Gov 1372 - Groups and Identities

Luke Kolar

September 23, 2021

Marriage and Partisan Polarization

Iyengar and Westwood (2014) use answers to questions about a child marrying an in-party or out-party spouse as one way of characterizing affective partisan polarization. Some authors have questioned if the way this question is framed too coarsely. In particular, Klar et al. (2018) argue that, by making the prospective child-in-law's partisanship salient, the marriage question may be picking up on respondents dislike of partisanship in general, rather than a dislike of the opposing party.

The in-class survey you took was a partial replication of the Klar et al. (2018) study. We randomized whether you were asked about a prospective child-in-law who "frequently talks about politics," "rarely talks about politics," or a person whose frequency of discussing politics was not mentioned. This last, control, condition matches the wording of the question used in Iyengar and Westwood (2014).

Data Details:

- File Name: `Sep23ClassData_clean.csv`
- Source: These data are from the survey you took in class. The questions are slightly adapted versions of some of the questions used in Klar et al (2018) (see here for the supplemental material of that study with the original questionnaire, if you are interested).

Variable Name	Variable Description
`pid3`	Political party preference
`pid_lean`	If a respondent didn't identify with the Democrats or Republicans in `QID1`, this indicates to which party (or neither) they feel closer
`strongGOP`	Indicator variable for whether the respondent identifies as a Strong Republican
`strongDEM`	Indicator variable for whether the respondent identifies as a Strong Democrat
`strongPARTISAN`	Indicator variable for whether the respondent identifies as a strong member of either major party
`party`	Party variable where those who lean toward either major party are counted as identifying with that party
`treatment`	Which treatment condition the respondent was randomly assigned to
`marryDemocrat`	The respondent's answer to how happy they would be if their child married a Democrat
`marryRepublican`	The respondent's answer to how happy they would be if their child married a Republican
`inPartyHappy`	Indicator variable for whether the respondent would be happy if their child married a member of their own party

Variable Name	Variable Description
`outPartyUnhappy`	Indicator variable for whether the respondent would be unhappy if their child married a member of the other major party
`polarized`	Indicator variable for whether the respondent was affectively polarized

Once again, the .Rmd version of this file has code you can use to load the data.

```
# remember to use the correct file name
marriage_data <- read_csv('Sep23ClassData_clean.csv')

messy <- read_csv('Sep23ClassData_messy.csv') %>%
  mutate(party = case_when(
    pid3 == "Democrat" | grepl("Democratic", pid_lean) ~ "Democrat",
    pid3 == "Republican" | grepl("Republican", pid_lean) ~ "Republican",
    pid3 == "Independent" & grepl("Neither", pid_lean) ~ "Indpendent",
    TRUE ~ NA_character_
  )) %>%
  mutate(treatment = case_when(
    !is.na(ControlRepublican) ~ "Control",
    !is.na(RarelyRepublican) ~ "Rarely",
    !is.na(FrequentlyRepublican) ~ "Frequently",
    TRUE ~ NA_character_
  )) %>%
  pivot_longer(c(ControlRepublican, RarelyRepublican, FrequentlyRepublican), values_to = "marryRepublica
  #pivot_longer(c(ControlDemocrat, RarelyDemocrat, FrequentlyDemocrat), values_to = "marryDemocrat")
```

These data are *not* the raw output from the survey you took. In particular, all of the indicator variables are the result of coding decisions and data processing done by the instructors (based on the procedures used in Klar et al. (2018)). For the first few questions, just open up the data and take a look at it (ask us if you need help viewing the data in spreadsheet format in RStudio).

Question 1

How were the `inPartyHappy` and `outPartyUnhappy` variables defined? Does this seem like a reasonable procedure? Do you notice any missing values? Why are they missing? How might the missing data affect researchers' inferences?

```
marriage_data %>%
  group_by(inPartyHappy) %>%
  summarize(n = n())
```

```
## # A tibble: 3 x 2
##   inPartyHappy     n
##   <lgl>        <int>
## 1 FALSE           32
## 2 TRUE            47
## 3 NA              10
```

ANSWER: *The `inPartyHappy` and `outPartyUnhappy` variables are boolean, meaning they take values of either `TRUE` or `FALSE`. This is reasonable for a simple regression, but it would be more thorough and interesting*

if they were scaled, so that researchers could consider a spectrum of happiness/dissatisfaction. Missing values are present as some respondents' party is neither Republican, Democrat, nor Independent with a partisan lean. If a respondent is a true Independent or third-party supporter, their "ingroup" and "outgroup" IDs are irrelevant for this experiment.

Question 2

How was the `polarized` variable defined? Is there another way you might consider coding this variable for individual polarization? What would be an advantage and a disadvantage of your alternative approach?

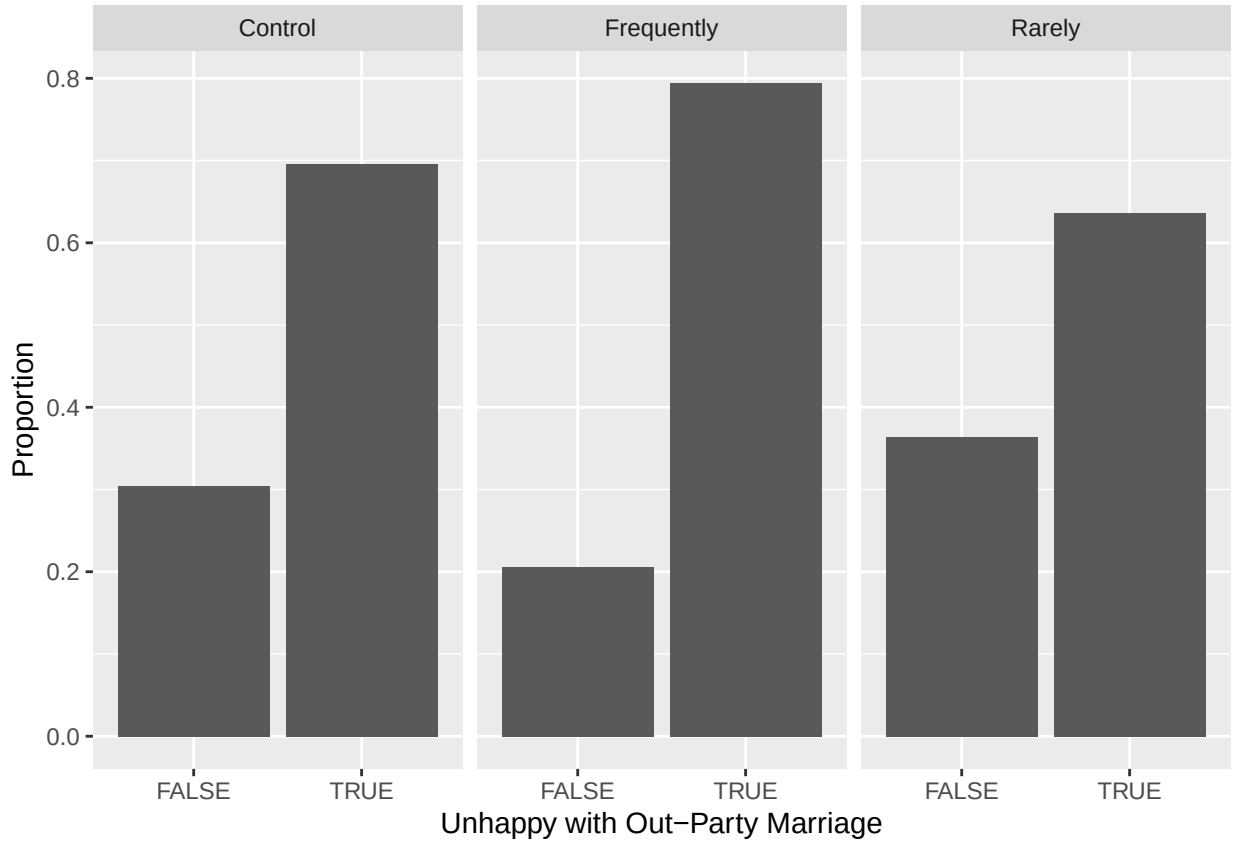
```
marriage_data %>%
  select(polarized) %>%
  glimpse()
```

```
## Rows: 89
## Columns: 1
## $ polarized <lgl> TRUE, TRUE, TRUE, FALSE, TRUE, FALSE, TRUE, NA, TRUE, TRU...
```

ANSWER: *The `polarized` variable is also boolean, and I'd make the same recommendation as above, but with greater emphasis here: this variable should be categorical, meaning it should be, say, a scale from 1 to 5. This would add more scientific rigor to the research.*

Now let's take a look at if there are differences in some of the key outcome variables depending on treatment status. Here is an example of how you might make a graph to look at if the rates of unhappiness with a prospective out-party in-law differ depending on the frequency with which they talk about politics.

```
ggplot(data = marriage_data %>% filter(is.na(outPartyUnhappy)==FALSE)) +
  geom_bar(mapping = aes(x = outPartyUnhappy, y = ..prop.., group = 1), stat = "count") +
  facet_wrap(~treatment) + ylab("Proportion") + xlab("Unhappy with Out-Party Marriage")
```



Question 3

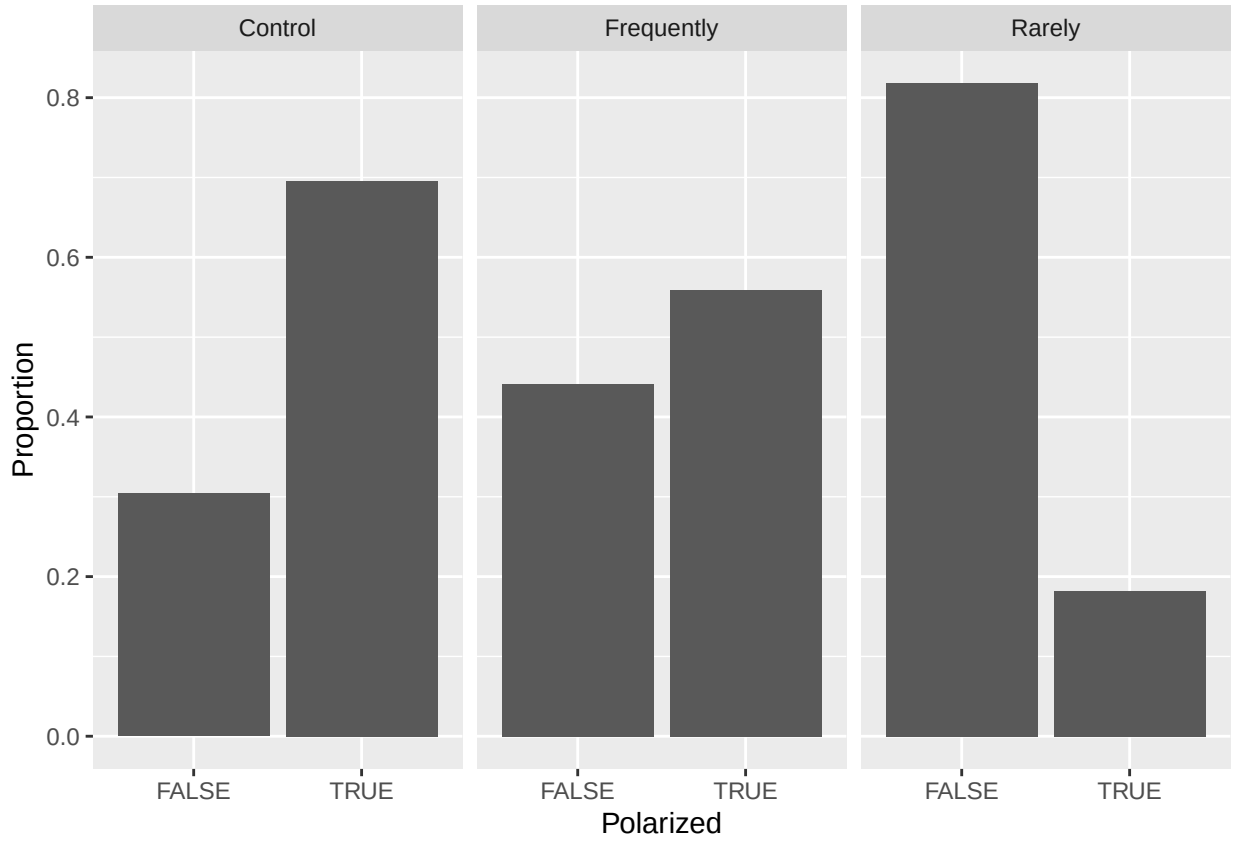
Comment on what you see in the example graph. Did the treatment affect unhappiness with a prospective out-party member coming into the family?

ANSWER: *The treatment does seem to have an effect on unhappiness with an out-party member coming into the family. About 10% fewer respondents in the "frequently" group (0.2) than in the control group (0.3) said they wouldn't be unhappy, while compared to the control group, almost 10% more respondents in the "rarely" group said they wouldn't be unhappy. This is about a 15 to 20% difference in the proportion of respondents who said they would be unhappy between the "frequently" and "rarely" groups.*

Question 4

Did the different treatment conditions affect the proportions of people who were affectively polarized? Make a plot and explain what you found.

```
marriage_data %>%
  filter(!is.na(polarized)) %>%
  ggplot(aes(x = polarized, y = ..prop.., group = 1)) + geom_bar() +
  facet_wrap(~treatment) + ylab("Proportion") + xlab("Polarized")
```



ANSWER: *My visualization shows that yes, a significantly higher proportion of those in the "frequently" group were polarized compared to the "rarely" group. However, both groups showed lower levels of polarization than the control group, an interesting finding. Nevertheless, the frequently and control groups are similar, and the disparity between their polarized proportions and that of the rarely group is striking.*

Question 5

Take a quick look at Figure 1 and Figure 2 in Klar et al. (2018). How do the results from our in-class data compare to their results? What might explain any differences? If there aren't an notable differences, is there a common pattern across the two datasets that is puzzling? What hypothesis do you have to explain it.

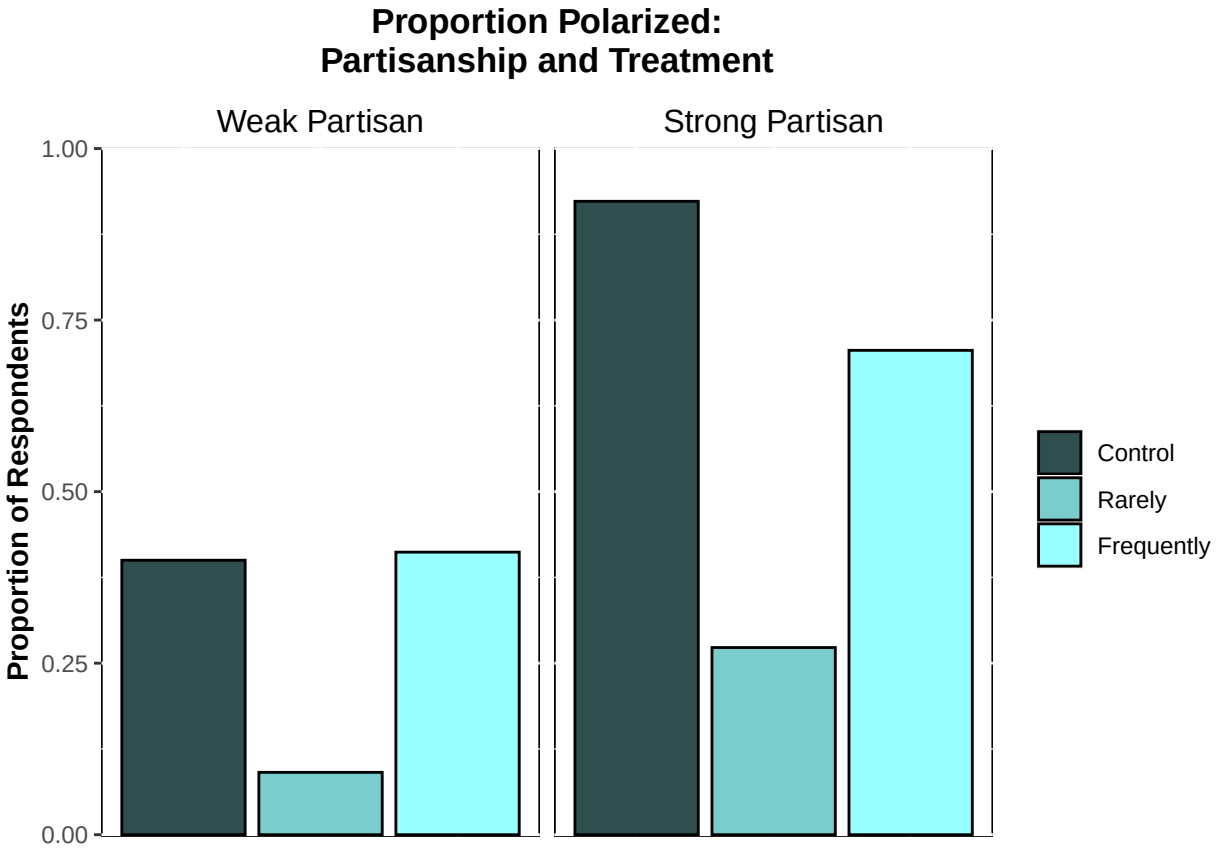
ANSWER: *The results are very similar. Unhappiness with out-party marriage shows the same relative proportions as my findings: the "frequently" group has the highest proportion unhappy, followed by the control and "rarely." Likewise, relative proportion sizes are ordered the same for polarization: control, "frequently," "rarely." The control group having the highest proportion polarized is the puzzling similarity between the experiments. The reason for this is likely that the polarization index here takes into account respondents' happiness with their child marrying someone of the same party who talks about politics frequently or rarely. The experiment found that people are biased against political discussion in general, so respondents might express unhappiness with in-party marriages, too, driving the polarization variable for "frequently" lower than the control group.*

Question 6 (Data Science Question)

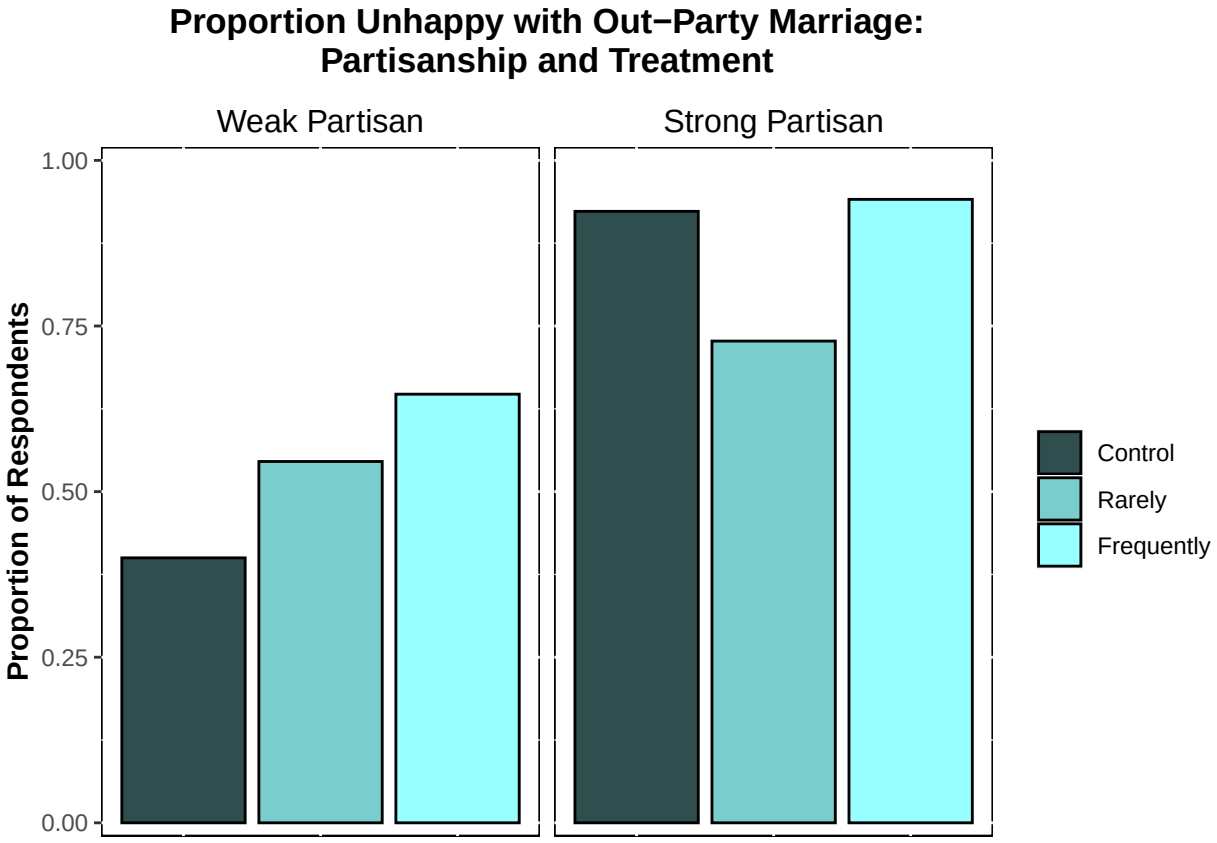
We might also be interested in if things looked different for weak vs. strong partisans. Pick one of the two outcome variables you just examined and make a plot that would help us

understand if responses within and across treatment groups differ between weak and strong partisans.

```
marriage_data %>%
  filter(!is.na(polarized)) %>%
  group_by(treatment, strongPARTISAN) %>%
  summarize(polarized = mean(polarized)) %>%
  mutate(strongPARTISAN = as.character(strongPARTISAN)) %>%
  mutate(strongPARTISAN = recode(strongPARTISAN,
                                 "TRUE" = "Strong Partisan",
                                 "FALSE" = "Weak Partisan")) %>%
  mutate(treatment = factor(treatment, levels = c("Control", "Rarely", "Frequently"))) %>%
  mutate(strongPARTISAN = factor(strongPARTISAN, levels = c("Weak Partisan", "Strong Partisan"))) %>%
  ggplot(aes(x = treatment, y = polarized, fill = treatment)) + geom_col(color = "black") +
    facet_wrap(~strongPARTISAN) +
    labs(x = NULL, y = "Proportion of Respondents",
         title = "Proportion Polarized:\nPartisanship and Treatment", fill = NULL) +
    theme(plot.title = element_text(face = "bold", hjust = 0.5),
          axis.text.x = element_blank(),
          axis.ticks.x = element_blank(),
          plot.background = element_rect(fill = "white"),
          axis.title = element_text(face = "bold"),
          panel.background = element_rect(fill = "white", color = "black"),
          strip.background = element_rect(color = "NA", fill = "NA"),
          strip.text.x = element_text(size = 12, color = "black")) +
  scale_y_continuous(expand = c(0.001, 0.001), limits = c(0, 1)) +
  scale_fill_manual(values = c("darkslategrey", "darkslategray3", "darkslategray1"))
```




```
marriage_data %>%
  filter(!is.na(outPartyUnhappy)) %>%
  group_by(treatment, strongPARTISAN) %>%
  mutate(strongPARTISAN = as.character(strongPARTISAN)) %>%
  mutate(strongPARTISAN = recode(strongPARTISAN,
                                 "TRUE" = "Strong Partisan",
                                 "FALSE" = "Weak Partisan")) %>%
  mutate(strongPARTISAN = factor(strongPARTISAN, levels = c("Weak Partisan", "Strong Partisan"))) %>%
  mutate(treatment = factor(treatment, levels = c("Control", "Rarely", "Frequently"))) %>%
  summarize(outPartyUnhappy = mean(outPartyUnhappy)) %>%
  ggplot(aes(x = treatment, y = outPartyUnhappy, fill = treatment)) + geom_col(color = "black") +
    facet_wrap(~strongPARTISAN) +
    labs(x = NULL, y = "Proportion of Respondents",
         title = "Proportion Unhappy with Out-Party Marriage:\nPartisanship and Treatment", fill = NULL)
    theme(plot.title = element_text(face = "bold", hjust = 0.5),
          axis.text.x = element_blank(),
          axis.ticks.x = element_blank(),
          plot.background = element_rect(fill = "white"),
          axis.title = element_text(face = "bold"),
          panel.background = element_rect(fill = "white", color = "black"),
          strip.background = element_rect(color = "NA", fill = "NA"),
          strip.text.x = element_text(size = 12, color = "black")) +
    scale_y_continuous(expand = c(0.01, 0.01), limits = c(0, 1)) +
    scale_fill_manual(values = c("darkslategrey", "darkslategray3", "darkslategray1"))
```



Question 7

Are there any other issues you can think of that might confound the utility of the marriage question as a measure of affective polarization? If you have any concerns, how might you design a study to evaluate your hypotheses?

ANSWER: *Answer here*

Question 8

Based on the data and your work on this assignment, are there any changes you would make to the Iyengar and Westwood (2014) study or the Klar et al. (2018) study or extensions of the research that you would like to see? (For example, would you alter the wording of any questions, change the experimental protocol, or come to any different conclusions?)

ANSWER: *Answer here*