Kasdin et al. (2025) show that dopamine in the brains of young zebra finches acts as a learning signal, increasing when they sing closer to their adult song and decreasing when they sing further away, effectively guiding their vocal development through trial-and-error. This suggests that complex natural behaviors, like learning to sing, are shaped by dopamine-driven reinforcement learning, similar to how artificial intelligence learns. You can find the paper at this link: https://www.nature.com/articles/s41586-025-08729-1..

Note they measure dopamine using fibre photometry, changes in the fluorescence indicate dopamine changes in realtime. Their specific measurement considers changes in flourescence in 100-ms windows between 200 and 300 ms from the start of singing, averaged across development.

1. Using the `pwr` package for `R` (Champely, 2020), conduct a power analysis. How many observations would the researchers need to detect a moderate-to-large effect ($d = 0.65$) when using $\alpha = 0.05$ and default power (0.80) for a two-sided one sample t test.

```
library(tidyverse)
library(pwr)

power.result <- power.t.test(delta=0.65, sig.level=0.05, power=0.80,
                             type="one.sample", alternative="two.sided")
(n.val <- power.result$n)


## [1] 20.58044
```

   To determine the required number of observations to detect a moderate-to-large effect($d = 0.65$) with a significance level of $\alpha = 0.05$ and a power of 0.80 for a two-sided one-sample t-test, we used the `pwr` package in `R` (Champely, 2020).The power analysis indicates that approximately 21 observations are needed to reliability detect this effect size. This sample size would give researchers an 80% chance of correctly rejecting the null hypothesis when the effect is truly present.

2. Click the link to go to the paper. Find the source data for Figure 2. Download the Excel file. Describe what you needed to do to collect the data for Figure 2(g). Note that you only need the `closer_vals` and `further_vals`. Ensure to `mutate()` the data to get a difference (e.g., `closer_vals - further_vals`).

```
dat <- read.csv("data.csv")

dat.clean <- dat |>
  mutate(difference = closer_vals - further_vals)
(dat.clean)


##    further_vals closer_vals difference
## 1   -0.19168505 0.275714146 0.46739919
## 2   -0.02993474 0.084044448 0.11397919
## 3   -0.18674609 0.131846414 0.31859251
## 4   -0.07381857 0.036630190 0.11044876
## 5   -0.08879365 0.145534134 0.23432778
## 6   -0.16044005 0.195669989 0.35611003
## 7   -0.30926059 0.158454214 0.46771480
## 8   -0.15890017 0.179967148 0.33886732
## 9   -0.12535659 0.131045171 0.25640176
## 10  -0.05447568 0.096865366 0.15134104
## 11  -0.60278594 0.329094427 0.93188037
## 12  -0.23766536 0.169982634 0.40764799
## 13  -0.12270538 0.074888134 0.19759351
## 14  -0.12990904 0.004057814 0.13396685
## 15  -0.04227093 0.001082100 0.04335303
## 16  -0.24306660 0.089018022 0.33208462
## 17  -0.34456488 0.260877405 0.60544228
## 18  -0.31220099 0.276972646 0.58917363
## 19  -0.17620137 0.072787838 0.24898921
## 20  -0.07446303 0.154177639 0.22864067
## 21  -0.34440941 0.266133888 0.61054330
## 22  -0.19040289 0.145314346 0.33571724
## 23  -0.20135921 0.095041639 0.29640085
## 24  -0.31355996 0.190888685 0.50444864
## 25  -0.35313445 0.339488076 0.69262252
```

   To collect data for Figure 2(g), we downloaded the Excel file containing the source data. We extracted only the relevant columns, namely `closer_vals` and `further_vals`, corresponding to 2(g), and saved them into a new CSV file titled `data.csv` for analysis. After loading the data into `R`, we used `mutate()` to create a new column with the calculated difference between the two conditions (`closer_vals - further_vals`), allowing us to examine how dopamine signals changed depending on vocal similarity.

3. Summarize the data.

   (a) Summarize the further data. Do the data suggest that dopamine in the brains of young zebra finches decreases when they sing further away?

```
summarize.further <- dat.clean |>
  summarize(
    mean = mean(further_vals, na.rm=T),
    sd = sd(further_vals, na.rm=T)
  )

(summarize.further)

##         mean        sd
## 1 -0.2027244 0.1303193
```

   The further data show a mean percent change in fluorescence of -0.20 with a standard deviation of 0.13, suggesting that dopamine levels tend to decrease when zebra finches sing notes that are further from their adult song. This negative mean fluorescence change aligns with the hypothesis that imprecise vocalizations lead to lower dopamine levels.

   (b) Summarize the closer data. Do the data suggest that dopamine in the brains of young zebra finches increases when they sing closer to their adult song?

```
summarize.closer <- dat.clean |>
  summarize(
    mean = mean(closer_vals, na.rm=T),
    sd = sd(closer_vals, na.rm=T)
  )

(summarize.closer)

##        mean         sd
## 1 0.1562231 0.09408297
```

   The closer data have a mean change in fluorescence of 0.16 with a standard deviation of 0.09, indicating a consistent increase in dopamine when zebra finches produce more accurate vocalizations. This relatively low standard deviation also suggests this effect is robust across individuals.

   (c) Summarize the paired differences. Do the data suggest that there is a difference between dopamine in the brains of young zebra finches when they sing further away compared to closer to their adult song?

```
summarize.differences <- dat.clean |>
  summarize(
    mean = mean(difference, na.rm=T),
    sd = sd(difference, na.rm=T)
  )
(summarize.differences)

##        mean        sd
## 1 0.3589475 0.2108744
```

   The paired differences between closer and further conditions show a mean difference of 0.36 and a standard deviation of 0.21. The magnitude of this effect (0.36) relative to the variability (0.21) suggests a strong and meaningful difference between the two conditions, potentially supporting the theory that learning to sing is shaped by dopamine-driven reinforcement learning.

   The boxplot below in Figure 1 visually reinforces this trend, clearly showing that fluorescence changes are consistently positive in the closer condition and consistently negative in the further condition. The paired differences stand out as they are far above zero, capturing overall increase in dopamine signaling.

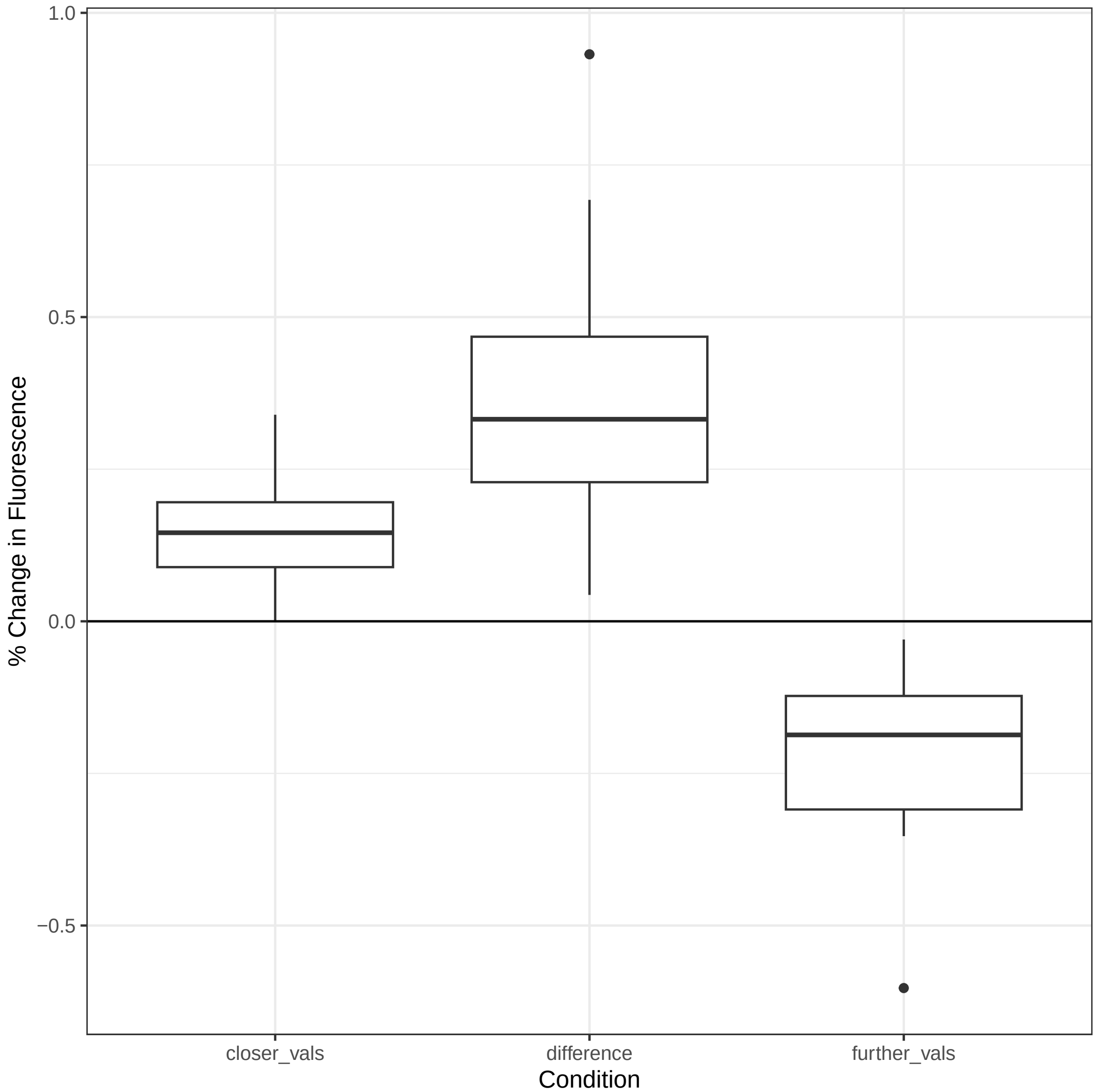

Figure 1: % Change in Flourescence Across Conditions

(d) **Optional Challenge:** Can you reproduce Figure 2(g)? Note that the you can use `geom_errorbar()` to plot the range created by adding the mean $\pm$ one standard deviation.

4. Conduct the inferences they do in the paper. Make sure to report the results a little more comprehensively – that is your parenthetical should look something like: ($t = 23.99$, $p < 0.0001$; $g = 1.34$; 95% CI: 4.43, 4.60).
Note: Your numbers may vary slightly as they performed some unclear correction of their p-values. I'm waiting to hear back from them via email!

(a) "The close responses differed significantly from 0 ($p = 1.63 \times 10^{-8}$)."

```
library(effectsize)
```

```r
# close
hedges_result_closer <- hedges_g(x = dat.clean$closer_vals, mu = 0, alternative = "greater")

help("hedges_g")

t_test_result_closer <- t.test(x=dat.clean$closer_vals, mu = 0, alternative = "greater")
t_test_result_closer.ci <- t.test(x=dat.clean$closer_vals, mu = 0, alternative = "two.sided")

t_stat_closer <- t_test_result_closer$statistic
p_val_closer <- t_test_result_closer$p.value
hedges_g_closer <- hedges_result_closer$Hedges_g
ci_low_closer <- t_test_result_closer.ci$conf.int[1]
ci_high_closer <- t_test_result_closer.ci$conf.int[2]

results.closer <- data.frame(
  t = c(t_stat_closer),
  p_value = c(p_val_closer),
  g = c(hedges_g_closer),
  CI_Lower = c(ci_low_closer),
  CI_Upper = c(ci_high_closer)
)

library(xtable)
results.closer.xtable <- xtable(results.closer,
                                caption="Closer Inferences")
```

t	p_value	g	CI_Lower	CI_Upper
8.30	0.00	1.61	0.12	0.20

Table 1: Closer Inferences

The close responses differed significantly from zero ($t = 8.30, p < 0.0001; g = 1.61; CI : 0.12, 0.20$). This provides statistically discernable support that the mean percent change in fluorescence in response to close sounds is greater than zero, indicating increased dopamine activity when a zebra finch sings in a way that more closely resembles the adult song. The large effect size ($g = 1.61$) reflects a substantial difference in dopamine response, consistent with a strong reinforcement signal.

(b) "The far responses differed significantly from 0 ($p = 5.17 \times 10^{-8}$)."

```r
library(effectsize)

hedges_result_further <- hedges_g(x = dat.clean$further_vals, mu = 0, alternative = "less")

t_test_result_further <- t.test(x=dat.clean$further_vals, mu = 0, alternative = "less")
t_test_result_further.ci <- t.test(x=dat.clean$further_vals, mu = 0, alternative = "two.sided")

t_stat_further <- t_test_result_further$statistic
p_val_further <- t_test_result_further$p.value
hedges_g_further <- hedges_result_further$Hedges_g
ci_low_further <- t_test_result_further.ci$conf.int[1]
ci_high_further <- t_test_result_further.ci$conf.int[2]

results.further <- data.frame(
```

```r
t = c(t_stat_further),
p_value = c(p_val_further),
g = c(hedges_g_further),
CI_Lower = c(ci_low_further),
CI_Upper = c(ci_high_further)
)

library(xtable)
results.further.xtable <- xtable(results.further,
                                caption="Further Inferences")
```

t	p_value	g	CI_Lower	CI_Upper
-7.78	0.00	-1.51	-0.26	-0.15

Table 2: Further Inferences

The far responses were significantly less than zero ($t = -7.78, p < 0.0001; g = -1.51; CI : -0.26, -0.15$). This provides statistically discernable support that the mean percent change in fluorescence in response to far sounds is less than zero, meaning dopamine activity tends to decrease when the bird sings songs further from the target sound. The large negative effect size ($g = -1.51$) highlights that this difference is large, indicating the idea that less successful vocal attempts are met with dopamine suppression.

(c) "The difference between populations was significant ($p = 1.04 \times 10^{-8}$)."

```r
# differences
hedges_result_difference <- hedges_g(x = dat.clean$difference, mu = 0, alternative = "two.sided

t_test_result_difference <- t.test(x=dat.clean$difference, mu = 0, alternative = "two.sided")

t_stat_difference <- t_test_result_difference$statistic
p_val_difference <- t_test_result_difference$p.value
hedges_g_difference <- hedges_result_difference$Hedges_g
ci_low_difference <- t_test_result_difference$conf.int[1]
ci_high_difference <- t_test_result_difference$conf.int[2]

results.difference <- data.frame(
t = c(t_stat_difference),
p_value = c(p_val_difference),
g = c(hedges_g_difference),
CI_Lower = c(ci_low_difference),
CI_Upper = c(ci_high_difference)
)

library(xtable)
results.difference.xtable <- xtable(results.difference,
                                   caption="Difference Inferences")
```

t	p_value	g	CI_Lower	CI_Upper
8.51	0.00	1.65	0.27	0.45

Table 3: Difference Inferences

The difference between populations was also significant ($t = 8.51, p < 0.0001; g = 1.65; CI : 0.27, 0.45$). This indicates a statistically discernable difference in dopamine-associated fluorescence between close and far vocalizations, with stronger responses for more accurate song attempts. The large effect size again underscores that dopamine encodes the degree of vocal accuracy, supporting the interpretation that zebra finches learn to sing via a reinforcement mechanism.

5. Reverse engineer the hypothesis test plot from Lecture 20 to create accurate hypothesis testing plots for each part of the previous question.

   (a) Question 4, part(a).

```r
# part a: question 4, part(a)
mu0 <- 0
x <- dat.clean$closer_vals
xbar <- mean(x)
s <- sd(x)
n <- length(x)
any(is.na(x)) # no missing data

## [1] FALSE

t.stat <- (xbar - mu0)/(s/sqrt(n))


# For plotting the null distribution
ggdat.t <- tibble(t=seq(-5,5,length.out=1000))|>
  mutate(pdf.null = dt(t, df=n-1))
# For plotting the observed point
ggdat.obs <- tibble(t    = t.stat,
                    y    = 0) # to plot on x-axis

# Resampling to approximate the sampling distribution
# on the data
R <- 1000
resamples <- tibble(t=numeric(R))
for(i in 1:R){
  curr.sample <- sample(x=x,
                        size=n,
                        replace=T)
  resamples$t[i] = (mean(curr.sample)-mu0)/(sd(curr.sample)/sqrt(n))
}

t.breaks <- c(-5, 0, qt(0.95, df = n-1), 5, t.stat)
xbar.breaks <- t.breaks * s/(sqrt(n)) + mu0

# Create Plot
close.hypothesis.plot <- ggplot() +
  # null distribution
  geom_line(data=ggdat.t,
            aes(x=t, y=pdf.null))+
  geom_hline(yintercept=0)+
  # rejection regions
  geom_ribbon(data=subset(ggdat.t, t<=qt(0.95, df=n-1)),
              aes(x=t, ymin=0, ymax=pdf.null),
              fill="grey", alpha=0.5)+
  # plot p-value (not visible)
  geom_ribbon(data=subset(ggdat.t, t>=t.stat),
              aes(x=t, ymin=0, ymax=pdf.null),
              fill="grey", alpha=0.25)+
  # plot observation point
  geom_point(data=ggdat.obs, aes(x=t, y=y), color="red")+
  # Resampling Distribution
  stat_density(data=resamples,
               aes(x=t),
               geom="line", color="grey")+
  # clean up aesthetics
  theme_bw()+
  scale_x_continuous("t",
                     breaks = round(t.breaks,2),
                     sec.axis = sec_axis(~.,
                                         name = bquote(bar(x)),
                                         breaks = t.breaks,
                                         labels = round(xbar.breaks,2)))+
  ylab("Density")+
  ggtitle("T-Test for Closer Responses",
          subtitle=bquote(H[0]==0*";"~H[a]>0))
```

Figure 2 below shows the null distribution for the t-statistic under the hypothesis that the mean percent change in fluorescence in response to close sounds is zero. The null hypothesis, $H_0 : \mu = 0$, represents no average dopamine response when the zebra finch sings in a way that closely matches the adult song. The alternative hypothesis, $H_a : \mu > 0$, reflects the researchers prediction that dopamine levels increase in response to these more accurate vocalizations. The unshaded shaded region in the right tail marks the 5% rejection region, while the red point marks the observed t-statistic, $t = 8.30$, which lies far into the right tail. The p value is effectively 0, providing statistically discernable support that the mean percent change in fluorescence is greater than zero. This confirms that singing closer to the adult song triggers a significant dopamine increase. The resampling distribution, overlaid in light grey, confirms the right-skewed location of the observed statistic relative to the null.

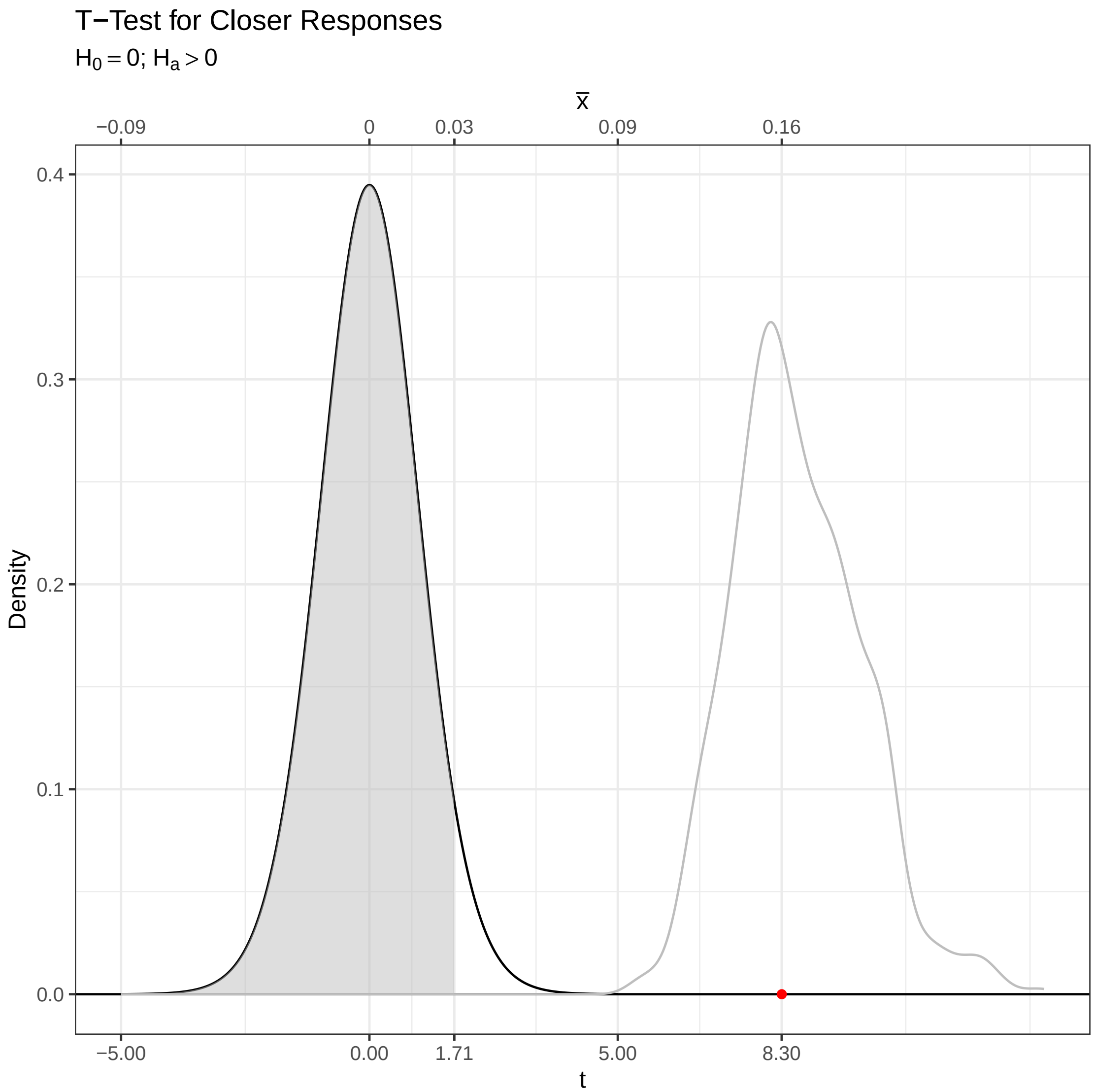

Figure 2: Hypothesis Test Plot for Close Responses

(b) Question 4, part(b).

Figure 3 tests whether dopamine levels significantly decrease when zebra finches sing in a way that diverges from the adult song. The hypotheses are $H_0 : \mu = 0$ and $H_a : \mu < 0$. Here, the observed t-statistic is $t = -7.78$, marked by the red point on the left tail of the null distribution. The rejection region (grey) and p-value area are both shaded to emphasize the extremity of the observed statistic. The p-value is again extremely small as you can see by where the red dot is located in relation to the tail. This result is consistent with a strong decrease in dopamine activity when vocalizations are less accurate, reinforcing the idea of negative reinforcement. The resampling distribution confirms the leftward shift and supports the conclusion that the observed t-statistic is unlikely under the null.

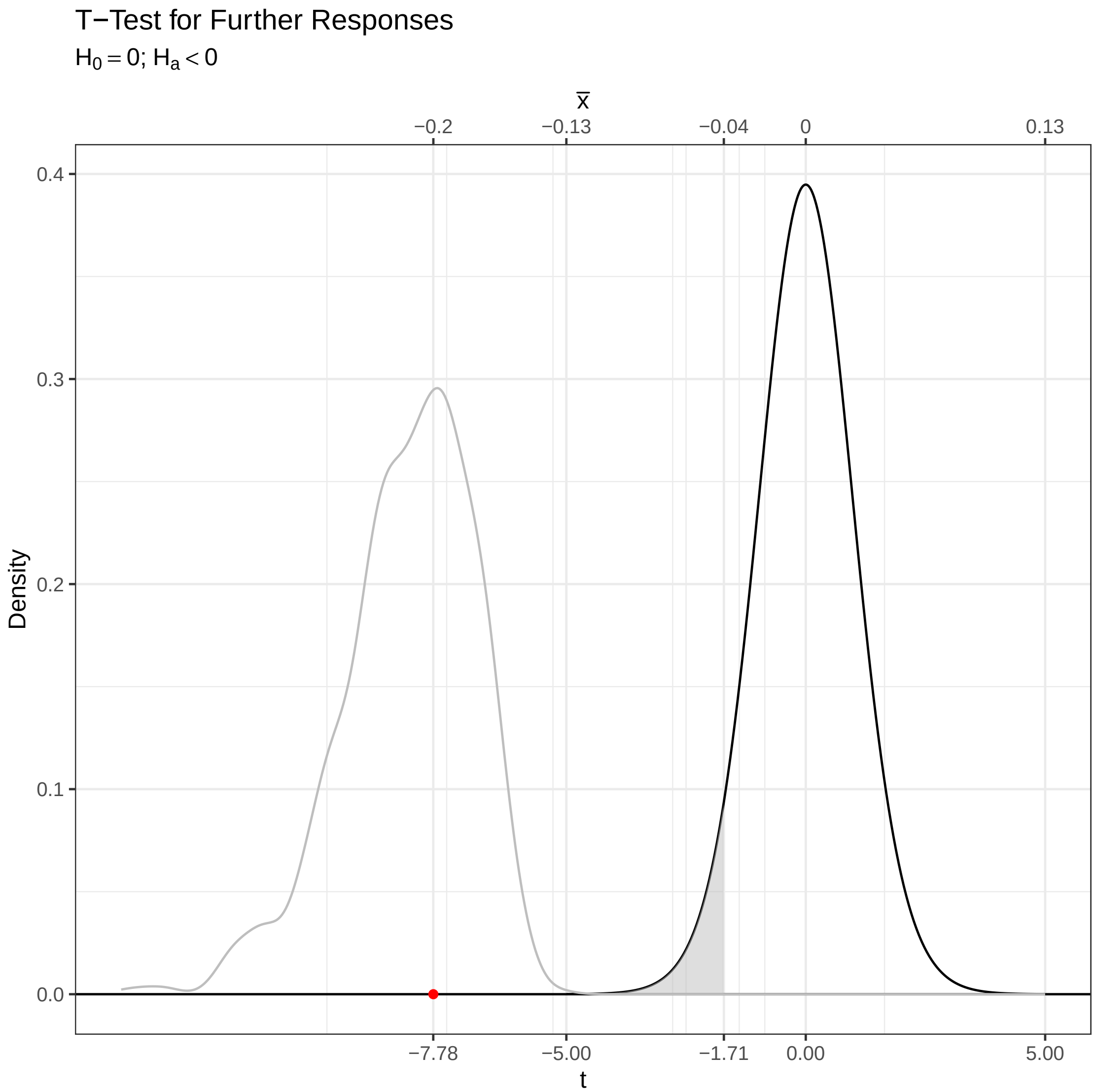

Figure 3: Hypothesis Test Plot for Far Responses

(c) Question 4, part(c).

Figure 4 compares dopamine responses directly between close and far vocalizations, testing the hypothesis $H_0 : \mu = 0$ versus $H_a : \mu \neq 0$. The null distribution assumes no difference in percent fluorescence change between the two conditions. The observed t-statistic falls significantly to the right, $t = 8.51$, indicating that dopamine responses are markedly stronger when the zebra finches produce closer approximations to the adult song. The p-value is again essentially zero, supporting a statistically discernable difference between close and far responses. This plot emphasizes that dopamine not only increases during accurate vocalizations but does so significantly more than during inaccurate ones– further reinforcing its role as a reinforcement signal during song learning.

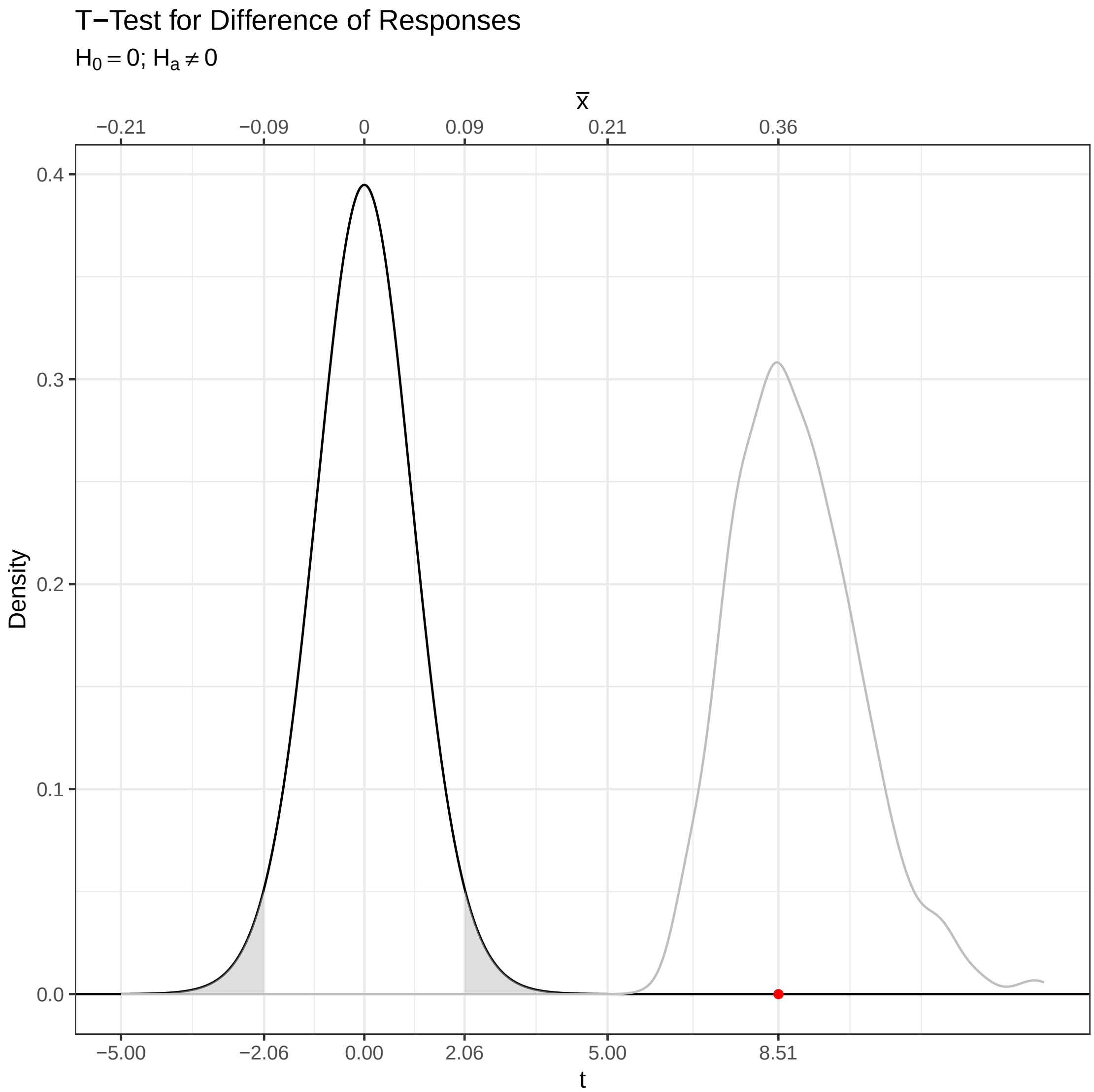

Figure 4: Hypothesis Test Plot for Difference Between Populations

References

Champely, S. (2020). *pwr: Basic Functions for Power Analysis*. R package version 1.3-0.

Kasdin, J., Duffy, A., Nadler, N., Raha, A., Fairhall, A. L., Stachenfeld, K. L., and Gadagkar, V. (2025). Natural behaviour is learned through dopamine-mediated reinforcement. *Nature*, pages 1–8.