**Name: Classwork 11.2**

**What’s the difference between these two questions?**

* *Maybe the true population mean is 25. Then what is the likelihood of taking another sample and getting a mean > 26?*
* *Maybe the true population mean is 25. Then what is the likelihood of taking another sample and finding another individual > 26?*

1. Which situations needs a sampling distribution? Why?

**Sampling distribution of b0 versus b1**

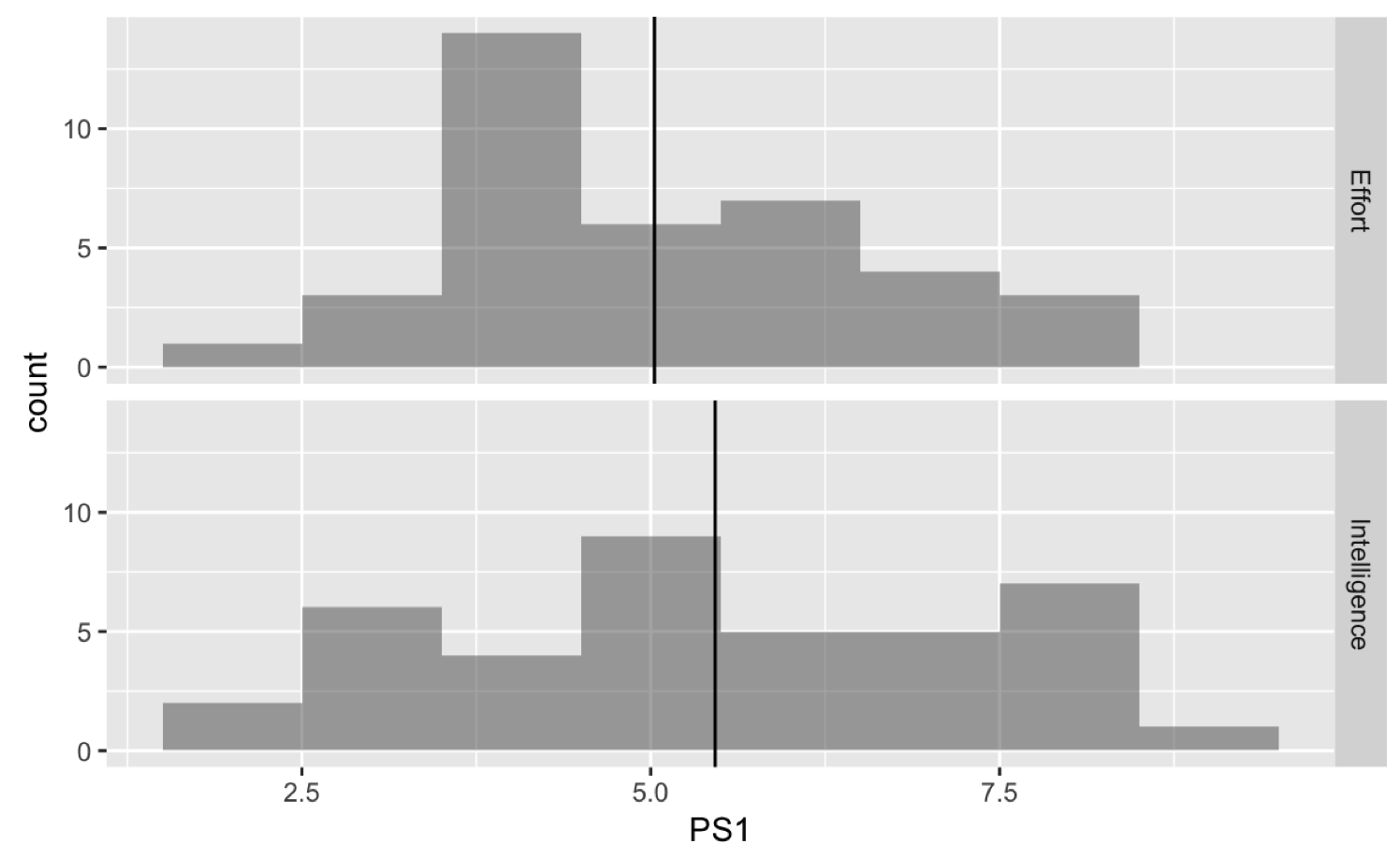
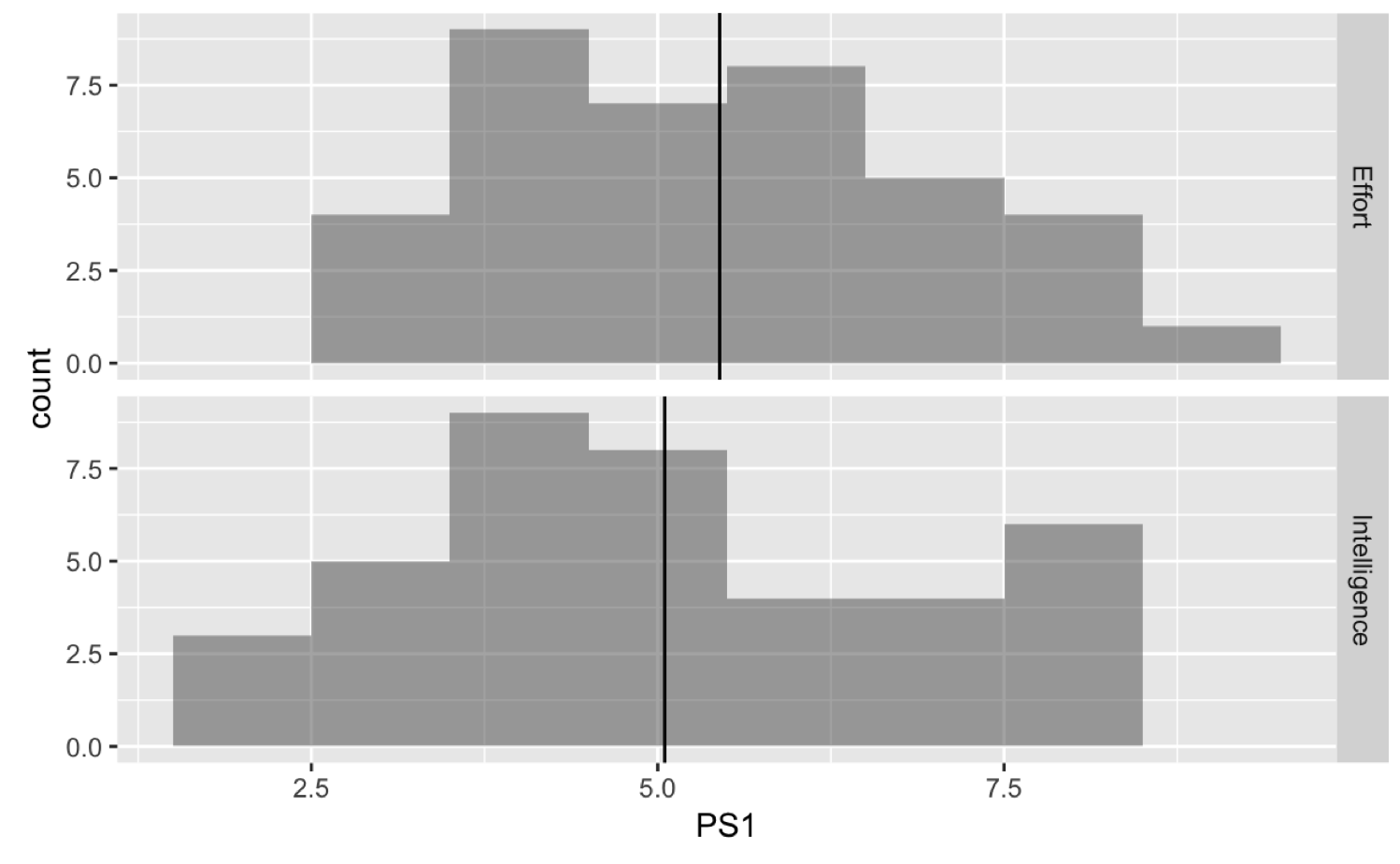
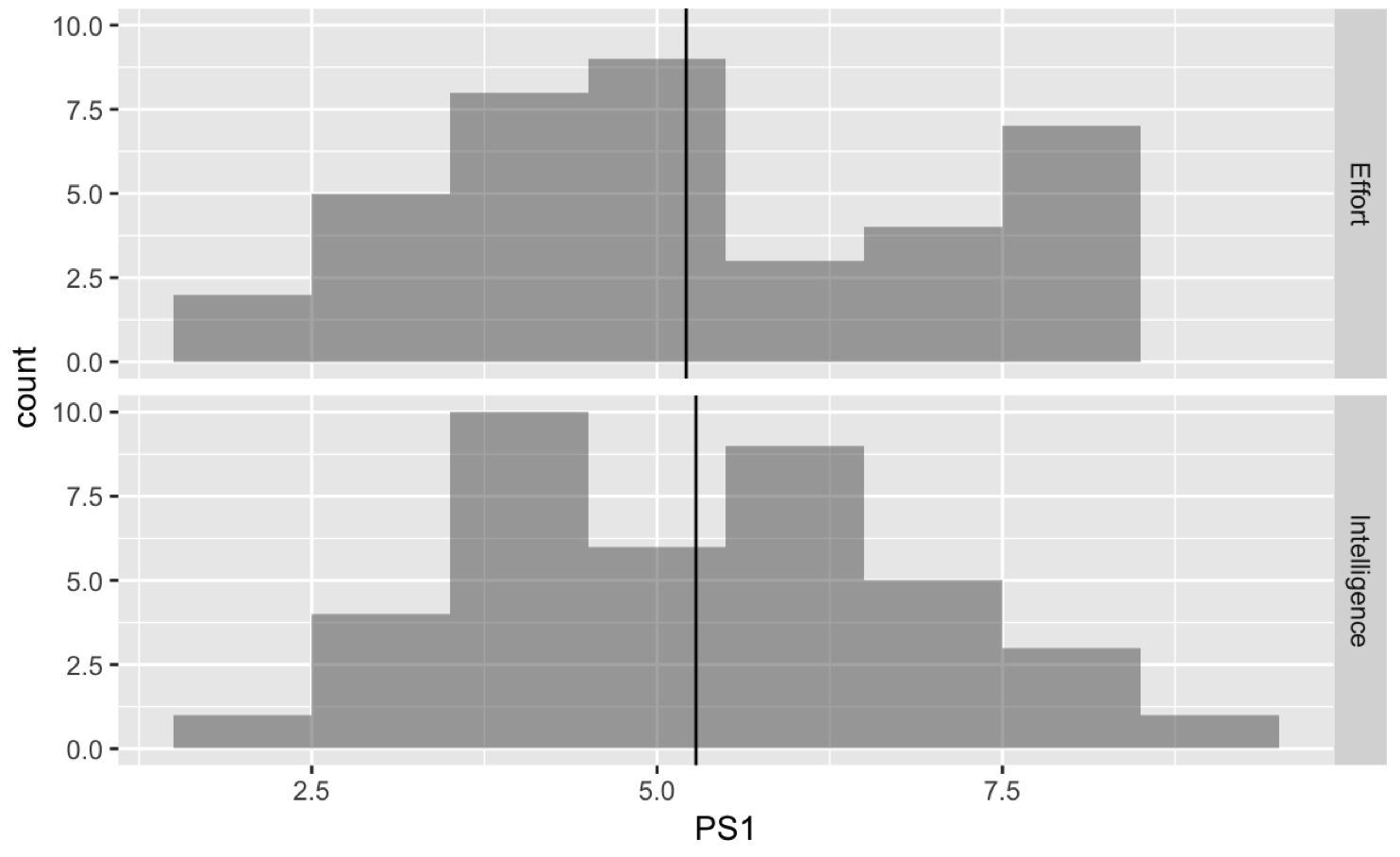
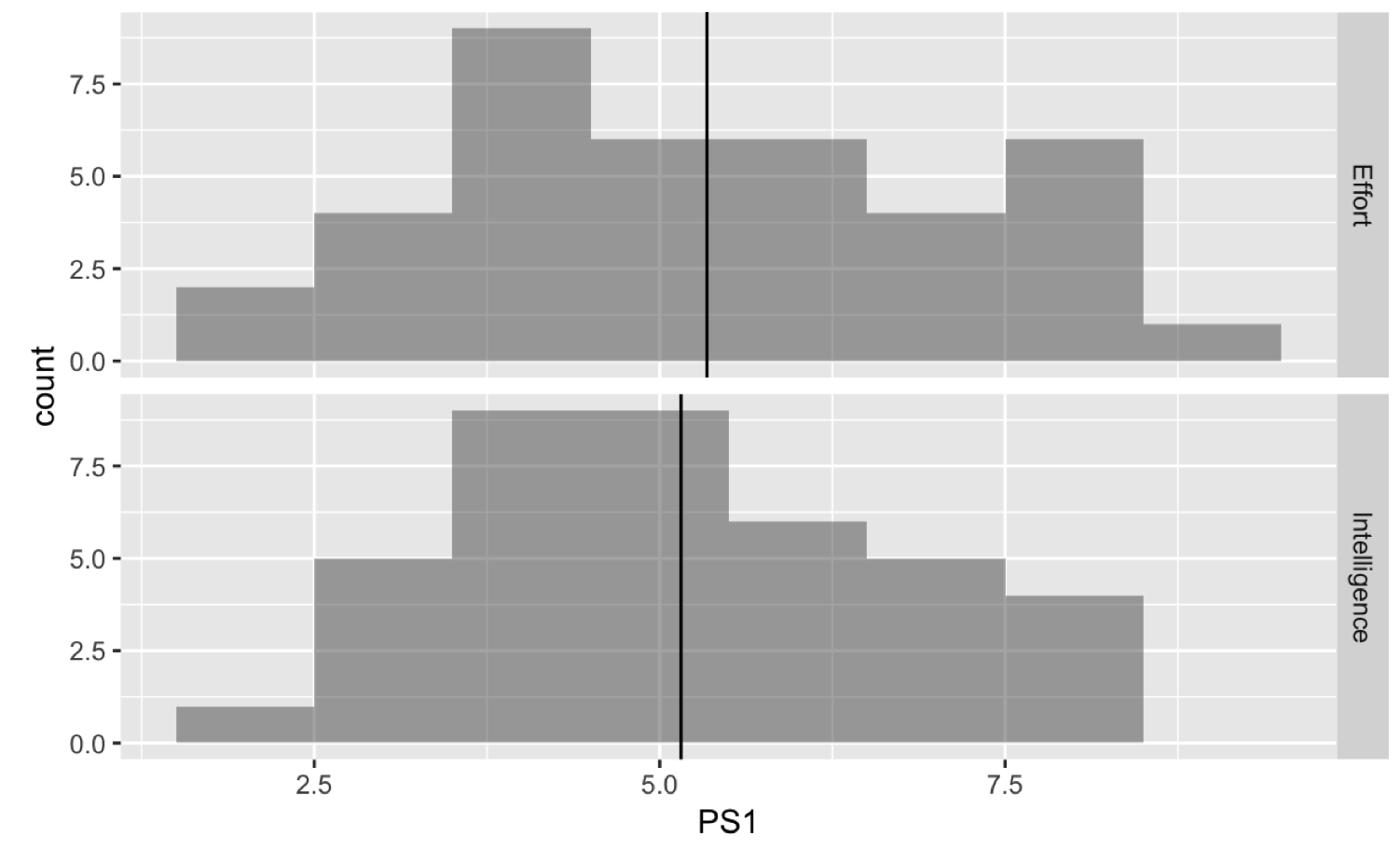
In week 5, we talked about the Mueller and Dweck study where they praised kids in different ways.

PRAISESTUDY <- read.csv("http://bit.ly/muellerdweck\_study1", header=TRUE)

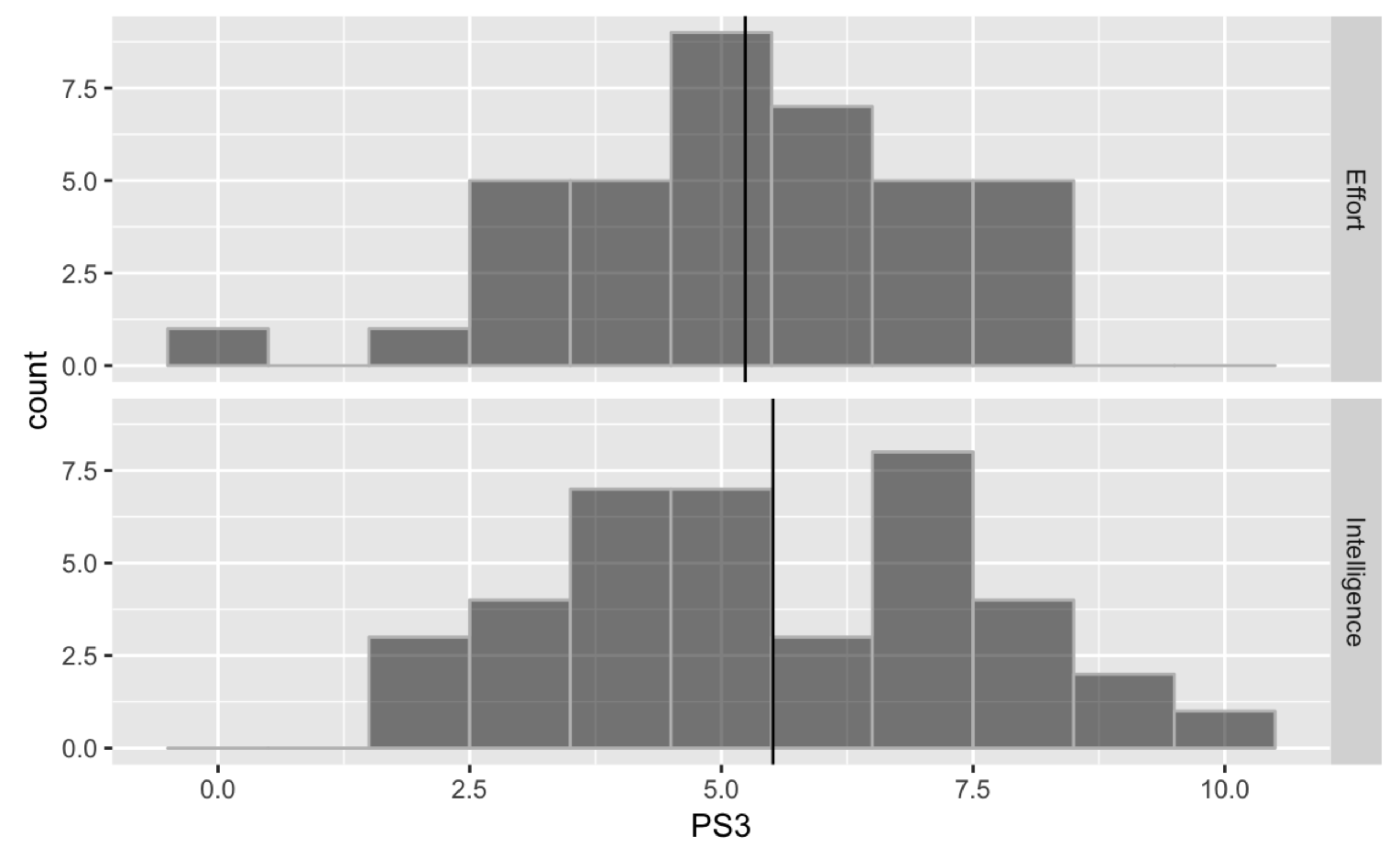
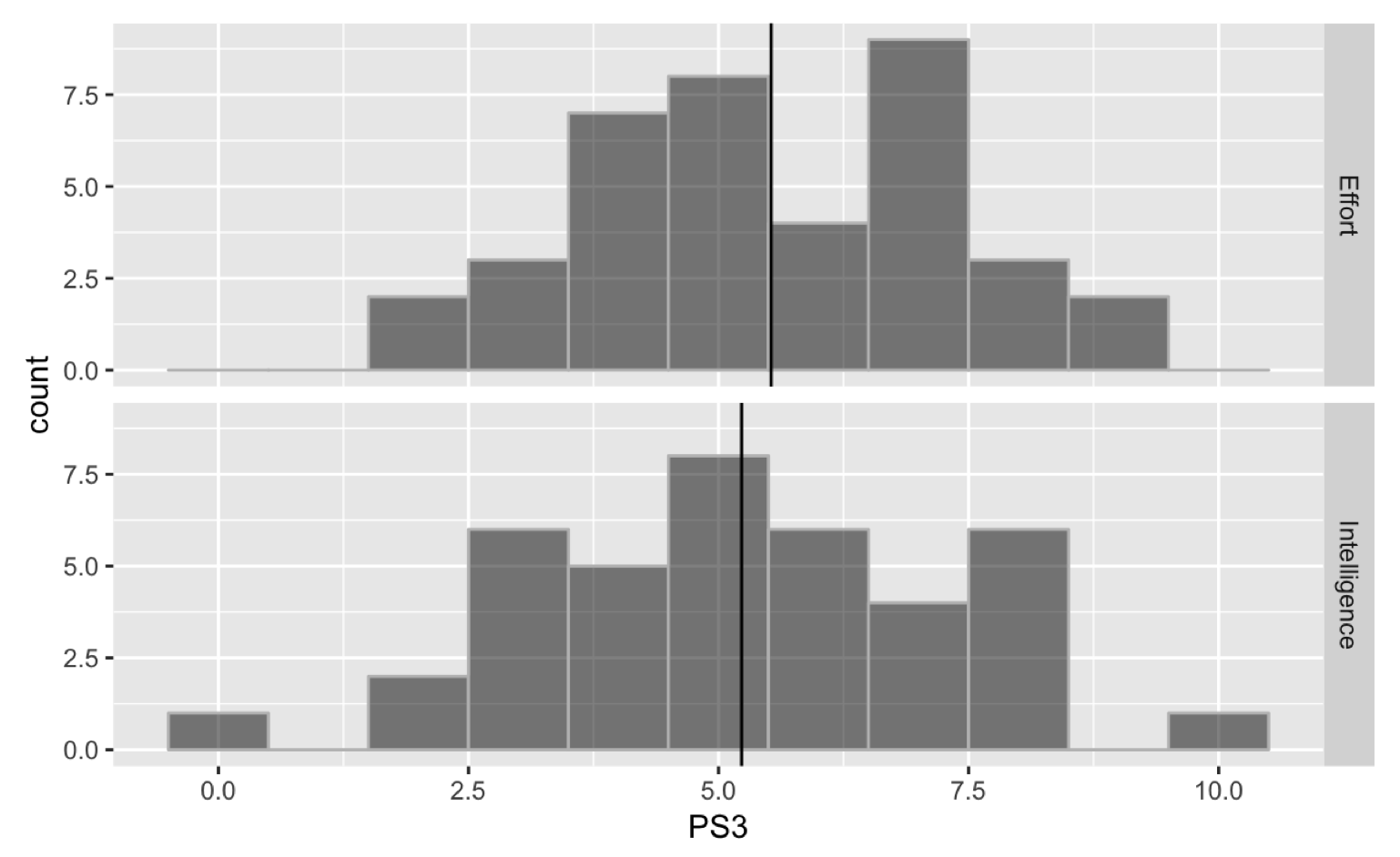
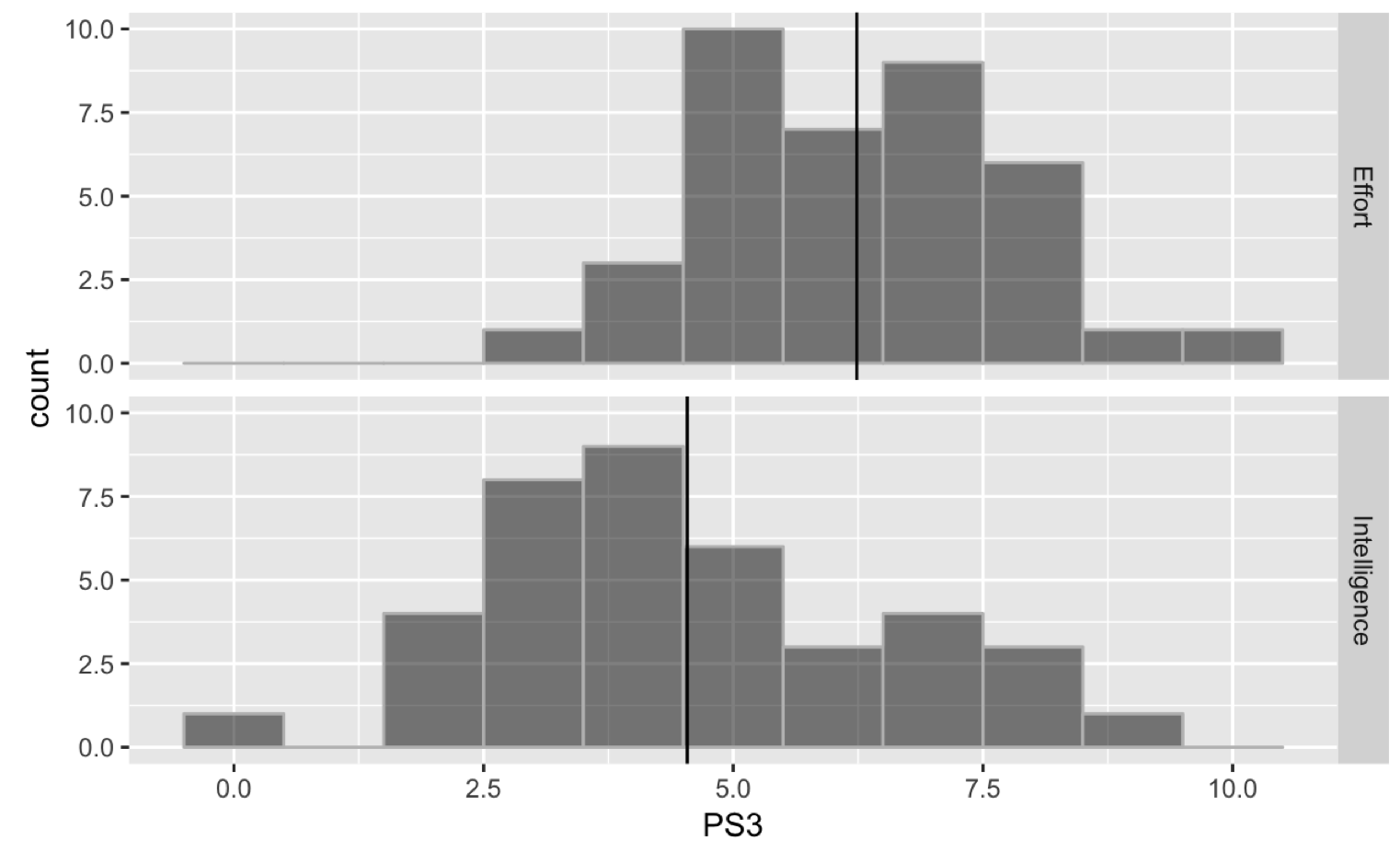
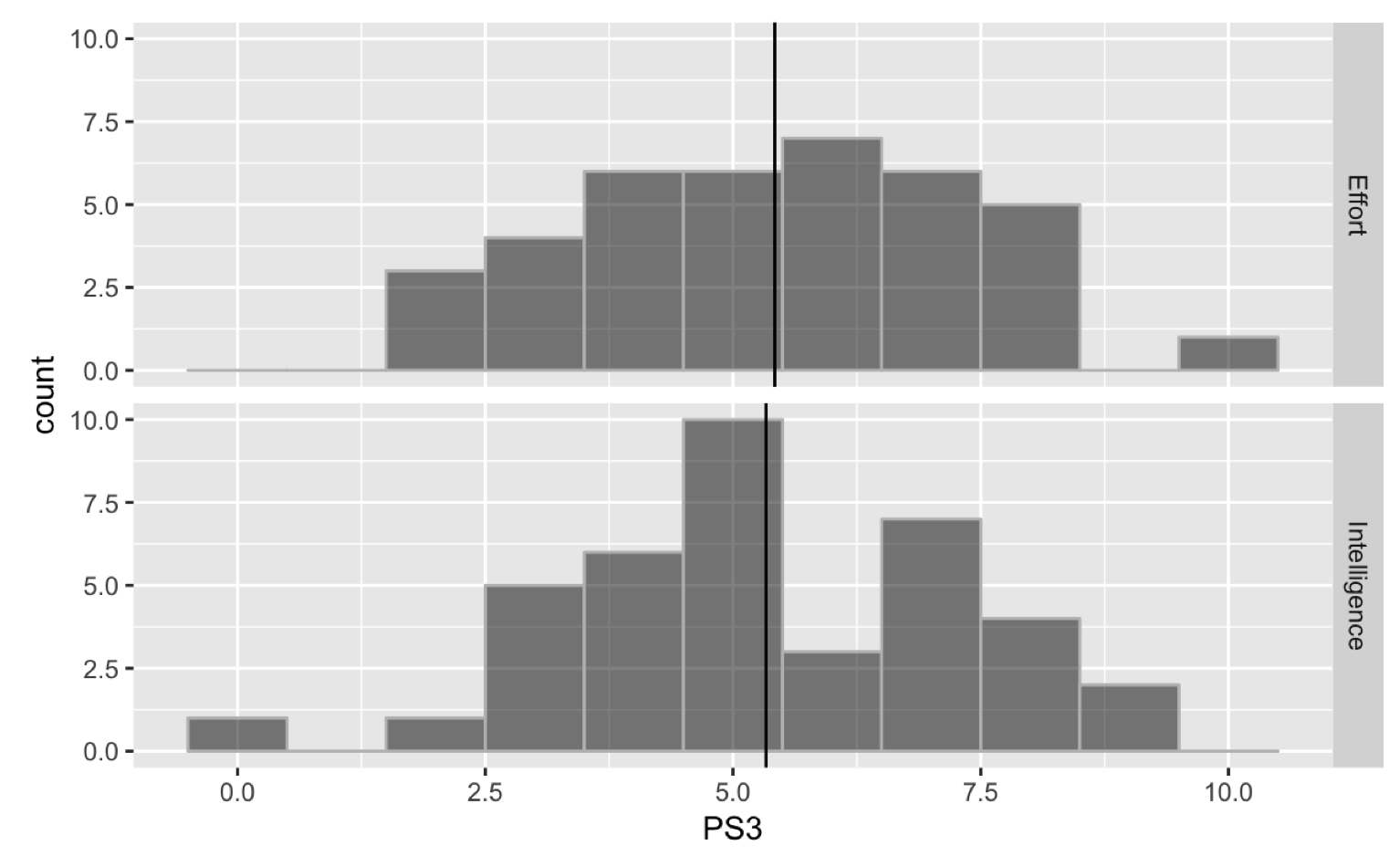
PRAISESTUDY <- filter(PRAISESTUDY, FEEDCODE != 2)

PRAISESTUDY$FEEDBACK <- factor(PRAISESTUDY$FEEDCODE, levels = c(1,0), labels = c("Effort", "Intelligence"))

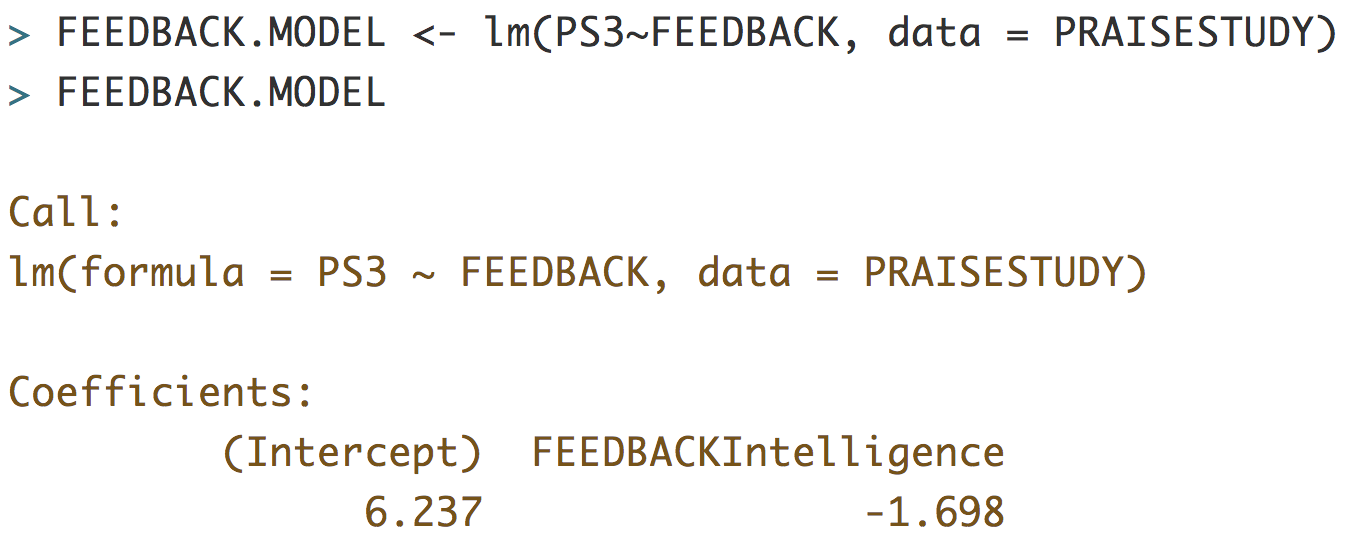
1. Remember, they did problem set #1 (PS1) before they praised the kids. So if the kids solve problems differently, could it logically be explained by the feedback they got?
2. Does that mean if we create a feedback model of PS1, will the PRE be 0? Why or why not?
3. If feedback doesn’t explain the group difference, why is there a difference between the two groups on their PS1 scores?
4. Remember we used **shuffle()** and made a bunch of shuffled histograms and tried to see whether our real sample could be detected from the shuffled samples? Can you pick out the “real” sample?

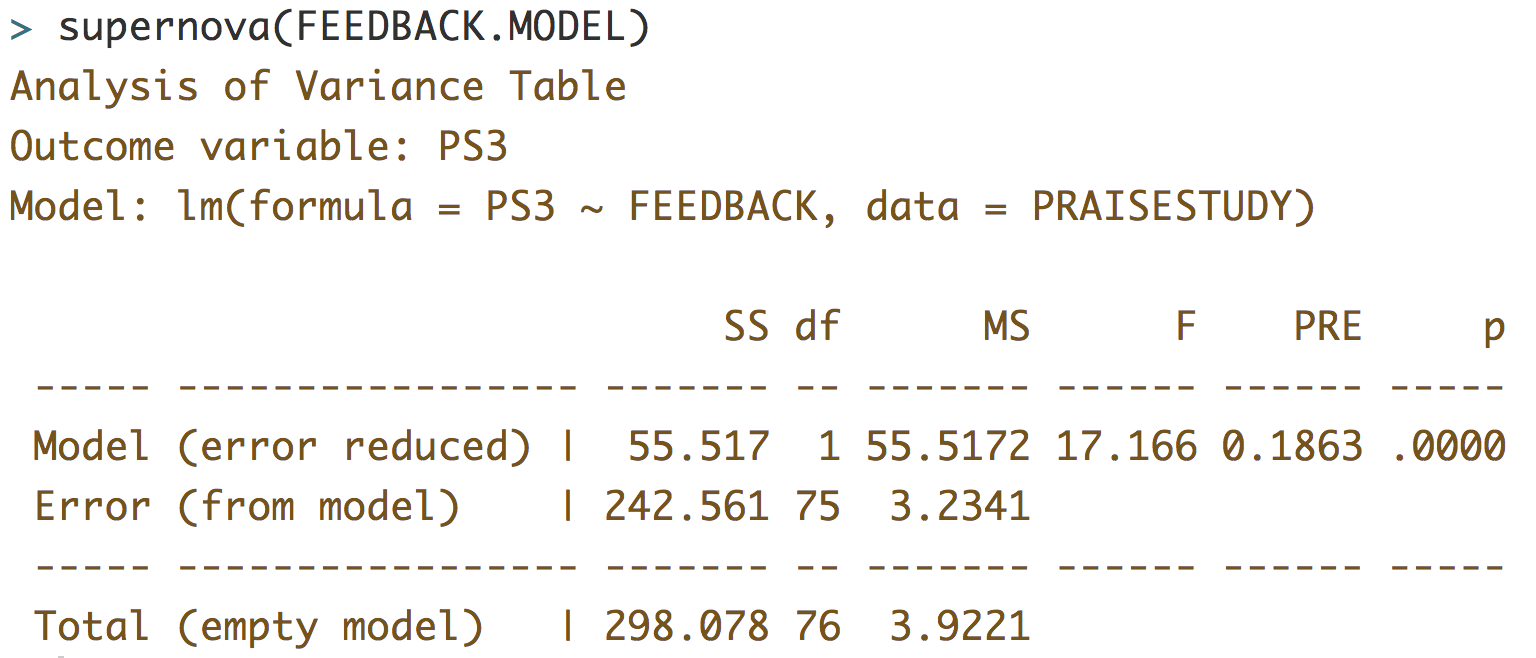
   

1. Remember we also used **shuffle()** on PS3? Can you pick out the real sample from the shuffled samples?

1. In which case was the real sample different: 6 or 7? How did it stand out?
2. If you can’t figure out whether the histogram of the real data is different from shuffled histograms, what kind of DGP might have produced the real data? (Which situation is like this: PS1 or PS3?)
3. Before we did this just by eyeballing. Now let’s figure out a quantitative way of trying to figure out whether our sample is really different from a random process (like shuffle). Here is the feedback model trying to explain variation in PS3 and the supernova table. Interpret the model fits and the PRE. Which of these statistics would be affected by sampling variation?





1. Refer back to the histograms shown in 6. What sample statistic is really different in the sample versus the shuffled histograms: ?
2. We have an intuition that really is different from the shuffled s. Let’s try and show exactly how different it is. Here we will create a plan with the distribution triad.
3. Let’s think about this as model comparison.

FEEDBACK model:

Empty model:

What if our sample estimate of could have come from a DGP with = 0? Write that FEEDBACK model here:

Is it any different from the empty model?

1. Let’s carry out the plan we conceived of using R.