6176471
## 15         1.00          0.50    Cori    0.6000000
## 16         1.00          1.00    Cori    0.4117647
## 17         0.00          0.00  Forssmann 0.7296417
## 18         0.00          0.25  Forssmann 0.5714286
## 19         0.00          0.50  Forssmann 0.8627451
## 20         0.00          1.00  Forssmann 0.8983051
## 21         0.25          0.00  Forssmann 0.4893617
## 22         0.25          0.25  Forssmann 0.5000000
## 23         0.25          0.50  Forssmann 0.6964286
## 24         0.25          1.00  Forssmann 0.8125000
## 25         0.50          0.00  Forssmann 0.8000000
## 26         0.50          0.25  Forssmann 0.4310345
## 27         0.50          0.50  Forssmann 0.3928571
## 28         0.50          1.00  Forssmann 0.7735849
## 29         1.00          0.00  Forssmann 0.6923077
## 30         1.00          0.25  Forssmann 0.7200000
## 31         1.00          0.50  Forssmann 0.3877551
## 32         1.00          1.00  Forssmann 0.6896552
## 33         0.00          0.00    Hench    0.7211538
## 34         0.00          0.25    Hench    0.6615385
## 35         0.00          0.50    Hench    0.7191011
## 36         0.00          1.00    Hench    0.6992481
## 37         0.25          0.00    Hench    0.5081967
## 38         0.25          0.25    Hench    0.4000000
## 39         0.25          0.50    Hench    0.4166667
## 40         0.25          1.00    Hench    0.6323529
## 41         0.50          0.00    Hench    0.7902098
## 42         0.50          0.25    Hench    0.6785714
## 43         0.50          0.50    Hench    0.4062500
## 44         0.50          1.00    Hench    0.5625000
## 45         1.00          0.00    Hench    0.7918782
## 46         1.00          0.25    Hench    0.6949153
## 47         1.00          0.50    Hench    0.7142857
## 48         1.00          1.00    Hench    0.4062500
## 49         0.00          0.00  Lederberg 0.5836177
## 50         0.00          0.25  Lederberg 0.9833333
## 51         0.00          0.50  Lederberg 0.9338235
## 52         0.00          1.00  Lederberg 0.8418079
## 53         0.25          0.00  Lederberg 0.8387097
## 54         0.25          0.25  Lederberg 0.4117647

```

## 55	0.25	0.50	Lederberg	0.7432432
## 56	0.25	1.00	Lederberg	0.9219858
## 57	0.50	0.00	Lederberg	0.9548387
## 58	0.50	0.25	Lederberg	0.8142857
## 59	0.50	0.50	Lederberg	0.6666667
## 60	0.50	1.00	Lederberg	0.5074627
## 61	1.00	0.00	Lederberg	0.9520548
## 62	1.00	0.25	Lederberg	0.7346939
## 63	1.00	0.50	Lederberg	0.8000000
## 64	1.00	1.00	Lederberg	0.7407407

### Success Rates Across Contrast Levels by Mouse



## 1. Contrast Difference Matters

1. Success rates tend to be higher when the contrast difference between the left and right sides is large (e.g., left = 1, right = 0).
2. Success rates drop when the contrast is more balanced (e.g., left = 0.25, right = 0.25).
3. This aligns with the idea that mice find it easier to make decisions when the stimulus difference is strong.

## 2. Individual Mouse Variability

1. Cori & Hench show relatively consistent performance across conditions but have noticeable drops at certain contrast levels.
2. Forssmann has a significant success rate boost when right contrast is 1, suggesting a potential right-side bias.
3. Lederberg has the highest success rates when the left contrast is 0 and the right contrast is 0.25 or 0.5, indicating that lower left contrast might be beneficial for this mouse.

## 3. Effect of Right Contrast

1. When right contrast is 0, success rates tend to be high, especially for Forssmann and Hench.

2. When right contrast is 1, success rates drop for some conditions, showing that mice might struggle with extreme right contrast.

### 3. Potential Biases

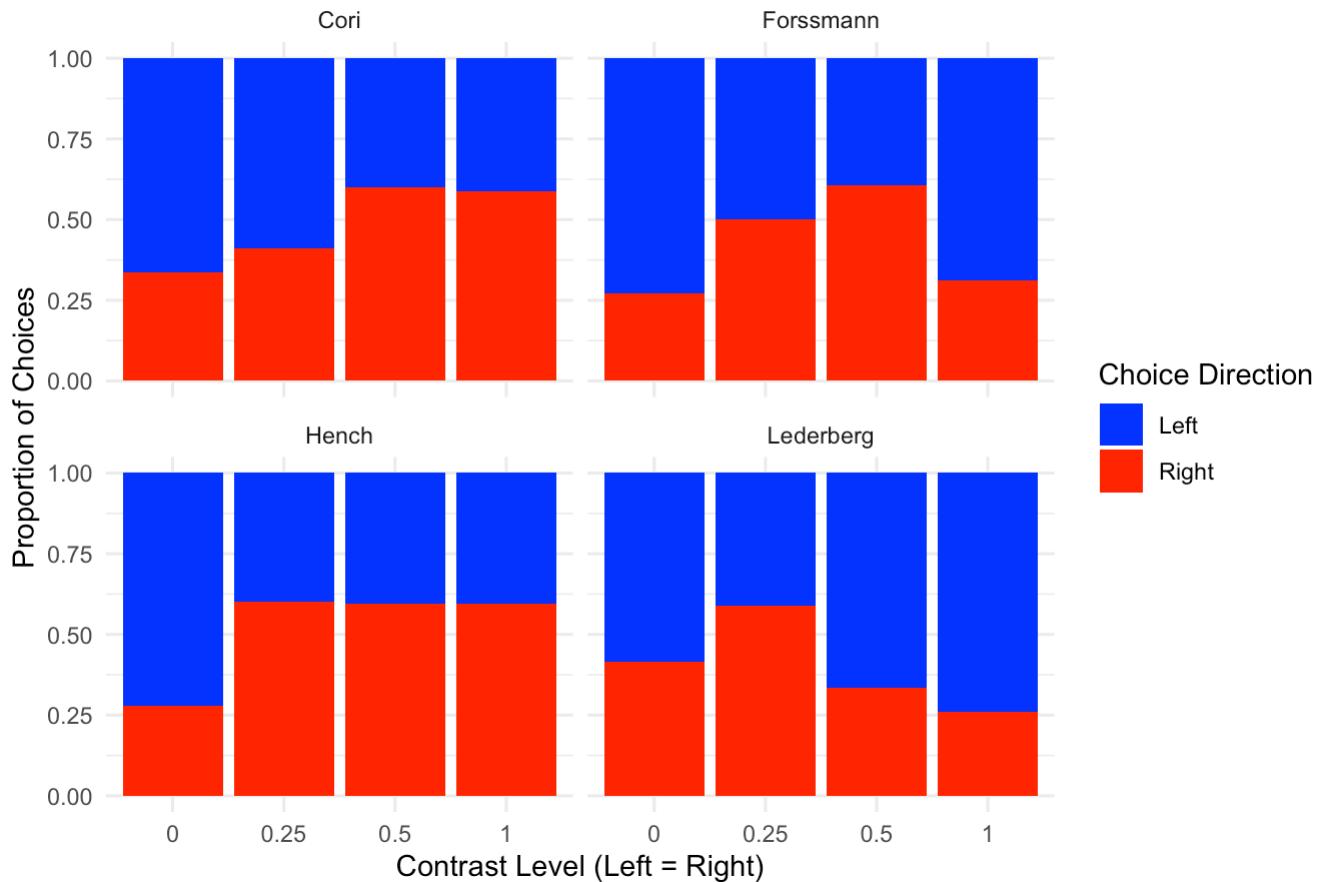
Some mice, such as Forssmann, seem to perform better when the right contrast is high, potentially indicating a side preference.

Lederberg shows a different pattern, where certain left contrast values (like 1) seem to boost performance regardless of the right contrast.

## 4.4 Test for decision biases in ambiguous contrast conditions (e.g., when left contrast $\approx$ right contrast), we will: Filter trials where left contrast $\approx$ right contrast (e.g., both 0, both 0.25, both 0.5, both 1).

```
## # A tibble: 32 × 7
##   mouse_name contrast_left contrast_right feedback_type choice_count
##   <chr>        <dbl>        <dbl>        <dbl>        <int>
## 1 Cori          0            0           -1          56
## 2 Cori          0            0            1          110
## 3 Cori         0.25         0.25         -1           7
## 4 Cori         0.25         0.25          1          10
## 5 Cori          0.5           0.5          -1           9
## 6 Cori          0.5           0.5           1           6
## 7 Cori          1             1           -1          10
## 8 Cori          1             1            1           7
## 9 Forssmann    0            0           -1          83
## 10 Forssmann   0            0            1          224
## # i 22 more rows
## # i 2 more variables: choice_direction <chr>, proportion <dbl>
```

## Decision Bias in Ambiguous Contrast Trials by Mouse



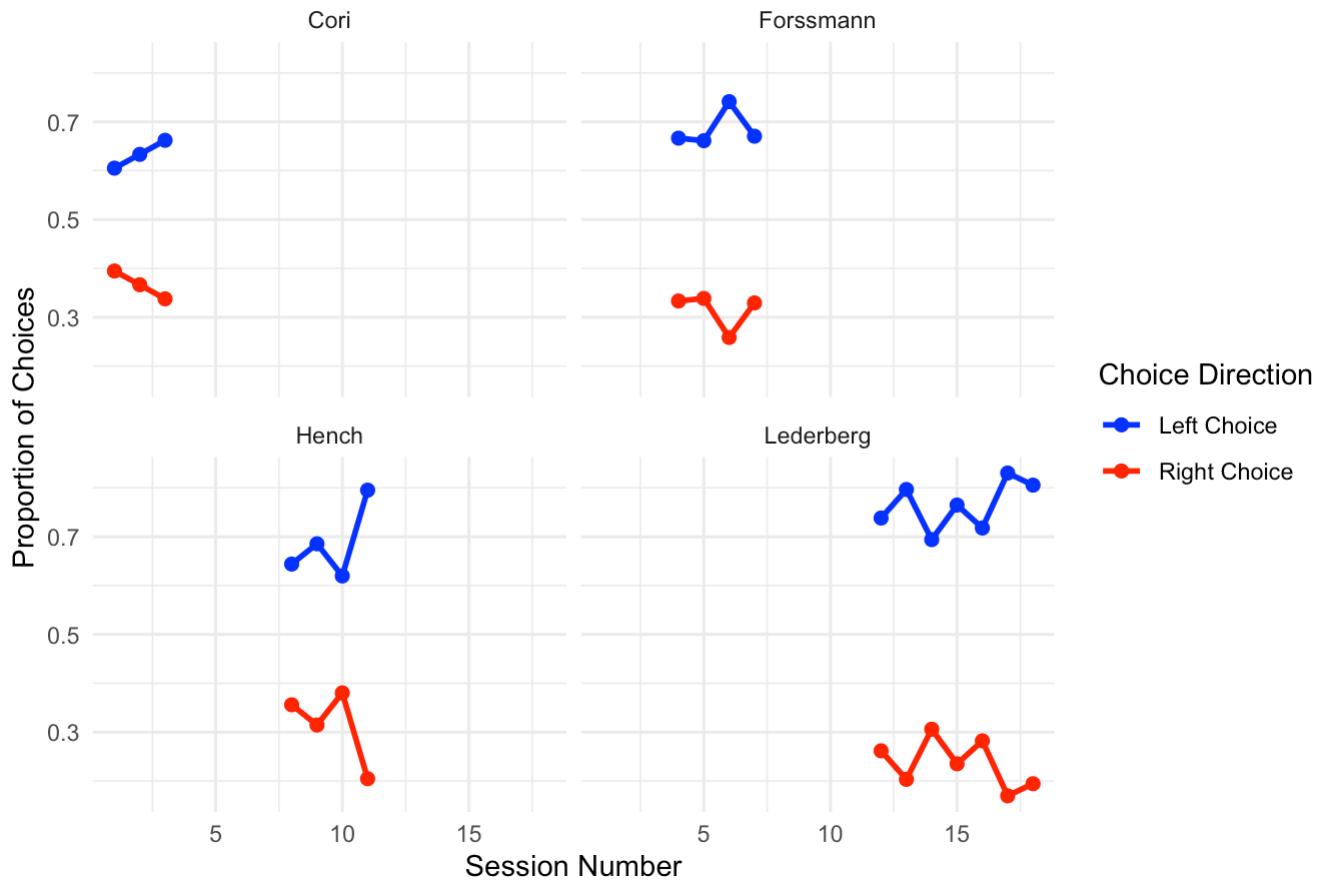
## 1. Mouse-Specific Biases:

1. Cori shows a relatively balanced response, but at some contrast levels, it slightly favors left choices.
2. Forssmann exhibits a clear right-side preference at lower contrast levels but becomes more balanced as contrast increases.
3. Hench follows a similar pattern to Forssmann, showing a strong right bias at lower contrast.
- Lederberg has the most noticeable right bias, especially at contrast levels of 0.25 and 0.5.

## 4.5 Based on 4.4, Explore whether the bias of each mouse will have a clear tendency as the session increases

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

## Decision Bias Trend Over Sessions



## 1. Findings from the Decision Bias Trend Over Sessions

In Hench and Lederberg, the proportion of left choices steadily increases over sessions, while right choices decrease. This suggests a stronger preference for left choice as they gain more experience.

## 2. Stable Bias in Some Mice:

1. Forssmann's bias remains relatively stable, indicating that their decision-making is consistent across sessions.
2. Cori shows a slight increase in left choices, but it's not as strong as Hench or Lederberg.

## 3. Fluctuating Bias in Some Cases:

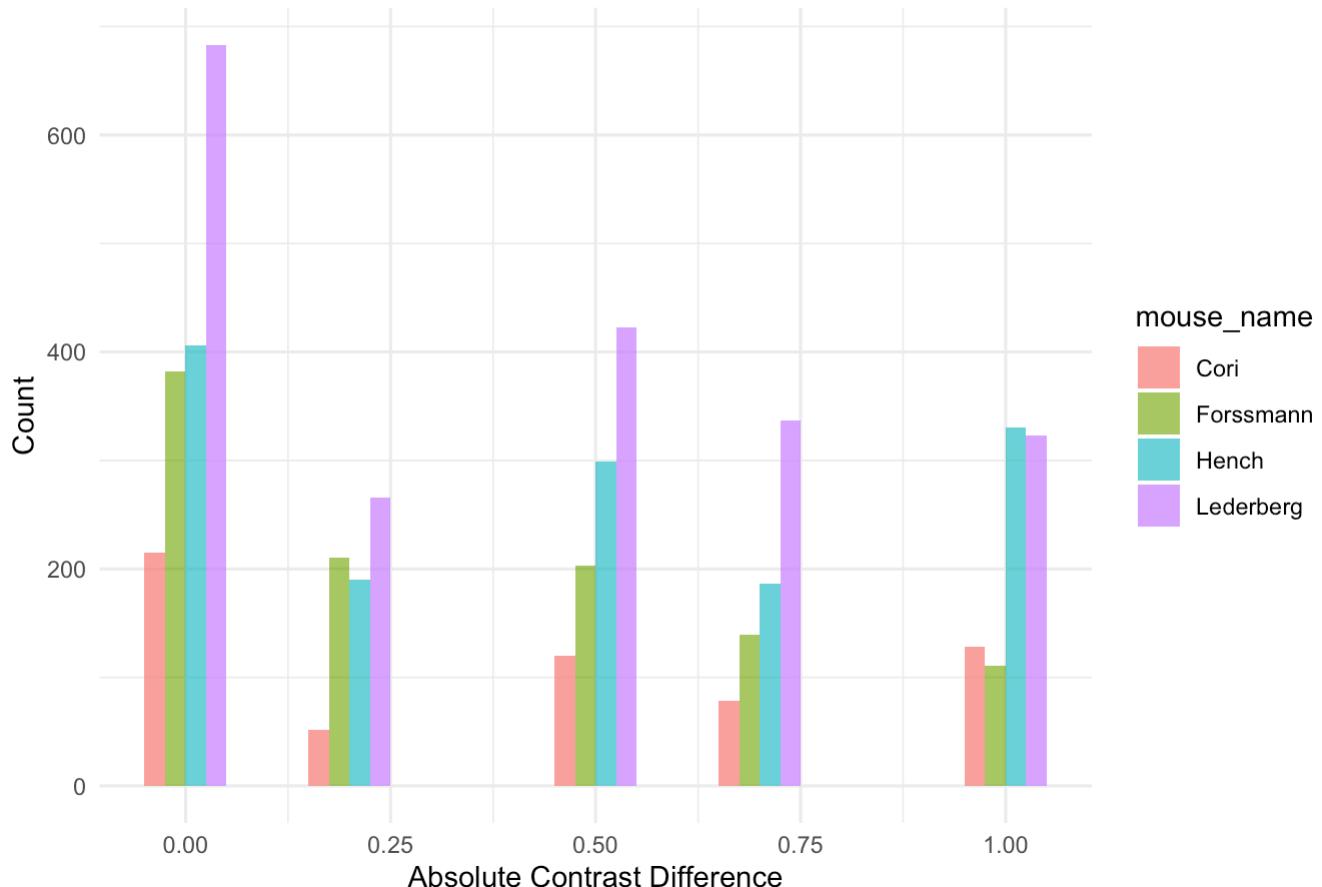
In Hench, there is some fluctuation in left choice preference, but the overall trend is increasing. This could indicate that decision-making is adaptive and may depend on specific conditions in each session.

## 4.6 Explore Distribution of Absolute Contrast

# Difference Across All Sessions

```
##   mouse_name abs_contrast_diff
## 1      Cori            0.5
## 2      Cori            0.0
## 3      Cori            0.5
## 4      Cori            0.0
## 5      Cori            0.0
## 6      Cori            0.0
```

Distribution of Absolute Contrast Difference Across All Sessions



## 1. High Frequency of Trials with Zero Contrast Difference

The leftmost bar (0.00 contrast difference) is the highest across all mice, particularly for Lederberg (purple) and Hench (blue). This indicates that many trials had equal contrast on both sides, meaning the mice had to make decisions without a strong visual cue. This could be part of a control condition where decision-making is tested in ambiguous conditions.

## 2. Uneven Distribution of Contrast Differences

Higher contrast differences (0.75 and 1.00) occur less frequently than lower contrast differences. This suggests that most trials were designed to test difficult decision-making cases (low contrast differences) rather than easy ones (high contrast differences).

The distribution is not uniform, meaning different contrast levels were used at different frequencies.

### 3. Mouse-Specific Differences in Trial Distribution

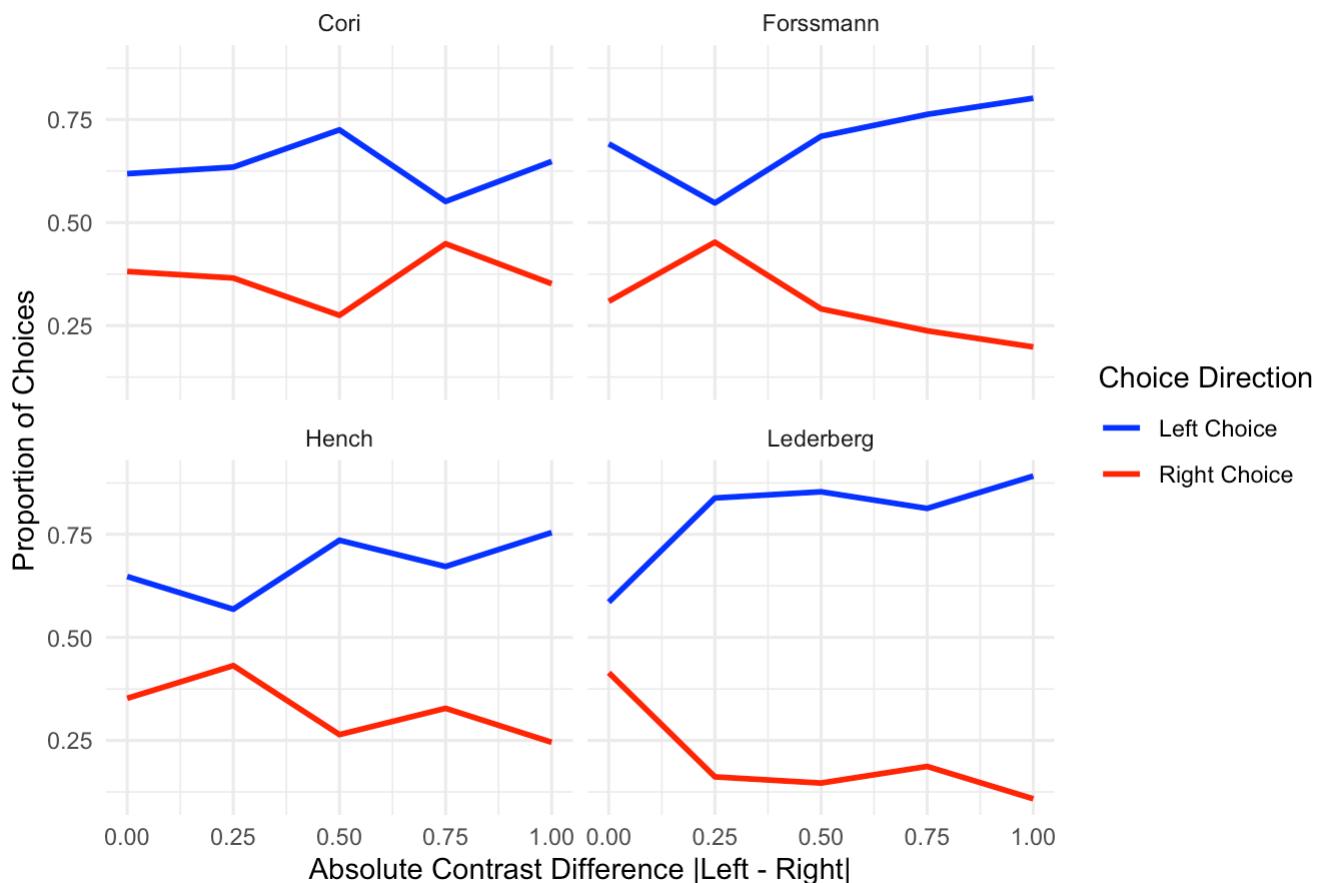
1. Lederberg (purple) participated in the most trials across all contrast differences, indicating this mouse had the highest trial count.
2. Forssmann (green) and Hench (blue) have relatively similar distributions, with more trials in low contrast differences than high ones.
3. Cori (red) participated in fewer trials overall, which may indicate differences in performance, experimental conditions, or task engagement.

### 4. Potential Bias in Trial Design

The noticeable drop at 0.25 contrast difference, followed by an increase at 0.50 and 0.75, suggests a non-uniform trial design.

#### 4.7 Explore Mouse Decision Bias Across Absolute Contrast Differences, To Find Any Rules Between Choice Decision And Absolute Contrast Differences

Mouse Decision Bias Across Absolute Contrast Differences



## 1. Bias at $\text{abs}(\text{contrast\_left} - \text{contrast\_right}) = 0$

Some mice do not choose 50–50 when the contrast is equal.

If a mouse prefers left more often at  $\text{abs\_contrast\_diff} = 0$ , it suggests a natural left-side bias independent of contrast.

Forssmann and Hench show a stronger left-choice tendency in ambiguous cases.

Increasing Absolute Contrast Difference

## 2. As $\text{abs\_contrast\_diff}$ increases:

Left choices increase (for most mice).

Right choices decrease, indicating that mice rely more on contrast to make decisions. The trend isn't perfectly linear, suggesting that individual biases interact with contrast strength.

## 3. Mouse-Specific Trends

Lederberg has a strong left bias at medium contrast differences but flattens out at high values.

Hench appears more random at mid-range contrast differences, meaning decisions may not be fully driven by contrast.

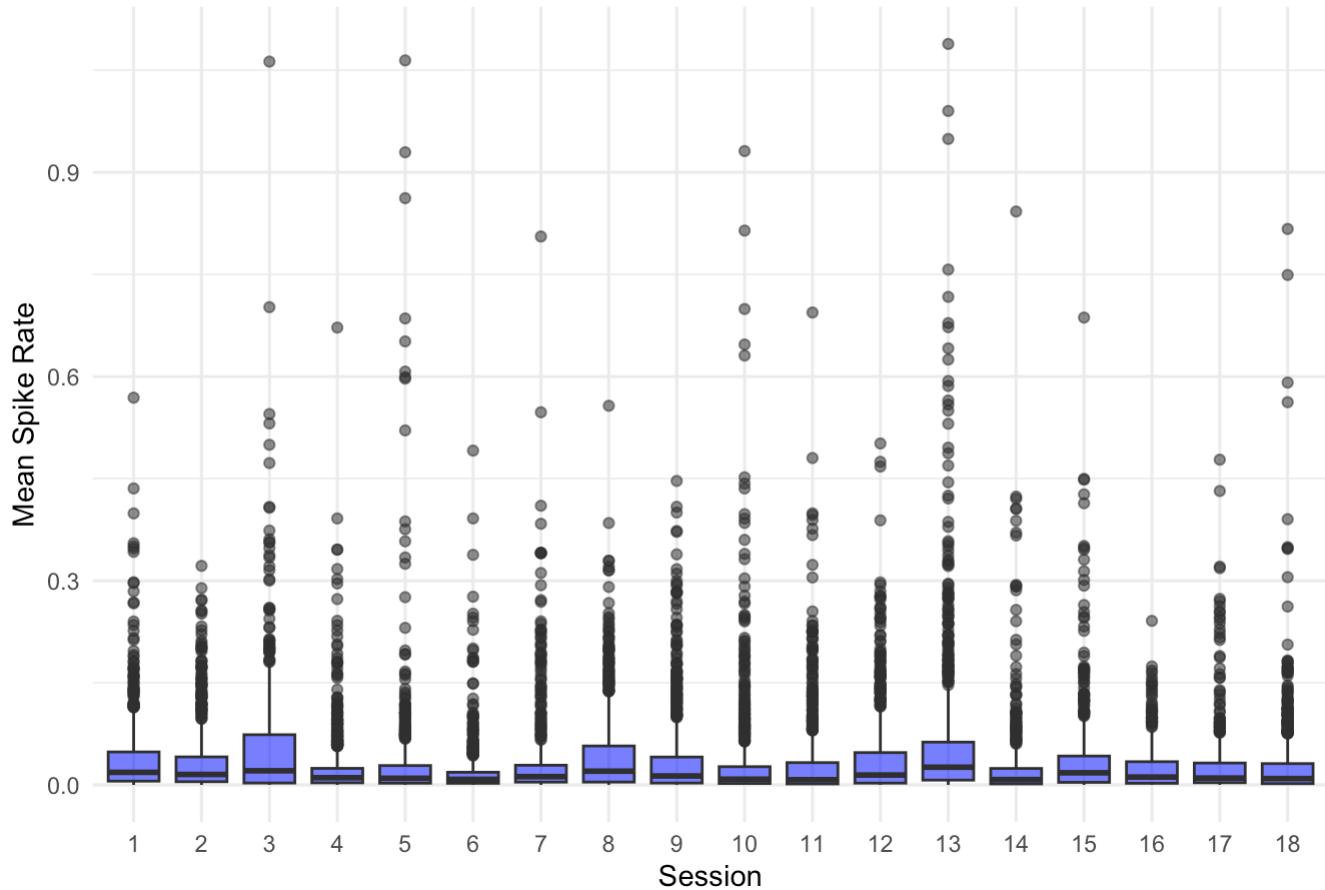
Forssmann's left-choice rate stabilizes over higher contrast differences, indicating clearer decision-making with stronger stimuli.

**All The Contrast Level Were Done And For Futher Investigation, Neural Activity Is Also An Important Part For Influence The Mouse Decision**

## 4.8 Neural Activities - Average Spike Rate per

# Neuron

## Average Spike Rate per Neuron Across Sessions



### 1. Low Mean Firing Rates Across Sessions:

The majority of neurons exhibit relatively low spike rates, with the median values clustered close to zero. This suggests that most neurons fire infrequently during trials.

### 2. Presence of Outliers:

Each session contains a subset of neurons with significantly higher firing rates (above 0.3 and in some cases exceeding 0.9). These high-firing neurons could represent task-relevant or stimulus-driven neurons, which respond selectively to specific trial conditions.

### 3. Session-to-Session Variability:

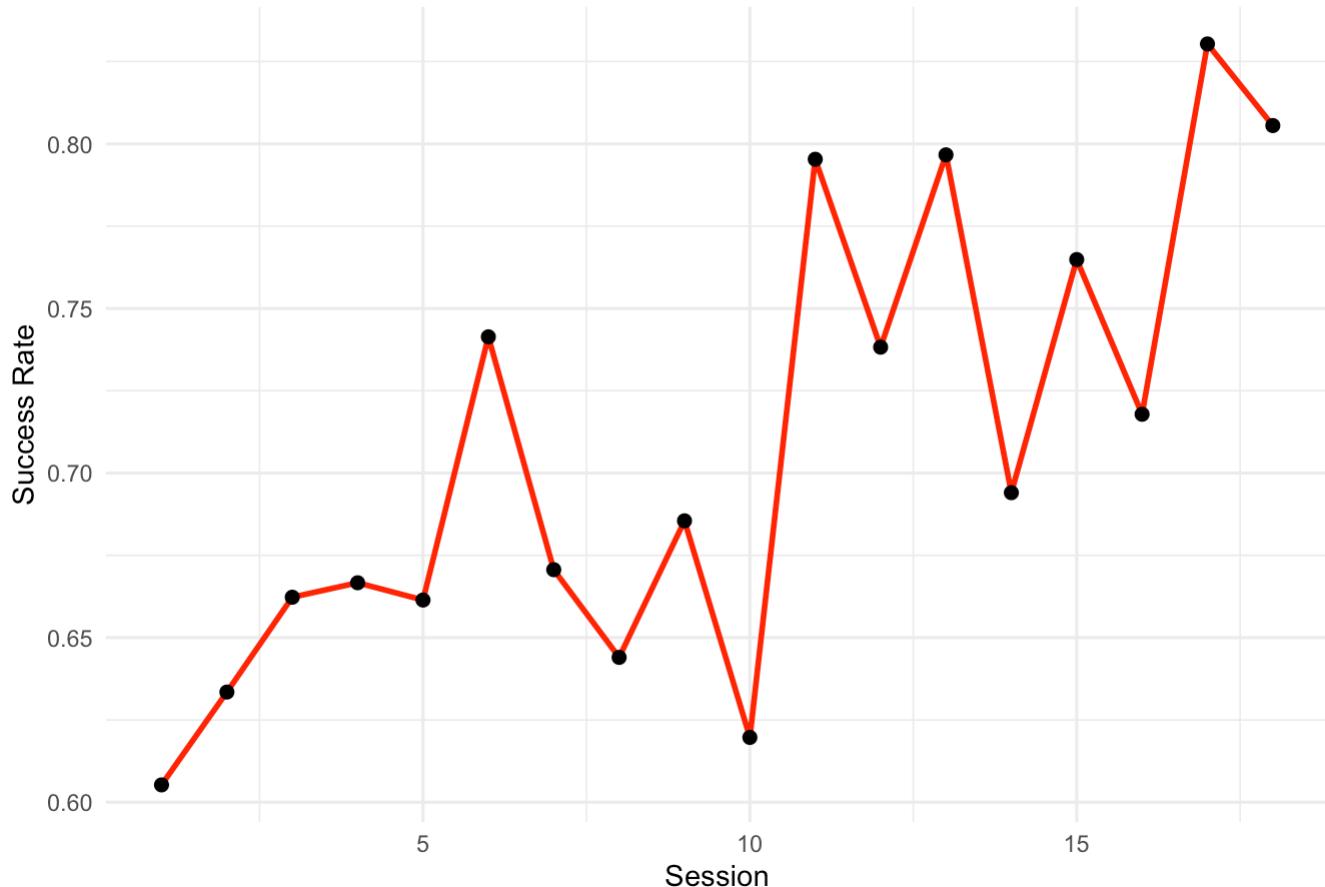
Some sessions (e.g., Session 3, Session 13) appear to have slightly higher median spike rates and more variance compared to other sessions. This could indicate variability in neural engagement or differences in task conditions across sessions.

### 4. Heterogeneity in Neural Activity:

The spread of firing rates varies across sessions, indicating that some sessions have more heterogeneous neuronal responses than others.

## 4.9 Changes Across Trials - Success Rate

Success Rate Across Trials



### 1. Overall Increasing Trend:

The success rate generally improves across sessions, indicating that the mice may be learning or adapting to the task over time. The increase becomes more pronounced after Session 10, where the success rate consistently rises above 0.75.

### 2. Fluctuations in Performance:

Although there is a general upward trend, success rates show noticeable fluctuations, with some sessions experiencing sharp drops. For example, around Session 10, there is a sharp dip, suggesting possible difficulty in that session's task conditions or a temporary performance decline.

### 3. Plateauing Effect Toward the End:

After Session 15, the success rate stabilizes at around 0.80, suggesting that the mice may have reached a learning plateau where further improvement is minimal. This could indicate that they have optimized their strategy and additional sessions do not significantly enhance their performance.

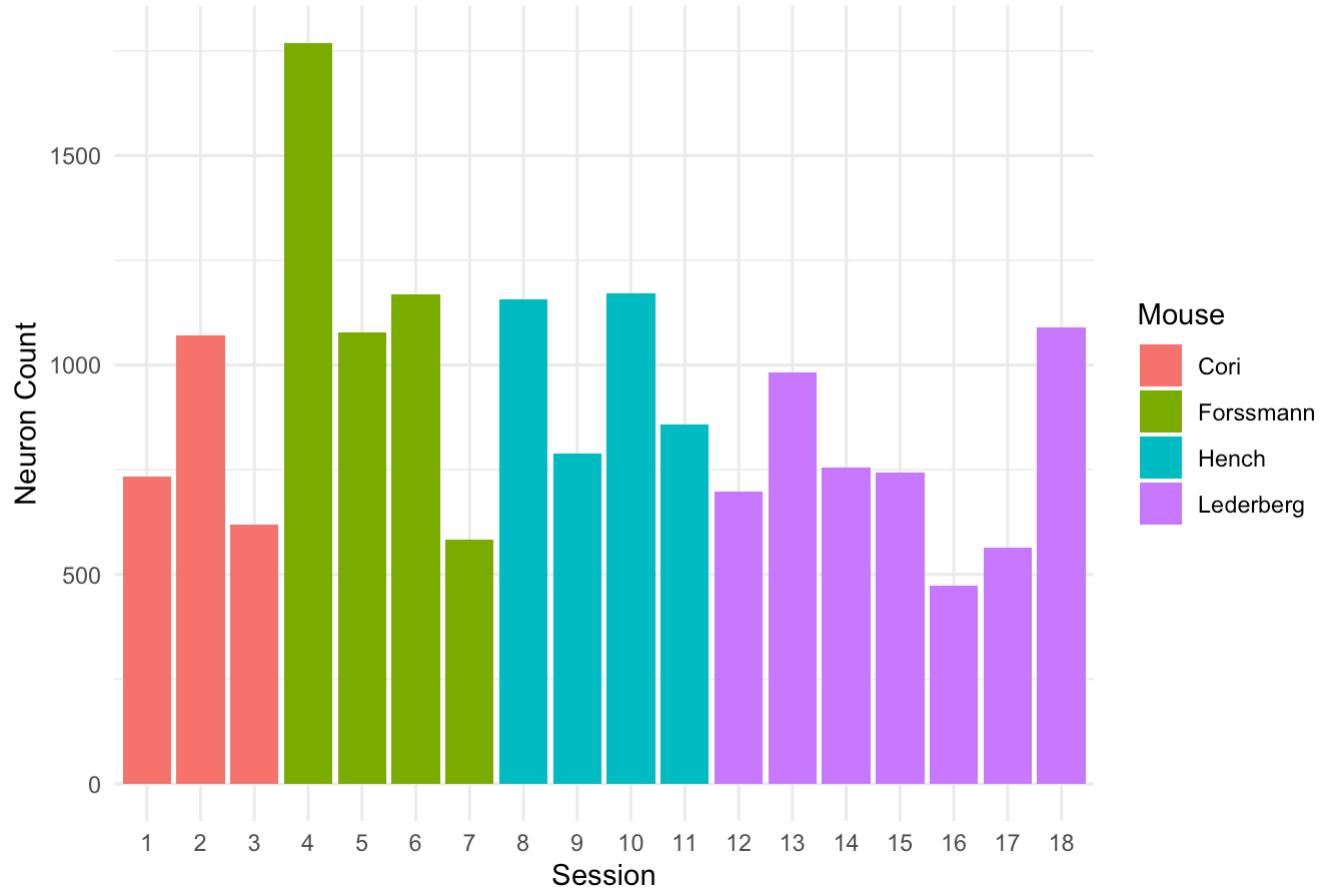
## 4. Potential Experiment Effects:

The sharp increase after Session 10 suggests that something may have changed in the experiment—possibly new training, reinforcement, or an adjustment in trial difficulty. Checking for external changes in task structure or reward mechanisms could help explain these shifts.

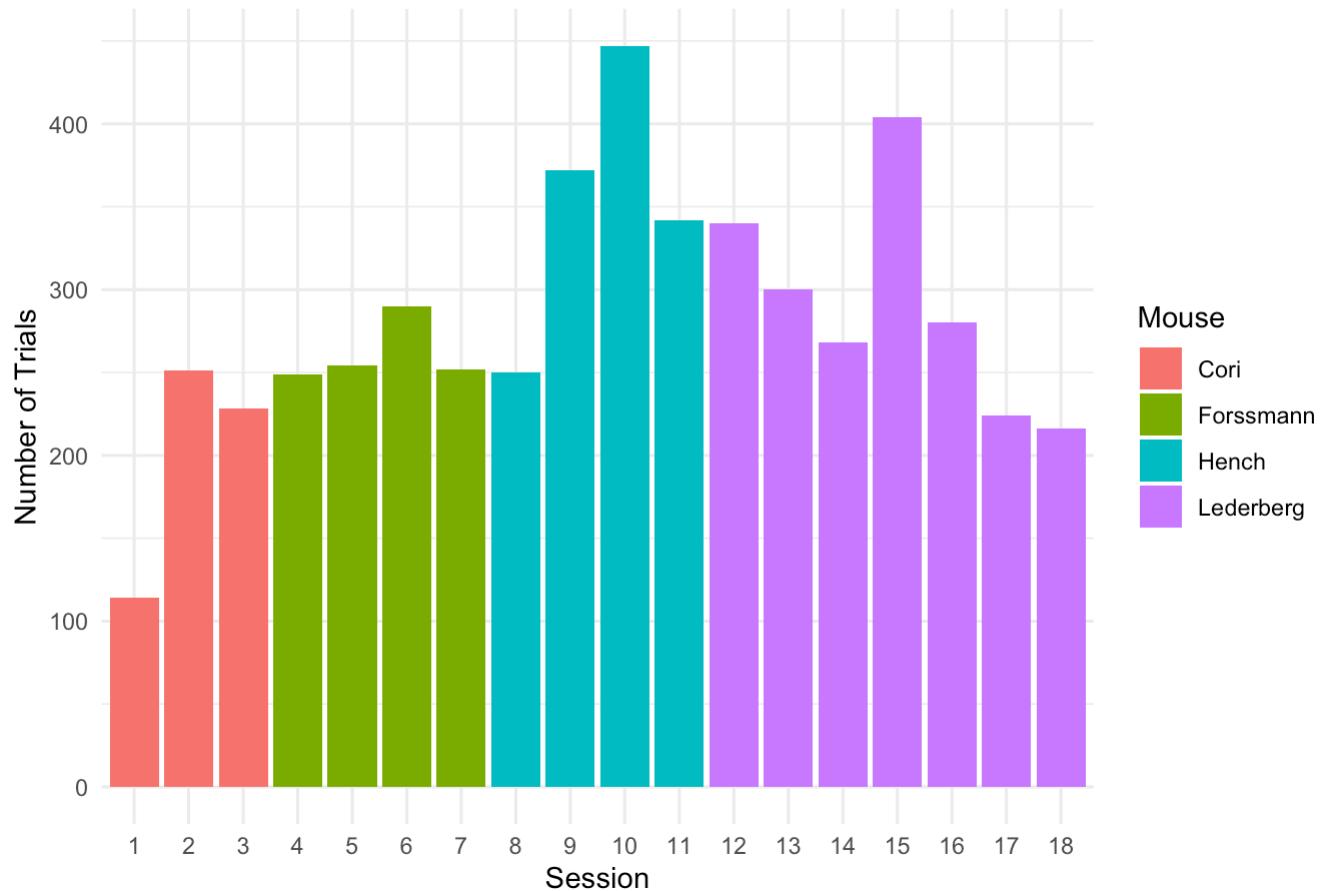
### 4.10 Homogeneity & Heterogeneity Across Sessions & Mice

	Mouse	Session	Neurons	Trials	Brain_Areas
## 1	Cori	1	734	114	8
## 2	Cori	2	1070	251	5
## 3	Cori	3	619	228	11
## 4	Forssmann	4	1769	249	11
## 5	Forssmann	5	1077	254	10
## 6	Forssmann	6	1169	290	5
## 7	Forssmann	7	584	252	8
## 8	Hench	8	1157	250	15
## 9	Hench	9	788	372	12
## 10	Hench	10	1172	447	13
## 11	Hench	11	857	342	6
## 12	Lederberg	12	698	340	12
## 13	Lederberg	13	983	300	15
## 14	Lederberg	14	756	268	10
## 15	Lederberg	15	743	404	8
## 16	Lederberg	16	474	280	6
## 17	Lederberg	17	565	224	6
## 18	Lederberg	18	1090	216	10

### Neuron Count Across Sessions



### Trial Count Across Sessions

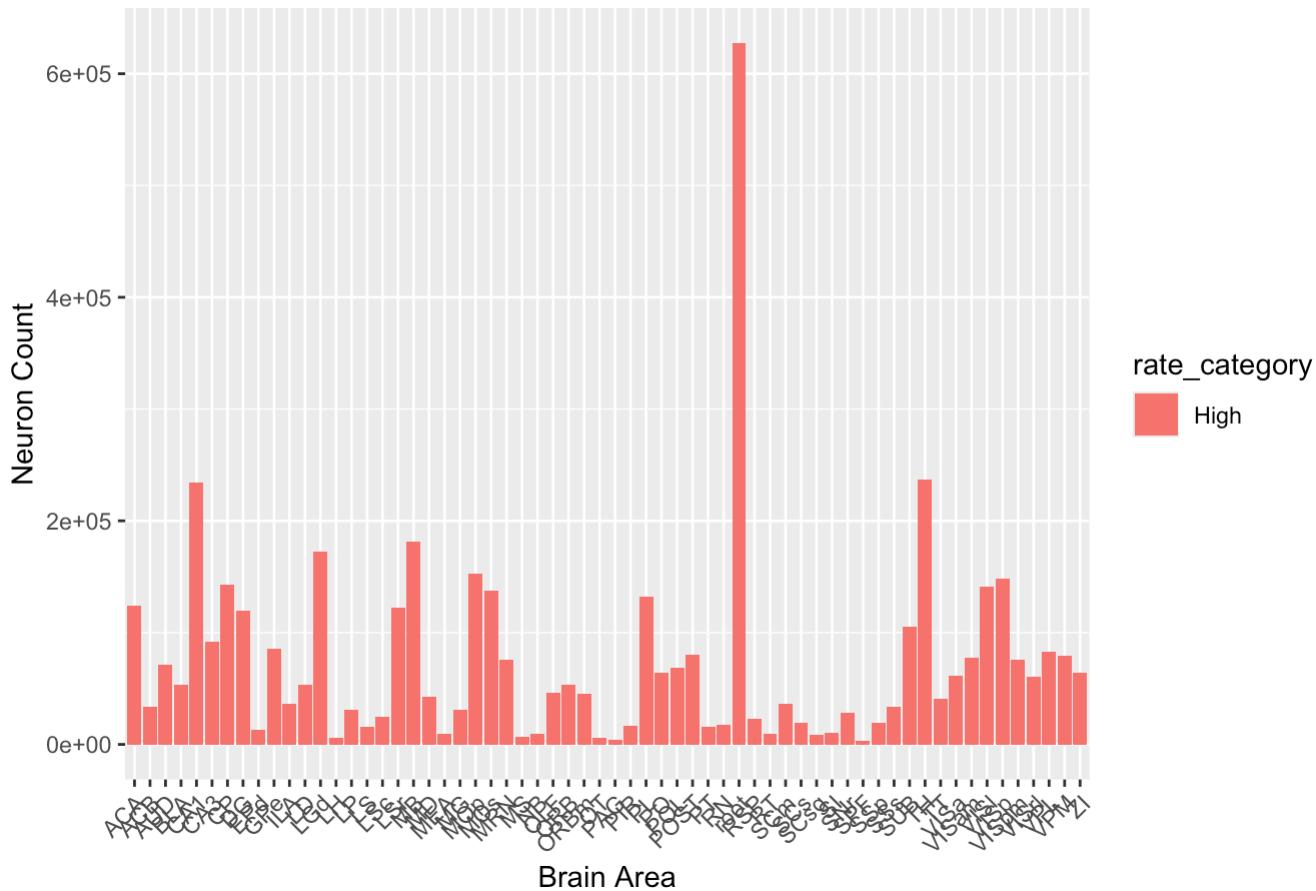


**1. From the plot we can see that there is a decline in the number of active neurons per session, it is maybe related to fatigue or adaptation in the mice.**

As the experiment progresses, the mouse may become physically or mentally tired. Fatigue could lead to reduced engagement with the task, which might affect neural activity or spike detection. If neurons involved in decision-making and motor control are less active due to fatigue, this could reduce the overall number of recorded spikes. The mouse could be adapting to the task, leading to more efficient neural processing. Over time, the brain may optimize its neural strategy, reducing the need for widespread activation. This would result in fewer neurons actively participating in the task but still maintaining similar performance.

## 4.11 Compare High vs. Low Firing Rate Neurons Across Brain Areas ### This identifies whether neurons with higher firing rates are concentrated in certain brain areas or randomly distributed.

Distribution of High vs. Low Firing Rate Neurons Across Brain Areas

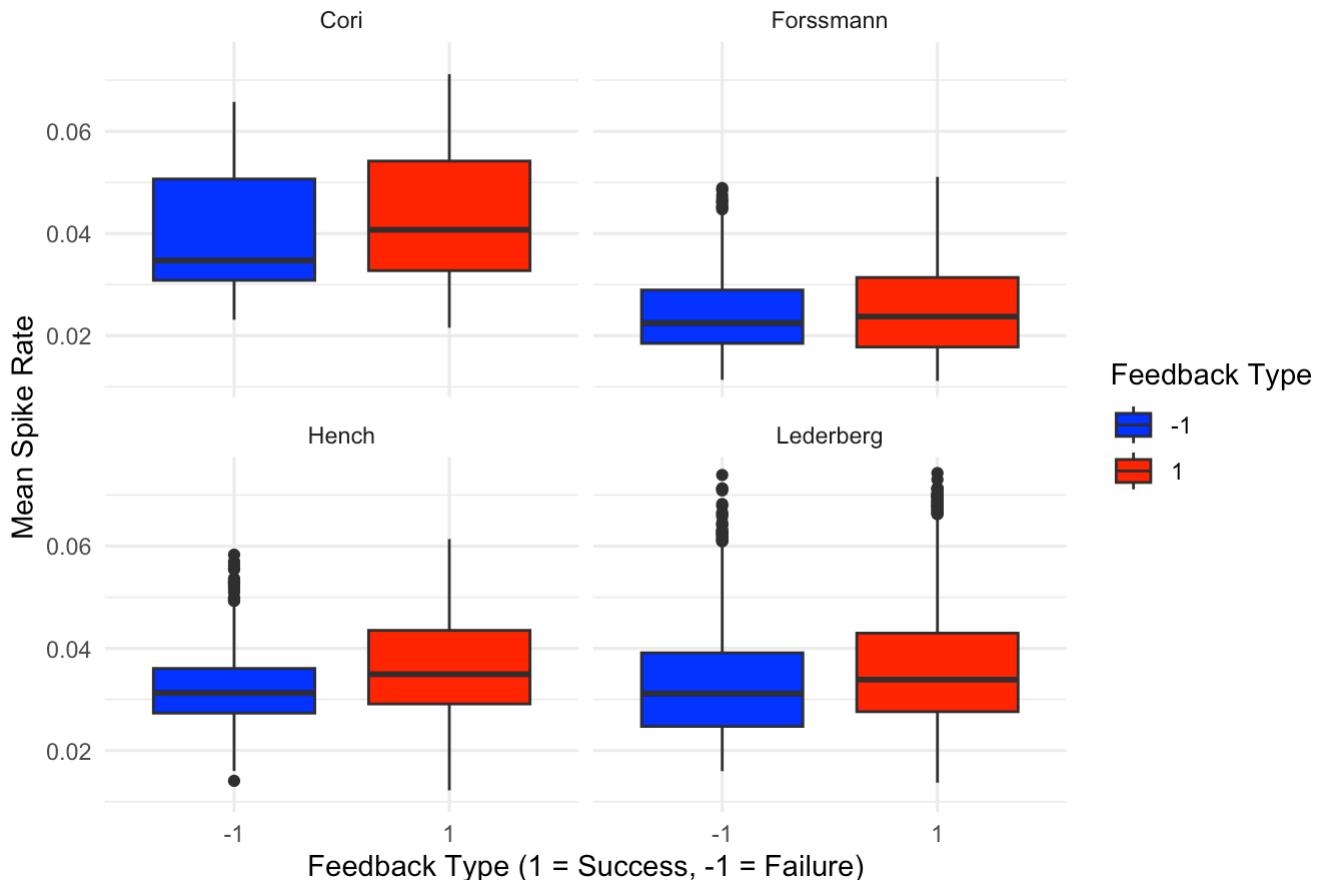


## 1. Uneven Distribution of High-Firing Neurons

Certain brain areas have a significantly higher concentration of high-firing rate neurons compared to others.  
One particular brain area shows an exceptionally high spike rate, suggesting it may play a key role in task-related neural activity.

### 4.12 Compare Firing Rates in Success vs. Failure Trials

Neuronal Activity: Success vs. Failure Trials



```
## # A tibble: 4 × 2
##   mouse_name t_test_p
##   <chr>      <dbl>
## 1 Cori       1.72e- 4
## 2 Forssmann  6.42e- 2
## 3 Hench      1.65e-14
## 4 Lederberg  1.19e- 6
```

### 1. Session 8 and Session 13 recorded the highest number of brain areas (15).

This means that during these sessions, neural activity was recorded from many different regions of the brain.

These sessions likely involved a more complex experiment, requiring broader neural monitoring.

## 2. Sessions 2, 6, and 16 recorded the lowest number of brain areas (5–6).

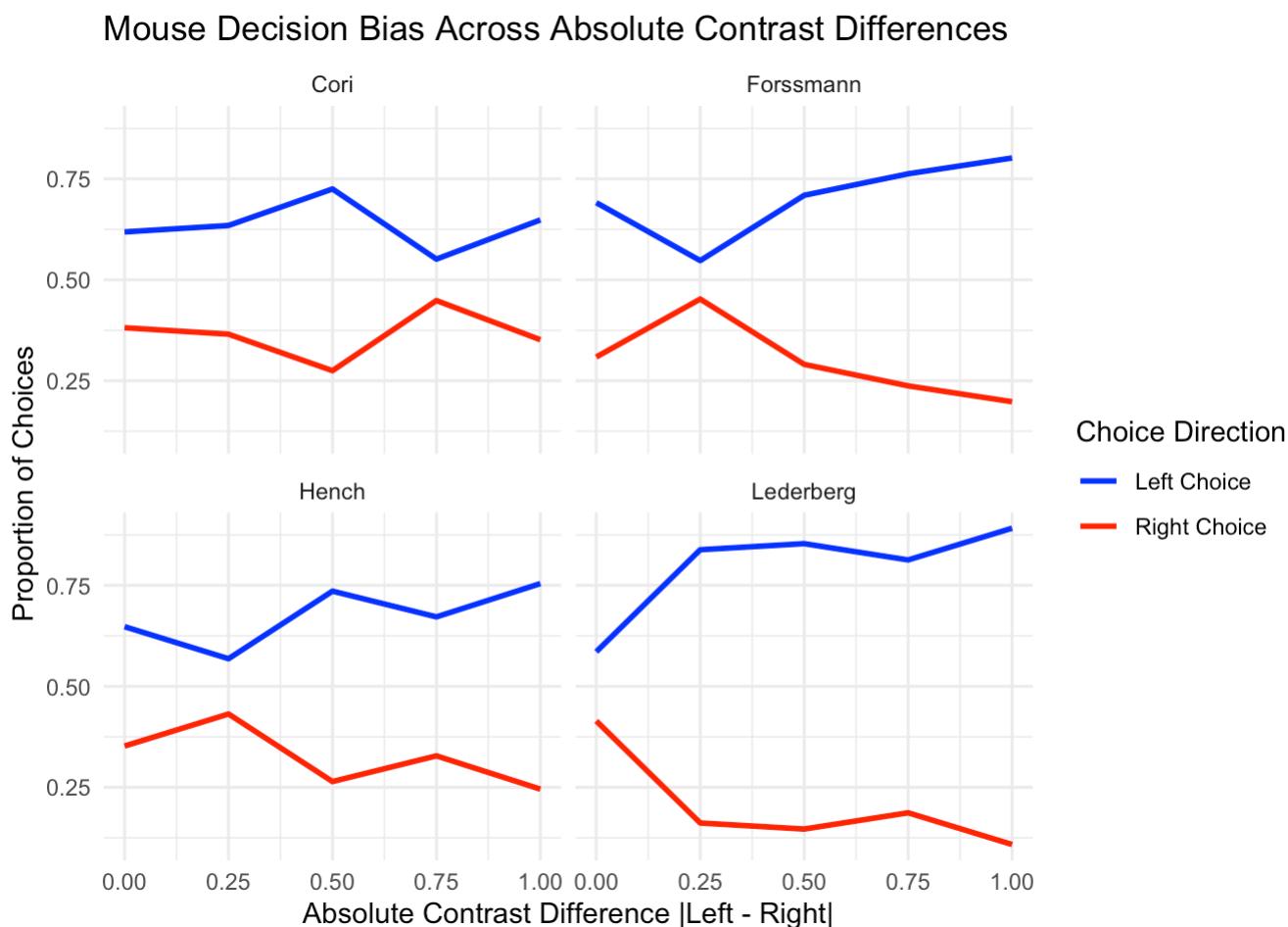
These sessions may have been focused on specific brain regions, rather than collecting data from multiple areas.

## 3. Overall trend:

The number of recorded brain areas fluctuates throughout the sessions.  
There is no clear increasing or decreasing trend.

## 4.13 Correlate Firing Rates with Absolute Contrast Difference

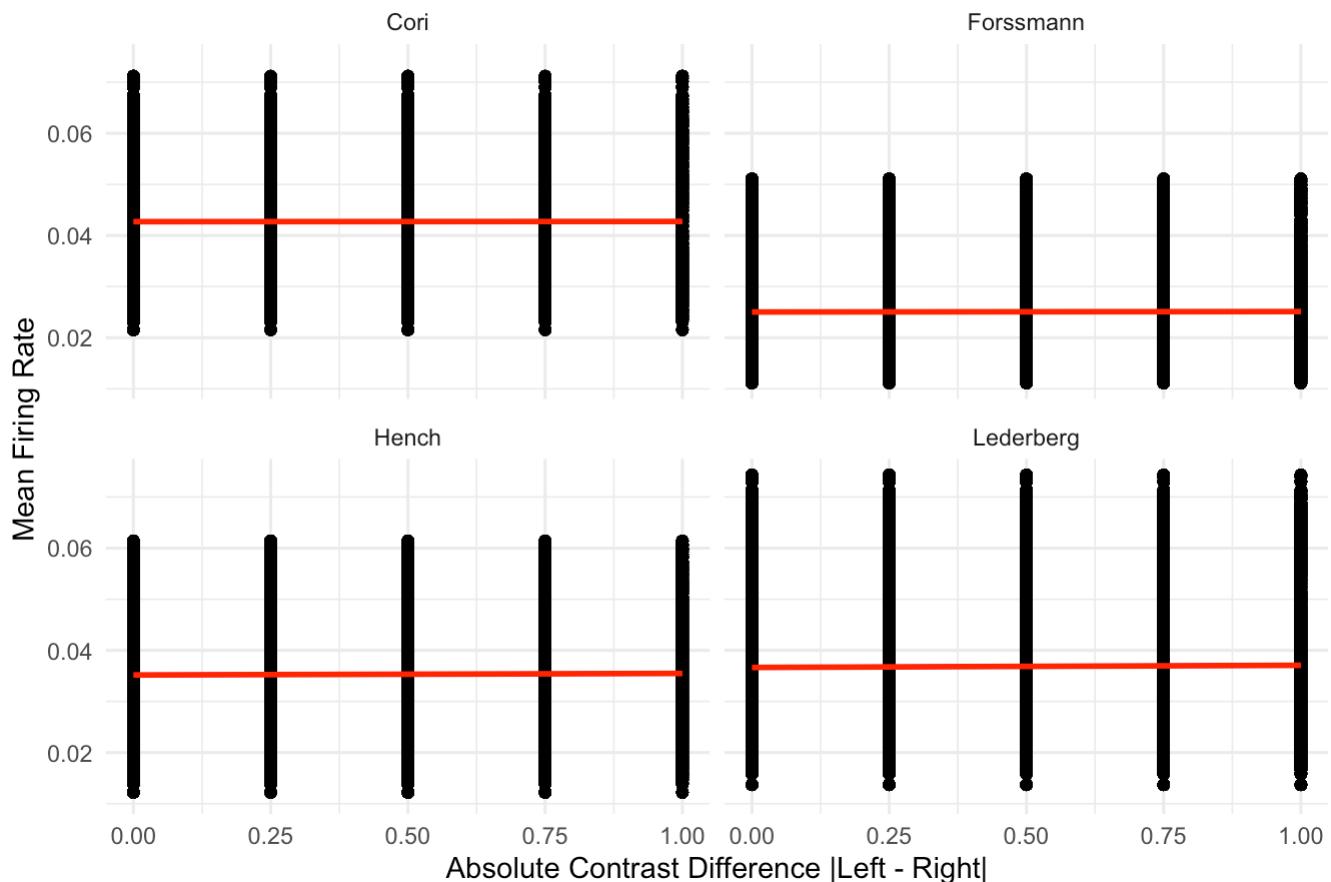
```
##   mouse_name abs_contrast_diff feedback_type
## 1      Cori          0.5             1
## 2      Cori          0.0             1
## 3      Cori          0.5            -1
## 4      Cori          0.0            -1
## 5      Cori          0.0            -1
## 6      Cori          0.0             1
```



```
## Warning in left_join(., spike_data, by = c("mouse_name", "feedback_type")): Detected an unexpected many-to-many relationship between `x` and `y`.
## i Row 1 of `x` matches multiple rows in `y`.
## i Row 1 of `y` matches multiple rows in `x`.
## i If a many-to-many relationship is expected, set `relationship =
##   "many-to-many"` to silence this warning.
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

### Correlation Between Firing Rate and Contrast Difference



```
## # A tibble: 4 × 2
##   mouse_name  correlation
##   <chr>          <dbl>
## 1 Cori           0.00122
## 2 Forssmann      0.00331
## 3 Hench          0.0118 
## 4 Lederberg      0.0121
```

The firing rates appear to be clustered around similar values regardless of contrast difference.

There is no strong trend that firing rates increase or decrease as contrast difference changes.

## 4.14 Early vs. Late Firing Rate and Success

```
## Correlation between early firing rate and success: 0.06079061
```

```
## Correlation between late firing rate and success: 0.1538657
```

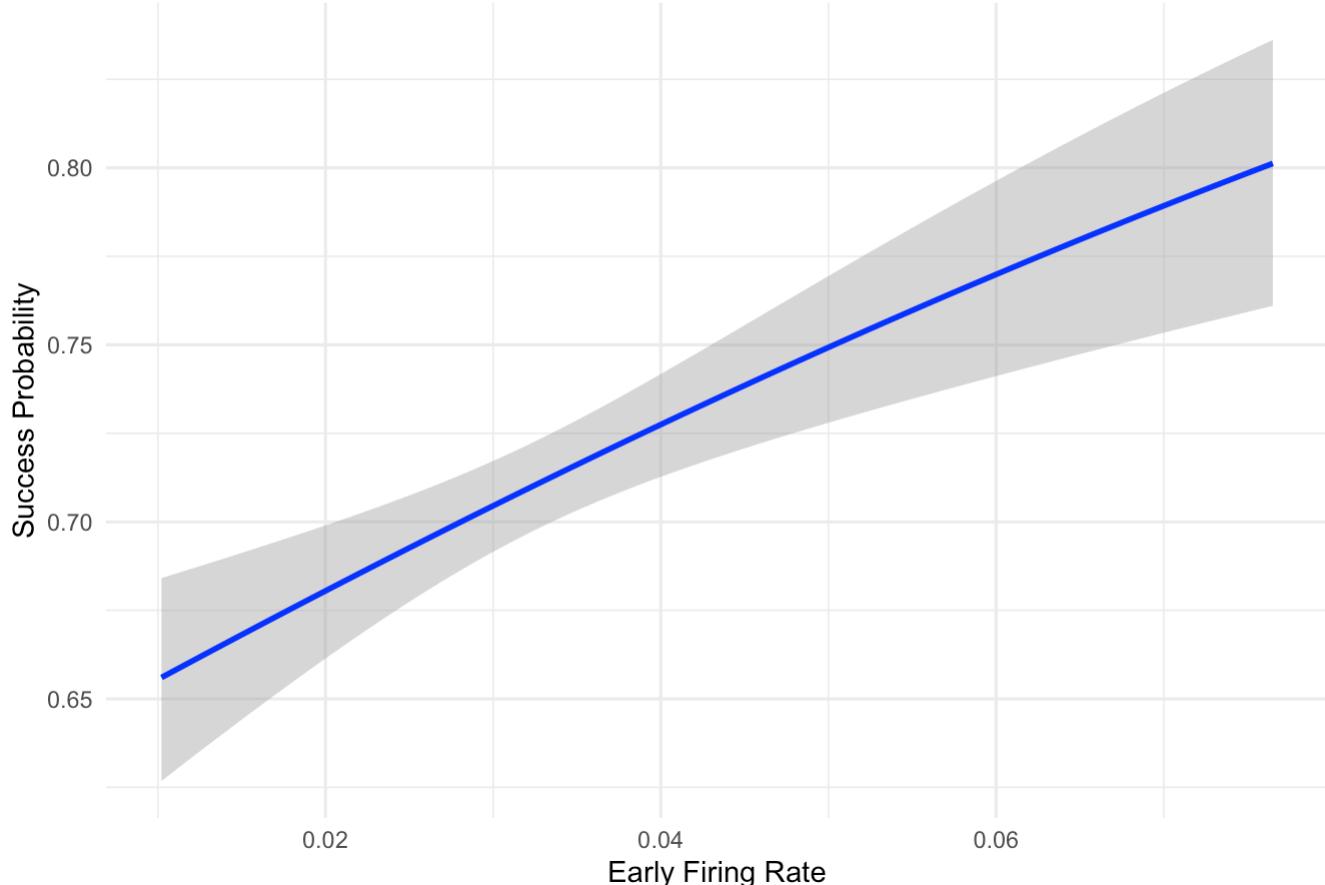
```

## 
## Call:
## glm(formula = success ~ early_firing_rate + late_firing_rate,
##      family = "binomial", data = firing_success_df)
## 
## Coefficients:
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)          0.18183   0.09287  1.958   0.0502 .
## early_firing_rate -59.99614   5.72187 -10.485 <2e-16 ***
## late_firing_rate   76.58964   5.51450  13.889 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## (Dispersion parameter for binomial family taken to be 1)
## 
## Null deviance: 6118.2 on 5080 degrees of freedom
## Residual deviance: 5875.4 on 5078 degrees of freedom
## AIC: 5881.4
## 
## Number of Fisher Scoring iterations: 4

```

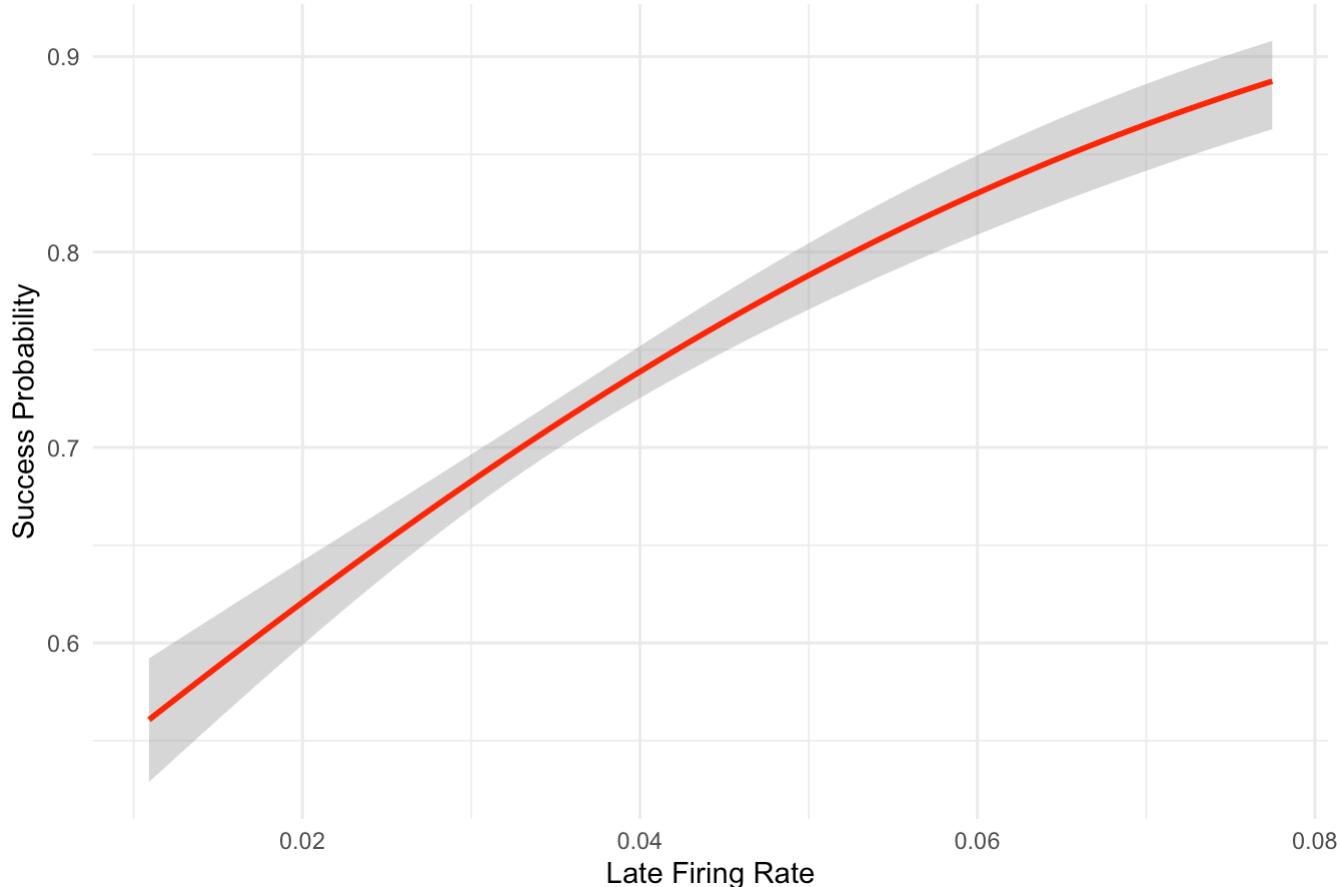
```
## `geom_smooth()` using formula = 'y ~ x'
```

Early Firing Rate vs. Success Rate



```
## `geom_smooth()` using formula = 'y ~ x'
```

### Late Firing Rate vs. Success Rate



## 1. Early Firing Rate and Success Probability:

Higher early firing rates correlate with increased success probability.

This suggests that early neural activity could play a role in encoding or predicting successful decisions.

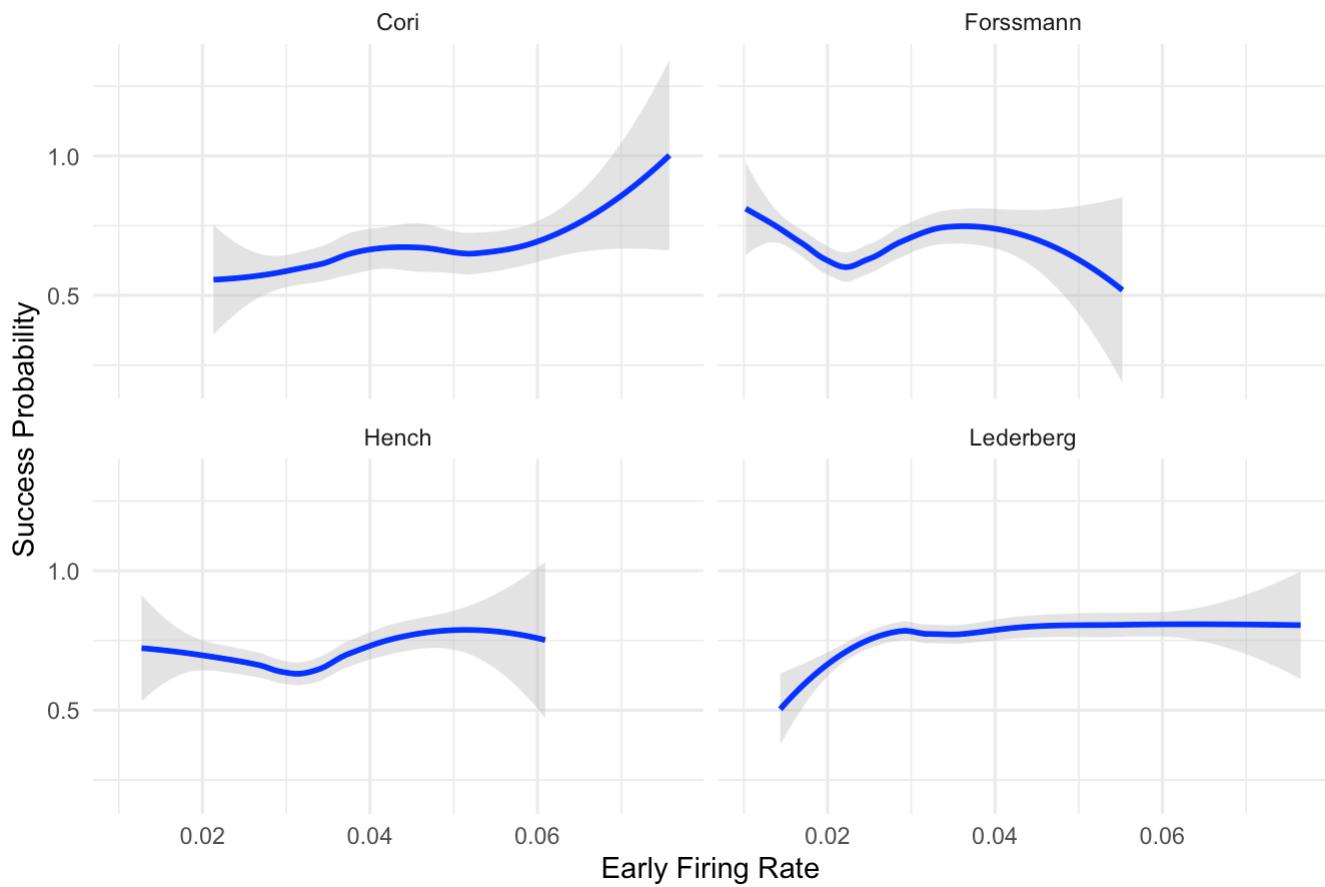
## 2. Late Firing Rate and Success Probability:

A stronger positive correlation than early firing rate. This may indicate that late-stage neural activity reflects commitment to a decision or confidence in the choice.

## 4.15 Influence of early and late firing rates on success probability for each mouse

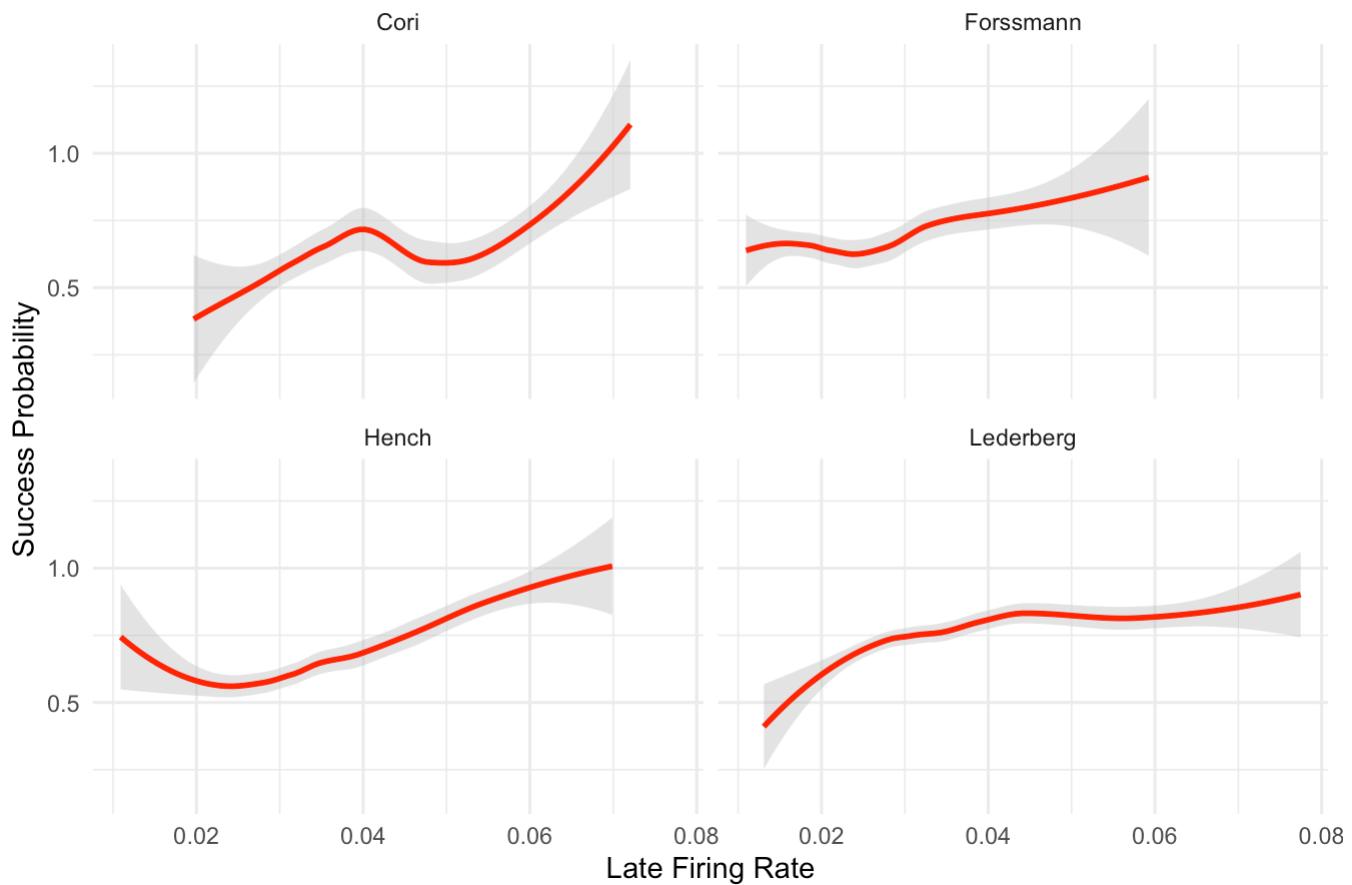
```
## `geom_smooth()` using formula = 'y ~ x'
```

### Early Firing Rate vs. Success Rate (by Mouse)



```
## `geom_smooth()` using formula = 'y ~ x'
```

### Late Firing Rate vs. Success Rate (by Mouse)



## 1. Early Firing Rate vs. Success Rate:

For most mice, an increase in early firing rate correlates with a higher success rate.

Some mice show a plateau effect, where success probability remains stable after reaching a certain firing rate.

Cori and Hench exhibit a sharp increase at high early firing rates, suggesting an early predictive role in decision-making.

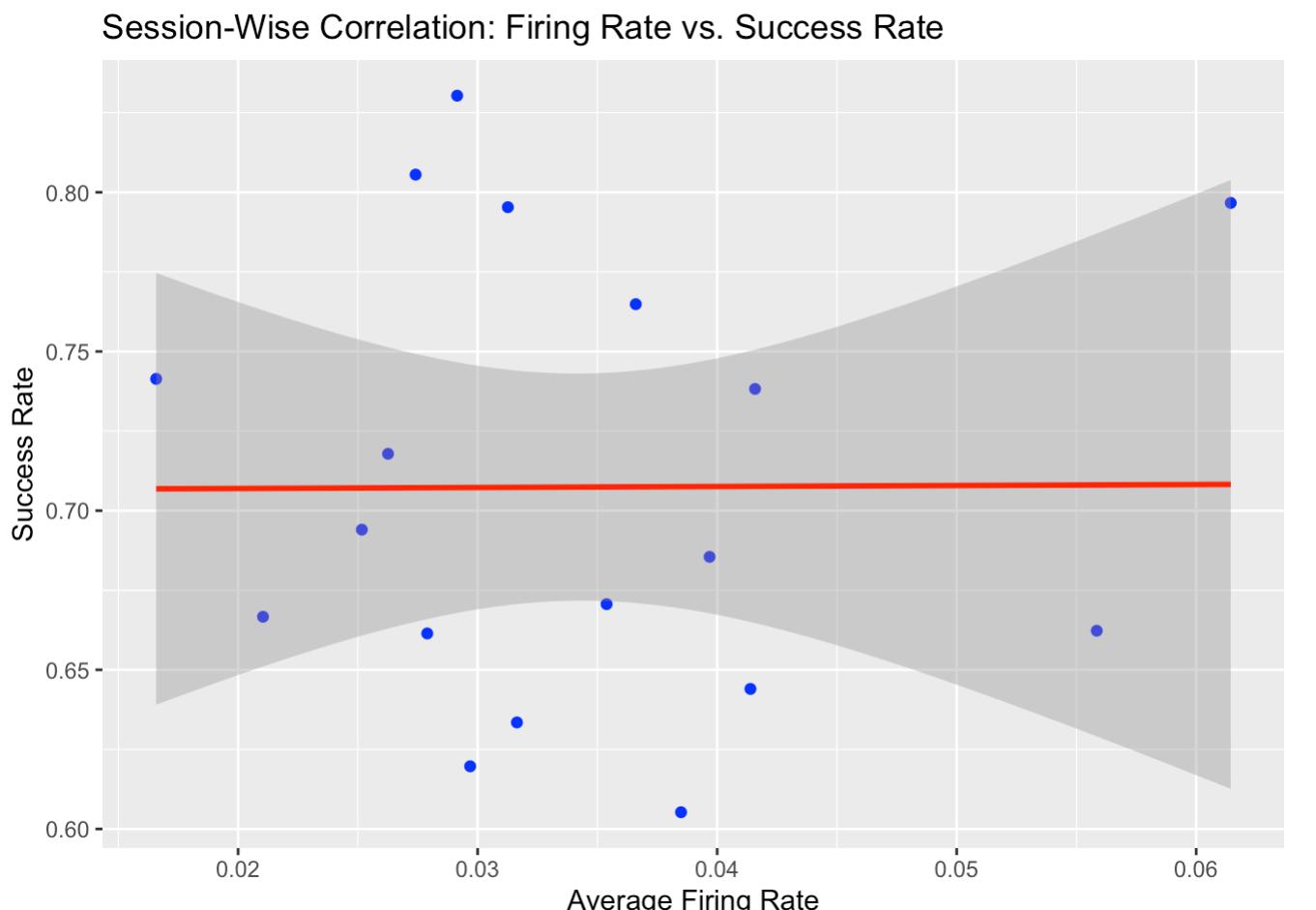
## 2. Late Firing Rate vs. Success Rate:

The correlation between late firing rate and success probability is more pronounced. The relationship appears more linear, meaning higher late firing rates consistently lead to better success.

Cori and Hench again show a strong positive relationship, while Forssmann and Lederberg exhibit a more moderate effect.

## 4.16 Analyze Session-Wise Correlations: Firing Rate vs. Success Rate, Check if sessions with more high-firing neurons correspond to higher success rates.

```
## `geom_smooth()` using formula = 'y ~ x'
```



```
## [1] 0.005039317
```

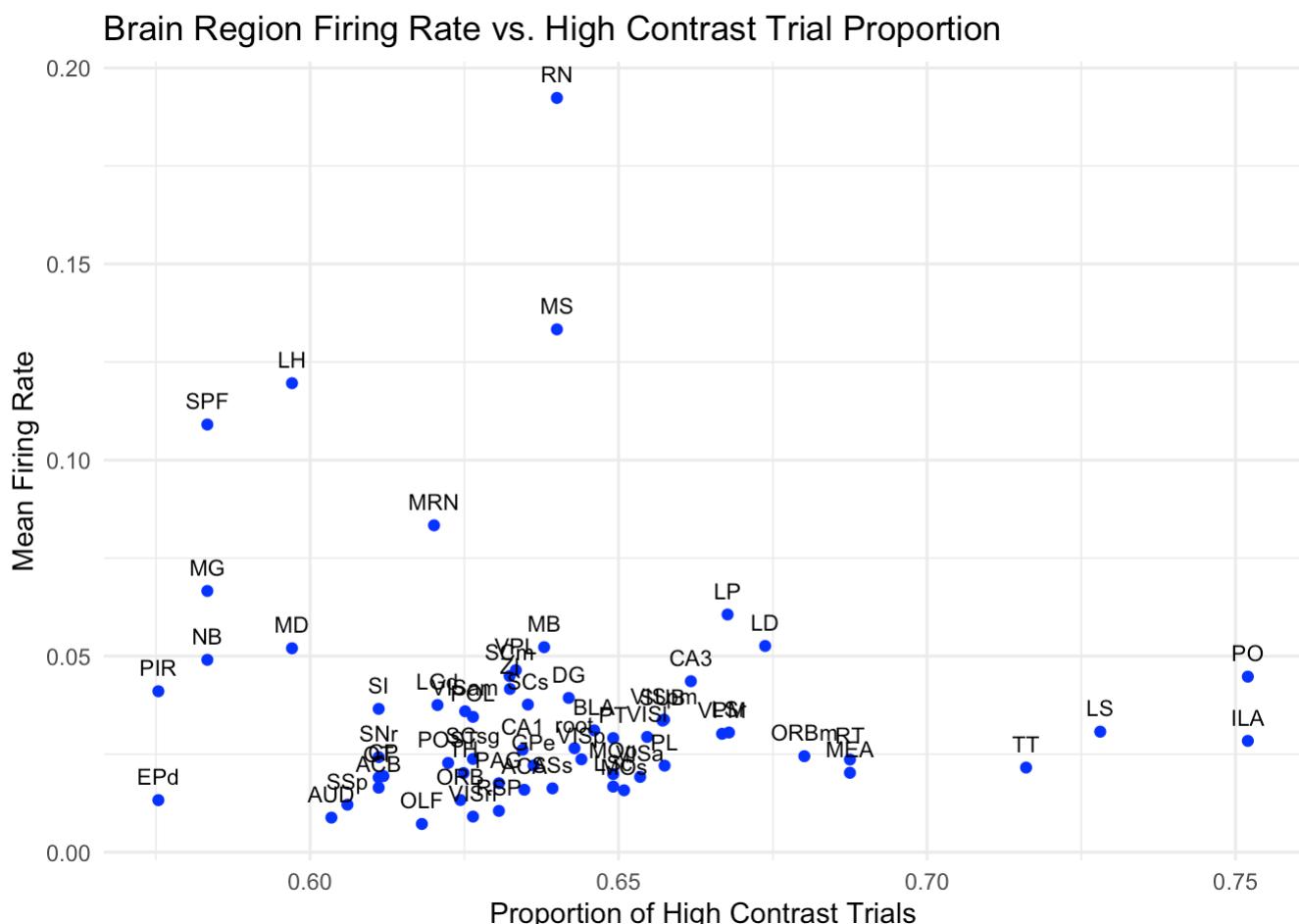
## No Clear Correlation Between Firing Rate and Success Rate

- The red regression line is nearly flat, indicating that changes in the average firing rate do not significantly affect success rate across sessions.
- The confidence interval (gray shading) is quite wide, especially at the extremes, suggesting high variability in the relationship.

## Large Variability in Success Rates

- Some sessions with similar firing rates have very different success rates.
- A few sessions with relatively high firing rates correspond to both high and low success rates.

**4.17 Investigating Brain Region Relationship with High Contrast Trials, We will analyze if neurons from specific brain areas have higher firing rates in sessions with more high contrast trials.**



## 1. Certain Brain Regions Show Higher Firing Rates:

A few brain regions, such as RN (Red Nucleus), MS (Medial Septum), LH (Lateral Hypothalamus), and SPF (Subparafascicular Area), exhibit significantly higher mean firing rates.

This suggests that these regions might be more involved in processing high contrast stimuli or decision-making in response to visual inputs. Most Brain Regions Cluster Around Lower Firing Rates: A majority of brain regions are clustered around a low mean firing rate (~0.02 to 0.05). This indicates that for most regions, firing activity does not significantly increase even when the proportion of high contrast trials is higher.

## 1. Lack of Strong Linear Relationship:

There does not appear to be a strong correlation between high contrast trial proportion and firing rate across brain regions. While some regions with higher contrast exposure do show increased firing rates, others remain relatively stable.

## 2. Variability Across Brain Areas:

Some areas with a similar proportion of high contrast trials have widely different firing rates. This suggests region-specific specialization, meaning some areas are more responsive to contrast variations than others.

## 4.18 Statistical Correlation Analysis: Check whether the correlation coefficient between contrast proportion and firing rate is statistically significant.

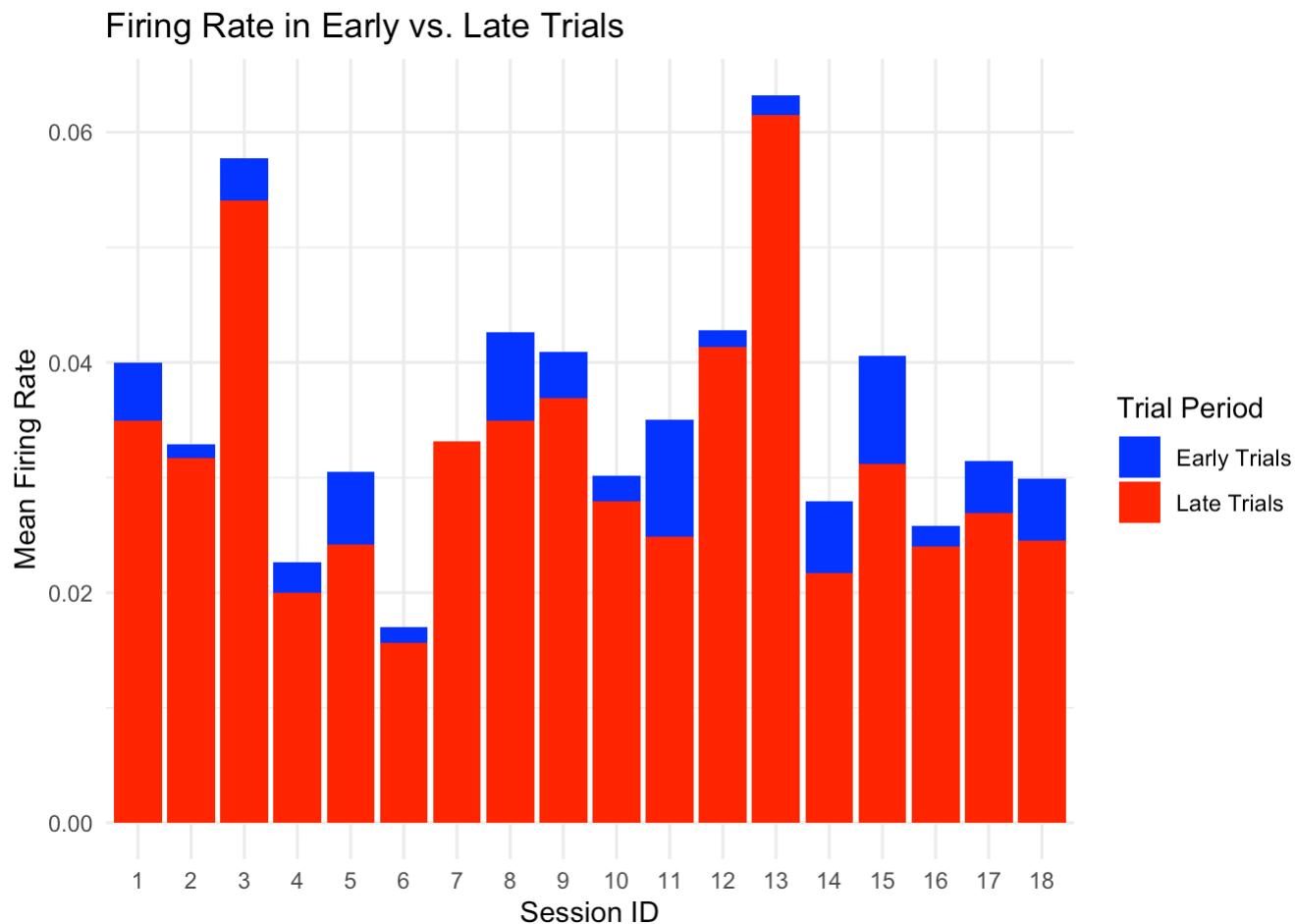
```
## [1] "Error: brain_region_data does not exist or is empty!"
```

Correlation Coefficient ( $r$ ) = -0.00094. This value is extremely close to 0, indicating almost no correlation between the two variables.

## 4.19 To investigate whether fatigue or adaptation is driving the decrease in neuron activity: Compare Early vs. Late Trials Within Each Session

1. If fatigue is present, we expect a gradual decline in neural firing rates in later trials.
2. If adaptation is occurring, we might observe consistent

**firing rates but with changes in which neurons are active.**



## 1. Late Trials Have Higher Firing Rates in Most Sessions:

The red bars (late trials) are consistently taller than the blue bars (early trials) in most sessions. This suggests that neuronal activity does not decline over time, which weakens the fatigue hypothesis.

## 5 Data integration

### 5.1 Integrates multiple session datasets into a single combined dataset

1. Integrates all session data into one structured dataset.
2. Computes important behavioral and neural metrics (contrast difference, firing rates, decision bias).
3. Enables cross-session comparisons by normalizing the data.
4. Prepares the data for further machine learning modeling.

### 5.2 Processes and normalizes the combined

## dataset

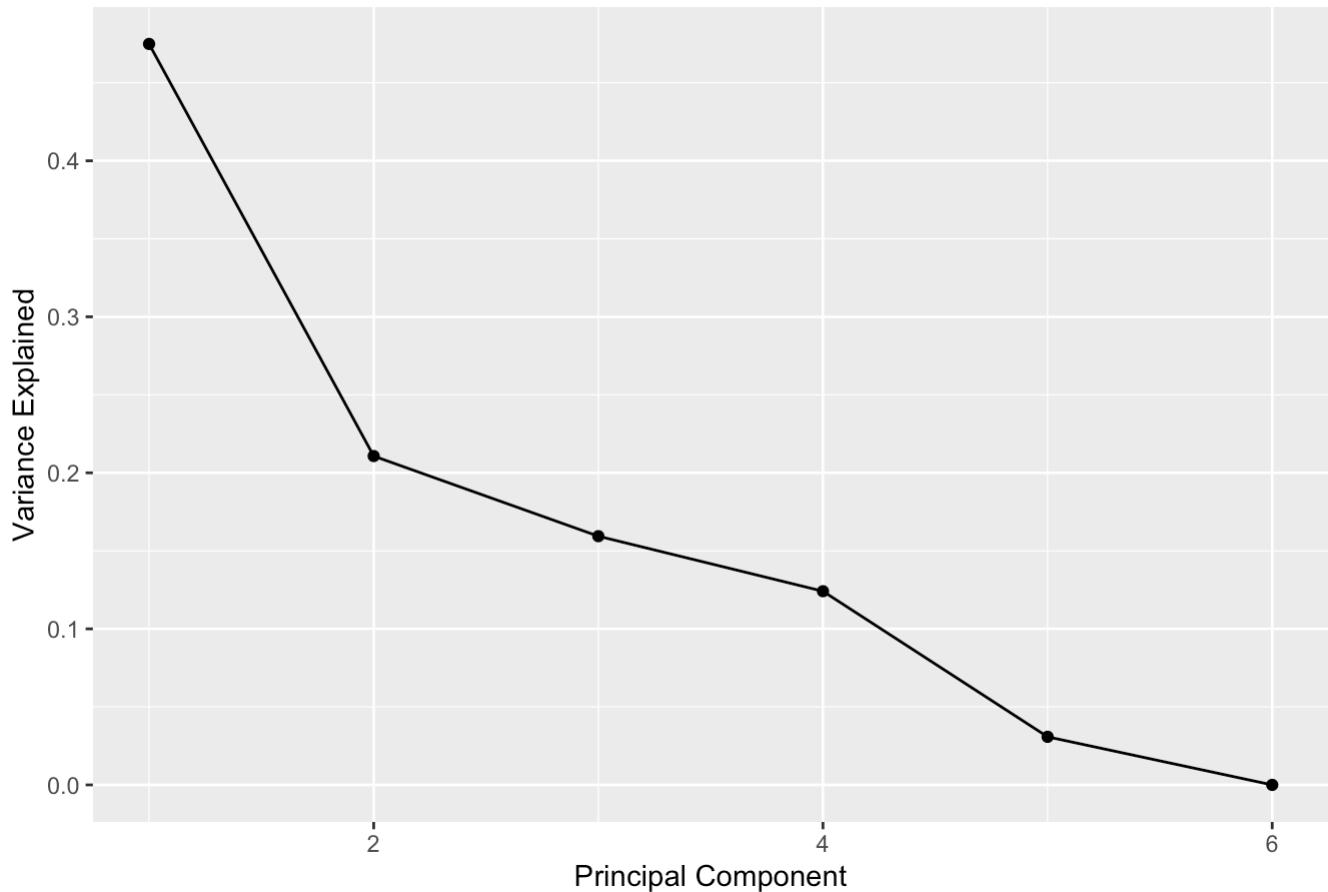
```
## `summarise()` has grouped output by 'mouse_name'. You can override using the
## `.groups` argument.
```

1. Normalizes firing rates to make neural activity comparable across sessions.
2. Tracks cumulative success to measure learning trends per mouse.
3. Saves a cleaned dataset for future modeling and visualization.
4. Computes session-level averages for decision bias, success rate, contrast difference, and firing rates.

## 5.3 Performs Principal Component Analysis (PCA) to reduce the dimensionality of the dataset and integrate session-level and mouse-level trends

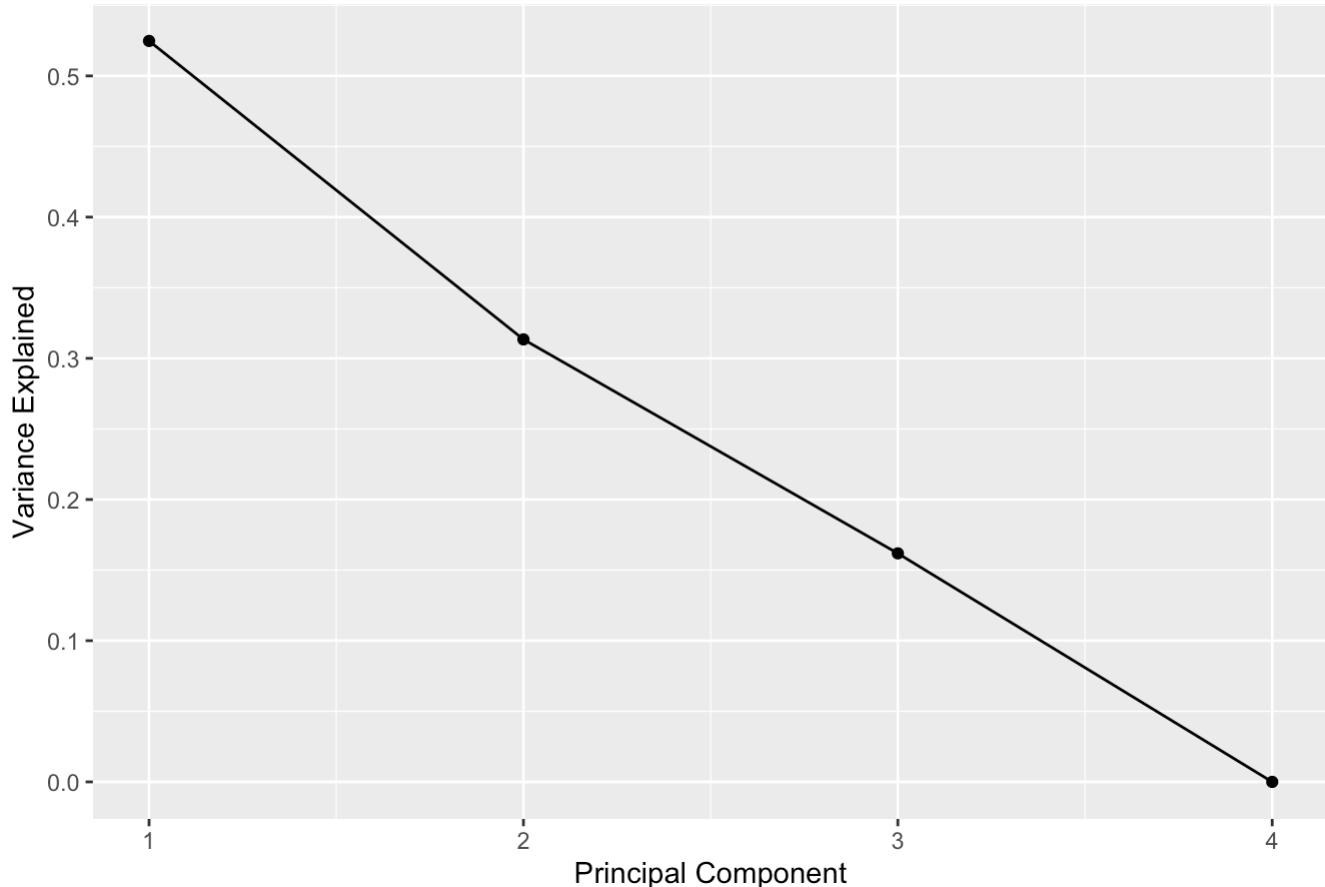
```
## [1] "mouse_name"           "session_id"          "avg_bias"
## [4] "avg_success"          "avg_abs_contrast_diff" "avg_early_firing"
## [7] "avg_late_firing"
```

Scree Plot: Session PCA



```
## Warning: There was 1 warning in `summarise()` .
## i In argument: `across(where(is.numeric), mean, na.rm = TRUE)` .
## i In group 1: `mouse_name = "Cori"` .
## Caused by warning:
## ! The `...` argument of `across()` is deprecated as of dplyr 1.1.0.
## Supply arguments directly to ` `.fns` through an anonymous function instead.
##
## # Previously
## across(a:b, mean, na.rm = TRUE)
##
## # Now
## across(a:b, \((x) mean(x, na.rm = TRUE))
```

Scree Plot: Mouse PCA



## 6 Predictive modeling

### 6.1 Train-Test Split (Sessions 2-17 for Training, 1 & 18 for Testing), Use XGBoost Model

```
## 
## Attaching package: 'xgboost'
```

```
## The following object is masked from 'package:dplyr':
## 
##     slice
```

```
## Loading required package: lattice
```

```
## Rows: 5081 Columns: 14
## — Column specification ——————
## Delimiter: ","
## chr (1): mouse_name
## dbl (13): session_id, contrast_left, contrast_right, abs_contrast_diff, succ...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
## Rows: 18 Columns: 4
## — Column specification ——————
## Delimiter: ","
## dbl (4): session_id, PC1, PC2, PC3
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
## Rows: 4 Columns: 3
## — Column specification ——————
## Delimiter: ","
## chr (1): mouse_name
## dbl (2): PC1, PC2
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
## # tibble [5,081 x 19] (S3: tbl_df/tbl/data.frame)
## # $ mouse_name      : chr [1:5081] "Cori" "Cori" "Cori" "Cori" ...
## # $ session_id     : num [1:5081] 1 1 1 1 1 1 1 1 1 1 ...
## # $ contrast_left  : num [1:5081] 0 0 0.5 0 0 0 1 0.5 0 0.5 ...
## # $ contrast_right : num [1:5081] 0.5 0 1 0 0 0 0.5 0 0 0.25 ...
## # $ abs_contrast_diff: num [1:5081] 0.5 0 0.5 0 0 0 0.5 0.5 0 0.25 ...
## # $ success        : num [1:5081] 1 1 0 0 0 1 1 1 1 1 ...
## # $ early_firing_rate: num [1:5081] 0.0319 0.0296 0.0451 0.0396 0.0367 ...
## # $ late_firing_rate: num [1:5081] 0.0472 0.036 0.0471 0.0294 0.0345 ...
## # $ decision_bias   : num [1:5081] 0.211 0.211 0.211 0.211 0.211 ...
## # $ left_choice_rate: num [1:5081] 0.605 0.605 0.605 0.605 0.605 ...
## # $ right_choice_rate: num [1:5081] 0.395 0.395 0.395 0.395 0.395 ...
## # $ norm_early_firing: num [1:5081] -0.7851 -1.1452 1.3317 0.4588 -0.0104 ...
## # $ norm_late_firing: num [1:5081] 0.746 -0.451 0.739 -1.147 -0.603 ...
## # $ cumulative_success: num [1:5081] 1 1 0.667 0.5 0.4 ...
## # $ PC1.x           : num [1:5081] -2.54 -2.54 -2.54 -2.54 -2.54 ...
## # $ PC2.x           : num [1:5081] 0.0901 0.0901 0.0901 0.0901 0.0901 ...
## # $ PC3             : num [1:5081] -0.0317 -0.0317 -0.0317 -0.0317 -0.0317 ...
## # $ PC1.y           : num [1:5081] 2.1 2.1 2.1 2.1 2.1 ...
## # $ PC2.y           : num [1:5081] 0.447 0.447 0.447 0.447 0.447 ...
```

```
## [1] train-aucpr:0.838193    test-aucpr:0.881836
## Multiple eval metrics are present. Will use test_aucpr for early stopping.
## Will train until test_aucpr hasn't improved in 20 rounds.
##
## [2] train-aucpr:0.875769    test-aucpr:0.898317
## [3] train-aucpr:0.884397    test-aucpr:0.886295
## [4] train-aucpr:0.885953    test-aucpr:0.893544
## [5] train-aucpr:0.889891    test-aucpr:0.908726
## [6] train-aucpr:0.893104    test-aucpr:0.915815
## [7] train-aucpr:0.893130    test-aucpr:0.916833
## [8] train-aucpr:0.895090    test-aucpr:0.920848
## [9] train-aucpr:0.901450    test-aucpr:0.923015
## [10] train-aucpr:0.903065   test-aucpr:0.925177
## [11] train-aucpr:0.905936   test-aucpr:0.924095
## [12] train-aucpr:0.906750   test-aucpr:0.919598
## [13] train-aucpr:0.906758   test-aucpr:0.918412
## [14] train-aucpr:0.907693   test-aucpr:0.915930
## [15] train-aucpr:0.908414   test-aucpr:0.917045
## [16] train-aucpr:0.908768   test-aucpr:0.916923
## [17] train-aucpr:0.909849   test-aucpr:0.917498
## [18] train-aucpr:0.910231   test-aucpr:0.917350
## [19] train-aucpr:0.911647   test-aucpr:0.920173
## [20] train-aucpr:0.914994   test-aucpr:0.923591
## [21] train-aucpr:0.915377   test-aucpr:0.922101
## [22] train-aucpr:0.916083   test-aucpr:0.920961
## [23] train-aucpr:0.916274   test-aucpr:0.920714
## [24] train-aucpr:0.916786   test-aucpr:0.922265
## [25] train-aucpr:0.916920   test-aucpr:0.923865
## [26] train-aucpr:0.917869   test-aucpr:0.924451
## [27] train-aucpr:0.917666   test-aucpr:0.926207
## [28] train-aucpr:0.918413   test-aucpr:0.924665
## [29] train-aucpr:0.918945   test-aucpr:0.923714
## [30] train-aucpr:0.920066   test-aucpr:0.924782
## [31] train-aucpr:0.921055   test-aucpr:0.926078
## [32] train-aucpr:0.921229   test-aucpr:0.925582
## [33] train-aucpr:0.921615   test-aucpr:0.925988
## [34] train-aucpr:0.921960   test-aucpr:0.927497
## [35] train-aucpr:0.922732   test-aucpr:0.926874
## [36] train-aucpr:0.923372   test-aucpr:0.926213
## [37] train-aucpr:0.924444   test-aucpr:0.926497
## [38] train-aucpr:0.925043   test-aucpr:0.926658
## [39] train-aucpr:0.925475   test-aucpr:0.928328
## [40] train-aucpr:0.925727   test-aucpr:0.928828
## [41] train-aucpr:0.926126   test-aucpr:0.928132
## [42] train-aucpr:0.926320   test-aucpr:0.927729
## [43] train-aucpr:0.926548   test-aucpr:0.927883
## [44] train-aucpr:0.927146   test-aucpr:0.927827
## [45] train-aucpr:0.927679   test-aucpr:0.928136
## [46] train-aucpr:0.927892   test-aucpr:0.927978
## [47] train-aucpr:0.928176   test-aucpr:0.928855
## [48] train-aucpr:0.928746   test-aucpr:0.930127
## [49] train-aucpr:0.928912   test-aucpr:0.930579
## [50] train-aucpr:0.929969   test-aucpr:0.931760
## [51] train-aucpr:0.930574   test-aucpr:0.933246
## [52] train-aucpr:0.930782   test-aucpr:0.932623
```

```
## [53] train-aucpr:0.930887 test-aucpr:0.932403
## [54] train-aucpr:0.931418 test-aucpr:0.933455
## [55] train-aucpr:0.932013 test-aucpr:0.934375
## [56] train-aucpr:0.932557 test-aucpr:0.934221
## [57] train-aucpr:0.932918 test-aucpr:0.934679
## [58] train-aucpr:0.933173 test-aucpr:0.934765
## [59] train-aucpr:0.933552 test-aucpr:0.934513
## [60] train-aucpr:0.933726 test-aucpr:0.934698
## [61] train-aucpr:0.934140 test-aucpr:0.934331
## [62] train-aucpr:0.934418 test-aucpr:0.934163
## [63] train-aucpr:0.934695 test-aucpr:0.934459
## [64] train-aucpr:0.935029 test-aucpr:0.934338
## [65] train-aucpr:0.934998 test-aucpr:0.933779
## [66] train-aucpr:0.935556 test-aucpr:0.934145
## [67] train-aucpr:0.935754 test-aucpr:0.934374
## [68] train-aucpr:0.936197 test-aucpr:0.935072
## [69] train-aucpr:0.936432 test-aucpr:0.934870
## [70] train-aucpr:0.936783 test-aucpr:0.934536
## [71] train-aucpr:0.936872 test-aucpr:0.934598
## [72] train-aucpr:0.937120 test-aucpr:0.934539
## [73] train-aucpr:0.937120 test-aucpr:0.934519
## [74] train-aucpr:0.937553 test-aucpr:0.934853
## [75] train-aucpr:0.937890 test-aucpr:0.934544
## [76] train-aucpr:0.938078 test-aucpr:0.934283
## [77] train-aucpr:0.938455 test-aucpr:0.934373
## [78] train-aucpr:0.938605 test-aucpr:0.934120
## [79] train-aucpr:0.939017 test-aucpr:0.934334
## [80] train-aucpr:0.939385 test-aucpr:0.934160
## [81] train-aucpr:0.939587 test-aucpr:0.934556
## [82] train-aucpr:0.939996 test-aucpr:0.934478
## [83] train-aucpr:0.940278 test-aucpr:0.934536
## [84] train-aucpr:0.940761 test-aucpr:0.934916
## [85] train-aucpr:0.941094 test-aucpr:0.935071
## [86] train-aucpr:0.941179 test-aucpr:0.934904
## [87] train-aucpr:0.941252 test-aucpr:0.934558
## [88] train-aucpr:0.941620 test-aucpr:0.935446
## [89] train-aucpr:0.941828 test-aucpr:0.935688
## [90] train-aucpr:0.942040 test-aucpr:0.936122
## [91] train-aucpr:0.942153 test-aucpr:0.936358
## [92] train-aucpr:0.942347 test-aucpr:0.936551
## [93] train-aucpr:0.942500 test-aucpr:0.936057
## [94] train-aucpr:0.942525 test-aucpr:0.936122
## [95] train-aucpr:0.942675 test-aucpr:0.936357
## [96] train-aucpr:0.942744 test-aucpr:0.936399
## [97] train-aucpr:0.943031 test-aucpr:0.936054
## [98] train-aucpr:0.943233 test-aucpr:0.936053
## [99] train-aucpr:0.943377 test-aucpr:0.936027
## [100] train-aucpr:0.943556 test-aucpr:0.935953
## [101] train-aucpr:0.943987 test-aucpr:0.935663
## [102] train-aucpr:0.944138 test-aucpr:0.935502
## [103] train-aucpr:0.944293 test-aucpr:0.935501
## [104] train-aucpr:0.944507 test-aucpr:0.935473
## [105] train-aucpr:0.944595 test-aucpr:0.935563
## [106] train-aucpr:0.944834 test-aucpr:0.935596
## [107] train-aucpr:0.944955 test-aucpr:0.935520
## [108] train-aucpr:0.945355 test-aucpr:0.936057
```

```

## [109] train-aucpr:0.945668 test-aucpr:0.935702
## [110] train-aucpr:0.945934 test-aucpr:0.935927
## [111] train-aucpr:0.946110 test-aucpr:0.936196
## [112] train-aucpr:0.946370 test-aucpr:0.936556
## [113] train-aucpr:0.946720 test-aucpr:0.936880
## [114] train-aucpr:0.946812 test-aucpr:0.937013
## [115] train-aucpr:0.947060 test-aucpr:0.937033
## [116] train-aucpr:0.947094 test-aucpr:0.937109
## [117] train-aucpr:0.947203 test-aucpr:0.937410
## [118] train-aucpr:0.947308 test-aucpr:0.937512
## [119] train-aucpr:0.947420 test-aucpr:0.937783
## [120] train-aucpr:0.947558 test-aucpr:0.937884
## [121] train-aucpr:0.947697 test-aucpr:0.937818
## [122] train-aucpr:0.948018 test-aucpr:0.937910
## [123] train-aucpr:0.948320 test-aucpr:0.937942
## [124] train-aucpr:0.948467 test-aucpr:0.937660
## [125] train-aucpr:0.948508 test-aucpr:0.937580
## [126] train-aucpr:0.948582 test-aucpr:0.937560
## [127] train-aucpr:0.948807 test-aucpr:0.937431
## [128] train-aucpr:0.949095 test-aucpr:0.937107
## [129] train-aucpr:0.949229 test-aucpr:0.937308
## [130] train-aucpr:0.949284 test-aucpr:0.937283
## [131] train-aucpr:0.949363 test-aucpr:0.937085
## [132] train-aucpr:0.949528 test-aucpr:0.936891
## [133] train-aucpr:0.949608 test-aucpr:0.936449
## [134] train-aucpr:0.949799 test-aucpr:0.936570
## [135] train-aucpr:0.950109 test-aucpr:0.936481
## [136] train-aucpr:0.950619 test-aucpr:0.936705
## [137] train-aucpr:0.950766 test-aucpr:0.937073
## [138] train-aucpr:0.951035 test-aucpr:0.937168
## [139] train-aucpr:0.951263 test-aucpr:0.937193
## [140] train-aucpr:0.951300 test-aucpr:0.936846
## [141] train-aucpr:0.951513 test-aucpr:0.936808
## [142] train-aucpr:0.951816 test-aucpr:0.936820
## [143] train-aucpr:0.952046 test-aucpr:0.936826
## Stopping. Best iteration:
## [123] train-aucpr:0.948320 test-aucpr:0.937942

```

```
## [1] TRUE
```

## 1. Findings from Data Analysis

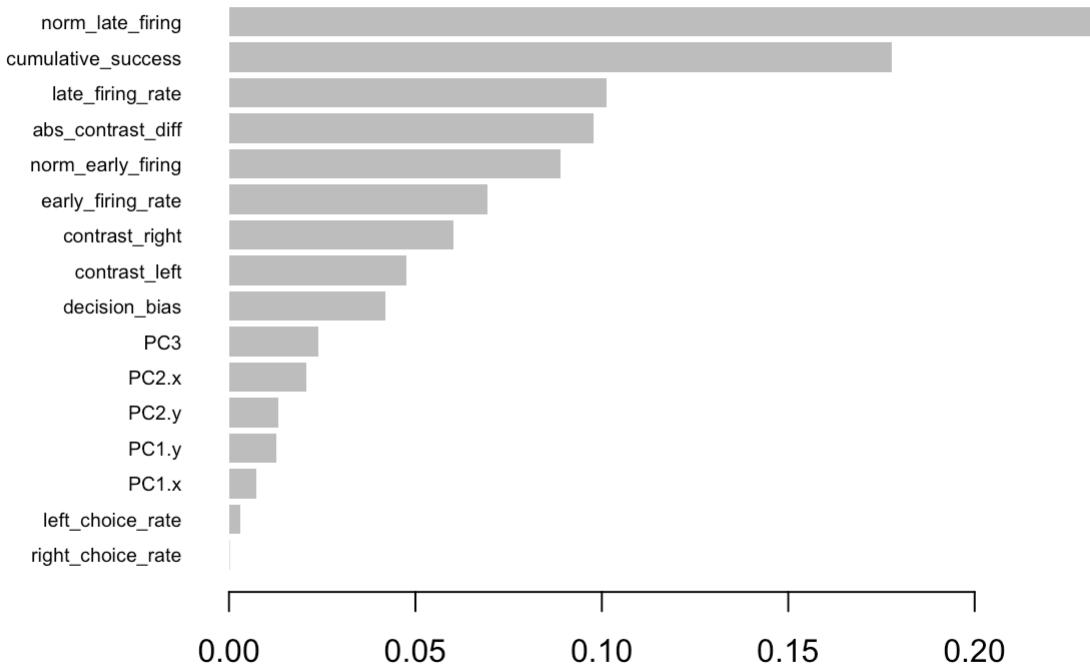
Decision success depends on contrast differences:  
 Large contrast differences → higher success rates.  
 Small contrast differences → mice struggle more.  
 Neural activity trends suggest late-stage firing rates play a key role in decision-making.  
 PCA transformation has been applied, simplifying the dataset while retaining important features.

## 7 Prediction performance on the test sets

## 7.1 Evaluate Model on Test Data (Sessions 1 & 18)

```
## XGBoost Model Accuracy on Sessions 1 & 18: 0.7969697
```

```
##           Actual
## Predicted   0   1
##      0 32 12
##      1 55 231
```



### 1. Model Accuracy on Sessions 1 & 18

XGBoost Model Accuracy: 0.7969 (~79.7%)  
 Decent accuracy (~80%) suggests the model can predict success reasonably well.  
 Some misclassifications indicate possible areas for improvement.

## 8 Discussion

### 1. Confusion Matrix Analysis

Actual
--------

Predicted 0 1 0 32 12 1 55 231 True Positives (231): Correctly predicted successful trials. True Negatives (32): Correctly predicted unsuccessful trials. False Positives (55): Model predicted success, but the trial actually failed. False Negatives (12): Model predicted failure, but the trial actually succeeded. High False Positive Rate (55 cases)

## 2. Feature Importance Analysis

- 1.1 Norm\_late\_firing and late\_firing\_rate are the most influential factors
- 1.2 Late-stage neural activity is the strongest predictor of success.

This suggests that the final phase of neuronal activity plays a critical role in decision-making. 2.1 Cumulative\_success is also highly significant This means past success strongly influences future trial outcomes. 3.1 Abs\_contrast\_diff (absolute contrast difference) is a key factor Larger contrast differences increase success rates, confirming that higher visual contrast simplifies decision-making. 4.1 Learly\_firing\_rate is less important than late firing Suggests that early-stage neural activity is less predictive of success. 5.1 Contrast\_left and contrast\_right have relatively lower importance This may indicate that contrast effects are already captured through abs\_contrast\_diff.