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.

```
#conduct power analysis
d <- 0.65 #effect size
type <- "one.sample"
alternative <- "two.sided"
alpha <- 0.05
power <- 0.80
power.analysis <- pwr.t.test(n = NULL, d=d, sig.level = alpha, power = power,
                             type = type, alternative = alternative)

#extract n argument and get integer value (rounded up)
num.obseravtions <- ceiling(power.analysis$n)
num.obseravtions


## [1] 21
```

   The researchers would need at least 21 observations 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.

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`).

```
#load the file with deleted sheets
fig.data <- read_csv("Fig2Dat.csv")

#mutate data to get the difference
fig.data <- fig.data|>
  mutate("difference" = closer_vals - further_vals)
```

   To collect the data for Figure 2(g), we need to download the source data for Figure 2. Then, we delete all sheets except those containing `closer_vals` and `further_vals`. We then combine these two sheets into one to contain `closer_vals` and `further_vals` side-by-side. Finally, we add new column of difference to the sheet by subtracting `further_vals` from `closer_vals`.

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?

```
further.dat <- fig.data$further_vals

#do numerical summary for the data
further.summary <- tibble(
  mean = mean(further.dat),
  sd = sd(further.dat),
  min = min(further.dat),
  q1 = quantile(further.dat, probs = 0.25),
  median = quantile(further.dat, probs = 0.50),
  q3 = quantile(further.dat, probs = 0.75),
  max = max(further.dat)
)

#do graphical summary for the data
furthest.boxplot <- ggplot(data = tibble(further.dat))+
```

Table 1: Numerical Summary of Dopamine Levels (Further Data)

mean	sd	min	q1	median	q3	max
-0.2027244	0.1303193	-0.6027859	-0.3092606	-0.1867461	-0.1227054	-0.0299347

```
  geom_boxplot(aes(x = "", y = further.dat*100),
               fill = "lightblue")+ #make the boxplot for the data
  theme_bw()+
  xlab("Further Data")+
  ylab("Dopanime percentage change")
```

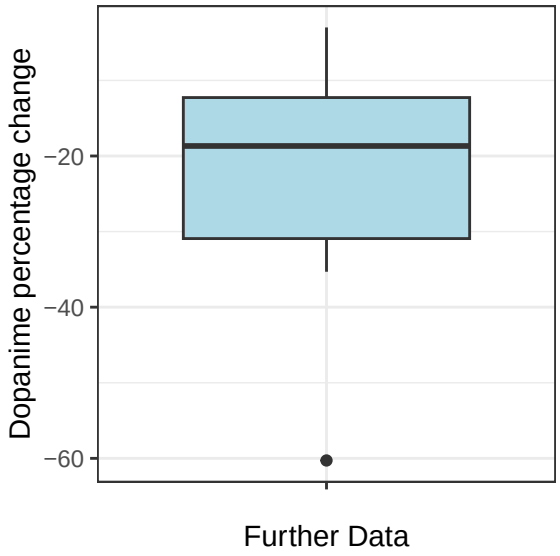

Figure 1: Boxplot showing percentage change in dopamine levels when zebra finches sing further away.

The data suggest that the dopamine in the brains of young zebra finches decreases when they sing further away. The numerical summary reflect the drop in the dopamine since all key measures of central tendency are negative. The maximum value is also negative. The entire boxplot lies below 0% and there are no positive outliers.

(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?

```
closer.dat <- fig.data$closer_vals

#do numerical summary for the data
closer.summary <- tibble(
  mean = mean(closer.dat),
  sd = sd(closer.dat),
  min = min(closer.dat),
  q1 = quantile(closer.dat, probs = 0.25),
  median = quantile(closer.dat, probs = 0.50),
  q3 = quantile(closer.dat, probs = 0.75),
  max = max(closer.dat)
)

#do graphical summary for the data
closer.boxplot <- ggplot(data = tibble(closer.dat))+
  geom_boxplot(aes(x = "", y = closer.dat*100),
               fill = "lightblue")+ #make the boxplot for the data
  theme_bw()+
  xlab("Closer Data")+
  ylab("Dopanime percentage change")
```

Table 2: Numerical Summary of Dopamine Levels (Closer Data)

mean	sd	min	q1	median	q3	max
0.1562231	0.094083	0.0010821	0.089018	0.1455341	0.19567	0.3394881

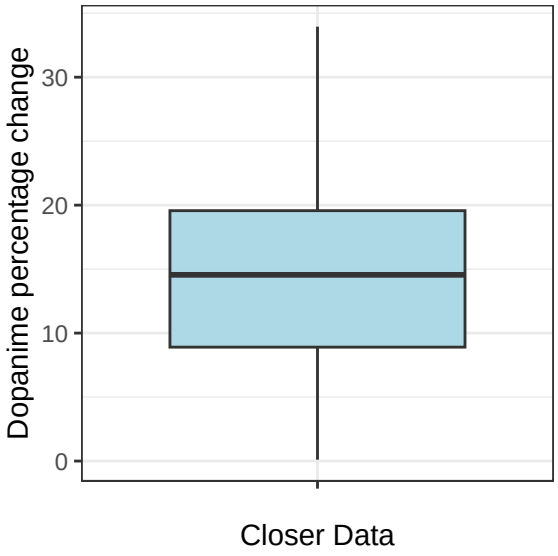

Figure 2: Boxplot showing percentage change in dopamine levels when zebra finches sing closer.

The data suggest that dopamine in the brains of young zebra finches increases when they sing closer to their adult song. The values in numerical summaries are positive, indicating the increase in dopamine levels. The whole boxplot lies above 0, so the dopamine levels increases when zebra finches sing closer to their adult song.

(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?

```
diff.dat <- fig.data$difference

#do numerical summary for the data
diff.summary <- tibble(
mean = mean(diff.dat),
sd = sd(diff.dat),
min = min(diff.dat),
q1 = quantile(diff.dat, probs = 0.25),
median = quantile(diff.dat, probs = 0.50),
q3 = quantile(diff.dat, probs = 0.75),
max = max(diff.dat)
)

#do graphical summary for the data
diff.boxplot <- ggplot(data = tibble(diff.dat))+
geom_boxplot(aes(x = "", y = diff.dat*100),
             fill = "lightblue")+ #make the boxplot for the data
theme_bw()+
xlab("Paired difference")+
ylab("Dopanime percentage change")
```

Table 3: Numerical Summary of Dopamine Levels (Paired Difference)

mean	sd	min	q1	median	q3	max
0.3589475	0.2108744	0.043353	0.2286407	0.3320846	0.4677148	0.9318804

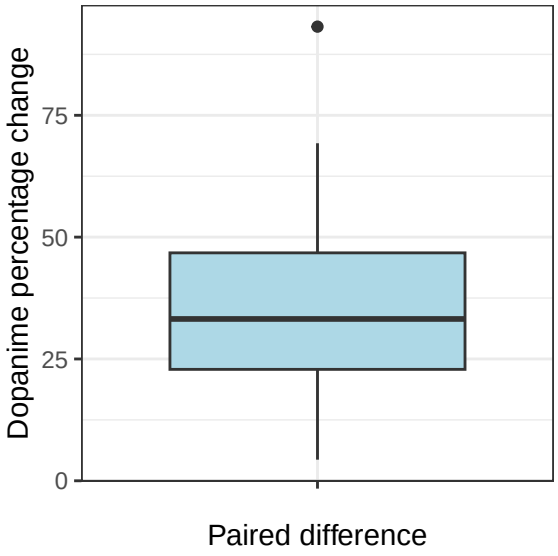

Figure 3: Boxplot showing percentage change in dopamine levels in paired difference of closer and further values.

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. The numerical summaries show that all paired differences are positive. This means that every bird had higher dopamine when it sang closer to its adult song compared to when it sang further away from its adult song. Also, the whole box plot lies above 0, which indicates dopamine increase for the closer conditions.

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 ($p = 1.63 \times 10^{-8}$)."

```
conf.level = 0.95
mu0 <- 0
#part a - conduct t-test for close responses
t.test.close <- t.test(x=closer.dat, mu = mu0, alternative = "greater")
#get the confidence interval using the two-sided test
t.test.close.interval <- t.test(x=closer.dat, mu = mu0, alternative = "two.sided")
conf.int.close <- t.test.close.interval$conf.int #get the confidence interval
conf.close.beg <- conf.int.close[1]
conf.close.end <- conf.int.close[2]
t.close <- t.test.close$statistic #get t
df.close <- t.test.close$parameter #get df
g.close <- hedges_g(x = closer.dat, mu = mu0, alternative = "greater") #get g
n.close <- t.test.close$parameter + 1 #get n
s.close <- t.test.close$stderr * sqrt(n.close) #get standard error
t.test.close <- t.test.close$p.value #get p-value
```

The close responses are statistically discernible from 0 ($t = 8.3$, $p < 0.0001$; $g = 1.61$; 95% CI: 0.12, 0.2).

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

```
t.test.far <- t.test(x=further.dat, mu = mu0, alternative = "less")
#get the confidence interval using the two-sided test
t.test.far.interval <- t.test(x=further.dat, mu = mu0, alternative = "two.sided")
conf.int.far <- t.test.far.interval$conf.int #get the confidence interval
conf.far.beg <- conf.int.far[1]
conf.far.end <- conf.int.far[2]
t.far <- t.test.far$statistic #get t
df.far <- t.test.far$parameter #get df
g.far <- hedges_g(x = further.dat, mu = mu0, alternative = "less") #get g
n.far <- t.test.far$parameter + 1 #get n
s.far <- t.test.far$stderr * sqrt(n.far) #get standard error
t.test.far <- t.test.far$p.value #get p-value
```

The far responses are statistically discernible from 0 ($t = -7.78$, $p < 0.0001$; $g = -1.51$; 95% CI: -0.26, -0.15).

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

```
t.test.diff <- t.test(x=diff.dat, mu = mu0, alternative = "two.sided")
#get the confidence interval using the two-sided test
t.test.diff.interval <- t.test(x=diff.dat, mu = mu0, alternative = "two.sided")
conf.int.diff <- t.test.diff.interval$conf.int #get the confidence interval
conf.diff.beg <- conf.int.diff[1]
conf.diff.end <- conf.int.diff[2]
t.diff <- t.test.diff$statistic #get t
df.diff <- t.test.diff$parameter #get df
g.diff <- hedges_g(x = diff.dat, mu = mu0, alternative = "two.sided") #get g
n.diff <- t.test.diff$parameter + 1 #get n
s.diff <- t.test.diff$stderr * sqrt(n.diff) #get standard error
t.test.diff <- t.test.diff$p.value #get p-value
```

The difference between populations are statistically discernible ($t = 8.51$, $p < 0.0001$; $g = 1.65$; 95% CI: 0.27, 0.45).

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).

```
# For plotting the null distribution
ggdat.t.close <- tibble(t=seq(-10,10,length.out=1000))|>
mutate(pdf.null = dt(x=t, df=df.close))
# For plotting the observed point
ggdat.obs.close <- tibble(t = t.close,
                          y = 0) # to plot on x-axis
t.breaks <- c(-5, qt(p = 1-0.05, df = df.close), # rejection region (left)
              0, 5, t.close)                     # t-statistic observed
t.breaks <- sort(unique(round(t.breaks, 2)))
xbar.breaks <- t.breaks * s.close/sqrt(n.close) + mu0

R <- 1000
resamples.close <- tibble(t=numeric(R))
for(i in 1:R){
curr.sample <- sample(x=closer.dat,
                      size=n.close,
                      replace=T)
resamples.close$t[i] = (mean(curr.sample)-mu0)/(sd(curr.sample)/sqrt(n.close))
}

#plot for part a - close responses
close.plot <- ggplot() +
# null distribution
geom_line(data=ggdat.t.close,
          aes(x=t, y=pdf.null, color = "Null Distribution"))+
# rejection regions
geom_ribbon(data=subset(ggdat.t.close, t>=qt(p = 1-0.05, df=df.close)),
            aes(x=t, ymin=0, ymax=pdf.null),
            fill="gray", alpha=0.5)+
# plot p-value (not visible)
geom_ribbon(data=subset(ggdat.t.close, t>=t.close),
            aes(x=t, ymin=0, ymax=pdf.null),
            fill="red", alpha=0.25)+
# Resampling Distribution
stat_density(data=resamples.close,
             aes(x=t, color="Resampling Distribution"),
```

```
            geom="line")+
theme_bw()+
ylab("Density")+
scale_color_manual(values = c("Null Distribution" = "black", "Resampling Distribution" = "grey"),
                   name = "") +
geom_hline(yintercept=0)+
# plot observation point
geom_point(data=ggdat.obs.close, aes(x=t, y=y), color="red")+
theme(legend.position = "bottom")+
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)))+
ggtitle("T-Test for Mean Dopamine Level of Close Responses",
        subtitle=bquote(H[0]: mu[X]==0*";"~H[a]: mu[X]>0))
```

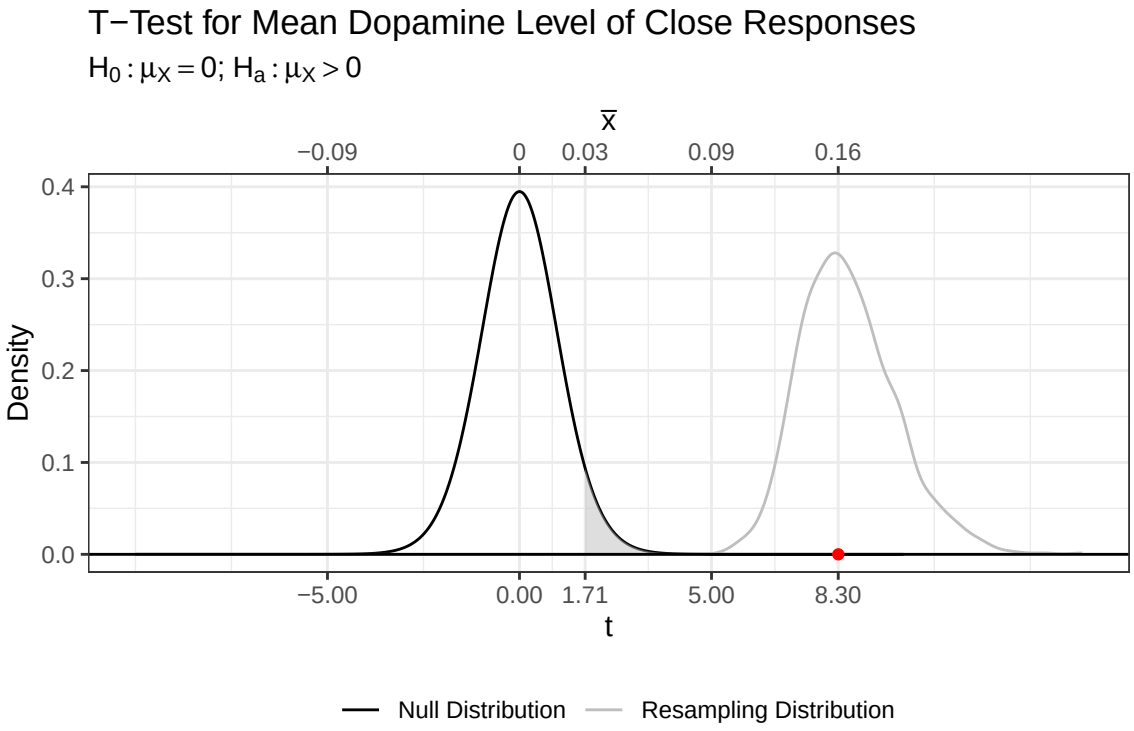

Figure 4: T-Test for Mean Dopamine Level of Difference in Close Responses for Zebra Finches.

(b) Question 4, part(b).

```
# For plotting the null distribution
ggdat.t.far <- tibble(t=seq(-10,10,length.out=1000))|>
mutate(pdf.null = dt(x=t, df=df.far))
# For plotting the observed point
ggdat.obs.far <- tibble(t = t.far,
                        y = 0) # to plot on x-axis
t.breaks <- c(t.far, -5, qt(p = 0.05, df = df.far), # rejection region (left)
              0, 5)                 # t-statistic observed
t.breaks <- sort(unique(round(t.breaks, 2)))
xbar.breaks <- t.breaks * s.far/sqrt(n.far) + mu0

resamples.far <- tibble(t=numeric(R))
for(i in 1:R){
curr.sample <- sample(x=further.dat,
                      size=n.far,
                      replace=T)
resamples.far$t[i] = (mean(curr.sample)-mu0)/(sd(curr.sample)/sqrt(n.far))
}
```

```r
#plot for part b - far responses
far.plot <- ggplot() +
# null distribution
geom_line(data=ggdat.t.far,
          aes(x=t, y=pdf.null, color = "Null Distribution"))+
# rejection regions
geom_ribbon(data=subset(ggdat.t.far, t<=qt(p = 0.05, df=df.far)),
            aes(x=t, ymin=0, ymax=pdf.null),
            fill="gray", alpha=0.5)+
# Resampling Distribution
stat_density(data=resamples.far,
             aes(x=t, color="Resampling Distribution"),
             geom="line")+
# plot observation point
geom_point(data=ggdat.obs.far, aes(x=t, y=y), color="red")+
geom_hline(yintercept=0)+
theme_bw()+
ylab("Density")+
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)))+
scale_color_manual(values = c("Null Distribution" = "black", "Resampling Distribution" = "grey"),
                   name = "") +
theme(legend.position = "bottom")+
ggtitle("T-Test for Mean Dopamine Level of Far Responses",
        subtitle=bquote(H[0]: mu[X]==0*";"~H[a]: mu[X] <0))
```

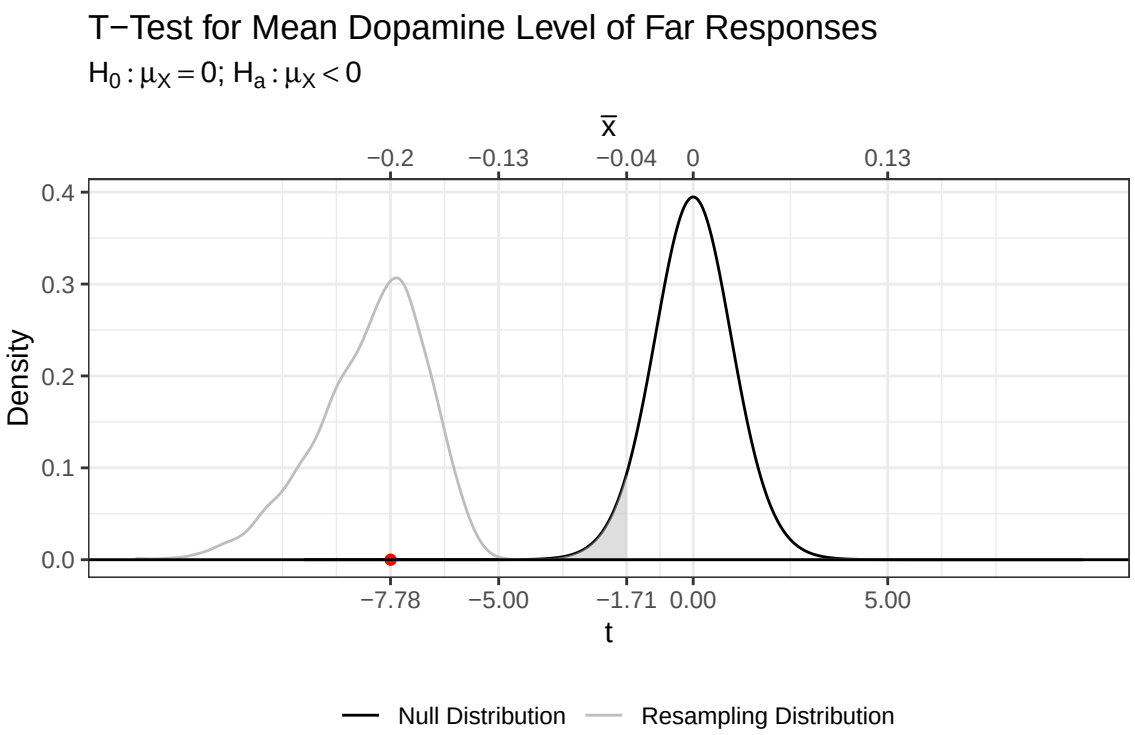

Figure 5: T-Test for Mean Dopamine Level of Far Responses for Zebra Finches.

(c) Question 4, part(c).

```r
# For plotting the null distribution
ggdat.t.diff <- tibble(t=seq(-10,10,length.out=1000))|>
mutate(pdf.null = dt(x=t, df=df.diff))
# For plotting the observed point
ggdat.obs.diff <- tibble(t = t.diff,
                         y = 0) # to plot on x-axis
```

```
t.breaks <- c(-5, qt(p = 0.025, df = df.diff), # rejection region (left)
              0, qt(p = 1-0.025, df = df.diff), 5, t.diff)                  # t-statistic observed
t.breaks <- sort(unique(round(t.breaks, 2)))
xbar.breaks <- t.breaks * s.diff/sqrt(n.diff) + mu0

resamples.diff <- tibble(t=numeric(R))
for(i in 1:R){
curr.sample <- sample(x=diff.dat,
                      size=n.diff,
                      replace=T)
resamples.diff$t[i] = (mean(curr.sample)-mu0)/(sd(curr.sample)/sqrt(n.diff))
}
#plot for part c - diff responses
diff.plot <- ggplot() +
# null distribution
geom_line(data=ggdat.t.diff,
          aes(x=t, y=pdf.null, color = "Null Distribution"))+
# rejection regions
geom_ribbon(data=subset(ggdat.t.diff, t<=qt(p = 0.025, df=df.diff)),
            aes(x=t, ymin=0, ymax=pdf.null),
            fill="gray", alpha=0.5)+
geom_ribbon(data=subset(ggdat.t.diff, t>=qt(0.975, df=df.diff)),
            aes(x=t, ymin=0, ymax=pdf.null),
            fill="grey", alpha=0.5)+
# Resampling Distribution
stat_density(data=resamples.diff,
             aes(x=t, color="Resampling Distribution"),
             geom="line")+
geom_hline(yintercept=0)+
# plot observation point
geom_point(data=ggdat.obs.diff, aes(x=t, y=y), color="red")+
theme_bw()+
ylab("Density")+
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)))+
scale_color_manual(values = c("Null Distribution" = "black", "Resampling Distribution" = "grey"),
                   name = "") +
theme(legend.position = "bottom")+
ggtitle("T-Test for Mean Dopamine Level of Difference in Responses",
        subtitle=bquote(H[0]: mu[X]==0*";"~H[a]: mu[X] != 0))
```

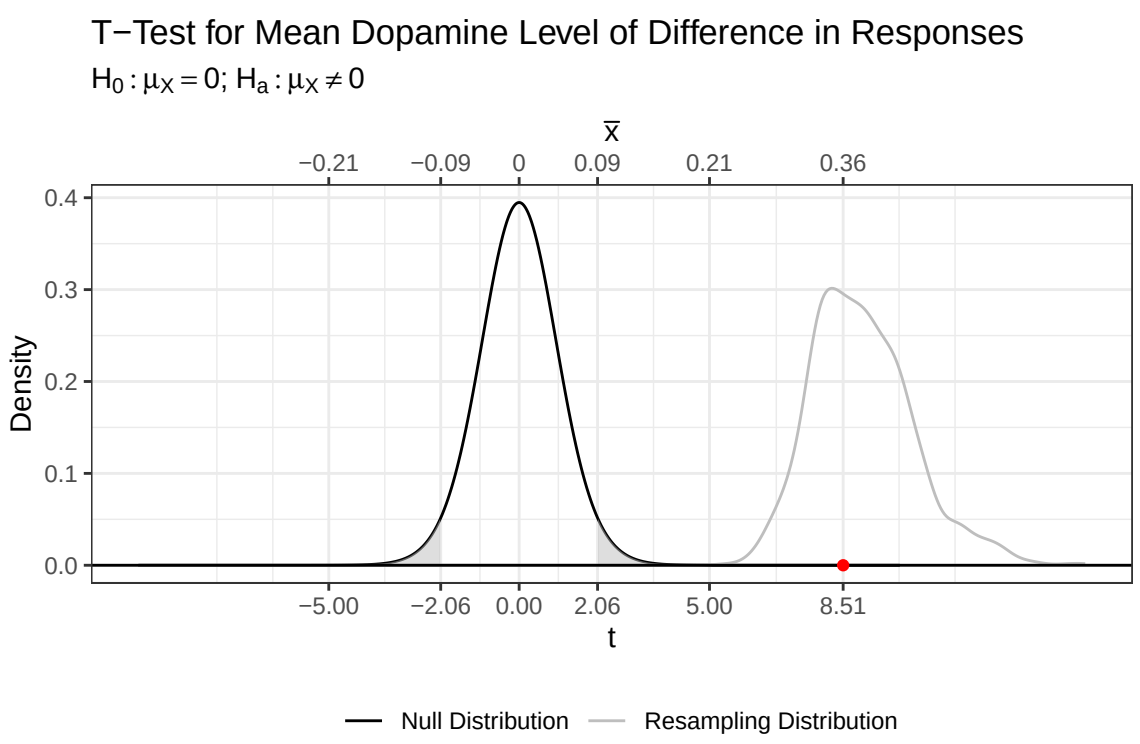

Figure 6: T-Test for Mean Dopamine Level of Difference in Responses for Zebra Finches.

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.