**Name: Katie Lee**

**Name of partner(s) if you worked with someone on this assignment: Elodia Lunn**

**Hands On 6 Questions**

**Experiment 1**

1. **What is the sample size for this experiment? (i.e., how many people participated in total?)**

**265 people participated in this experiment**

1. **How many participants took the experiment in each semester? What is the benefit of analyzing the data from this larger sample of students across multiple classes/semesters, instead of just analyzing the data from our class this semester?**

**111 People took it in prior semesters (fall and spring 22) 154 People took it in spring 24 semester**

Having a larger sample of students across multiple class can external validity and thus higher generalizability and its a more representative sample by collecting data across multiple points of time.

1. **What were the conditions of the experiment and how many participants were in each condition?**

**Conditions: write or type**

**Type (145 participants), Write (120 participants)**

1. **Using the descriptive statistics and/or histograms, briefly describe how participants scored on the memory test overall, on the factual questions, and on the conceptual questions.** For your reference, the total number of points possible was 19 overall, 14 on the factual questions, and 5 on the conceptual questions.

**Participants generally scored around 5.02 out of 14 on the factually questions, and around 1.96 out of 5 on the conceptual questions. Also the total score (which was the the score that combined the factual and conceptual questions) is 6.98/19.**

**With this in mind, participants generally scored low on the memory test questions**

1. **Based on the means broken down by condition and the box plots, does there appear to be a difference in how participants scored on the memory test based on condition?** Briefly describe whether and how scores appear to differ by condition overall, on the factual questions, and on the conceptual questions. Note that we are not talking about whether the difference is significant here – just examining whether the mean scores differ by condition and how (i.e., which condition scored higher?).

**Writing did better overall and the mean in both conceptual and factual scores were higher than the mean scores of typing. In terms of total score the writing mean was higher than the typing mean about 2+.**

1. **Were the differences between conditions on memory test performance statistically significant?** Indicate whether participants’ notetaking format significantly affected their performance overall, on the factual questions, and on the conceptual questions. Explain how you know, reporting both *p*-values and confidence intervals.

**The total score and factual score differences were statistically significant, but not the conceptual score differences, because the p-value for the conceptual scores was greater than 0.05.**

**We found this through the analysis of confidence intervals: for both the total score (-1.1886 to -0.290) and factual score (-1.380 to -0.2029), the lower and upper bounds of the confidence intervals do not include 0, indicating significant differences. In contrast, the confidence interval for the conceptual scores (-0.657 to 0.064) includes 0, suggesting that the differences observed here are not statistically significant.**

1. **Did our class experiment replicate the results of Mueller & Oppenheimer (2014), Urry et al. (2021), or neither?** Briefly explain your reasoning.

**Yes it for the most part replicates Mueller and Oppenheimer’s results that writing provides better memory retention than typing.**

**Mueller and Oppenheimer found on conceptual scores, notetaking was higher than writing but we found that there was not much significance.**

**Other than the conceptual scores which did not show much significance, the other ones (total and factual) here significant and agreed with Mueller**

1. **How many participants correctly guessed the purpose of the study or got close? Explain how these participants guessing the purpose of the study could threaten the internal validity of the study.**

**62 correctly guessed the purpose of the study and 21 got close to guessing the purpose of the study.**

**People who guessed the purpose of the study may threat internal validity because they may exude the Good-participant / please-you effect (when participants adjust their behavior to conform to the perceived expectations of the researcher, rather than behaving naturally or according to their true reactions) By having participants adjust their behavior it is not representative of the population they are trying to measure.**

1. Thinking about the experiment’s design and/or your experience participating in the study, **identify one additional potential threat to the internal validity of the study and one potential threat to the external validity of the study.** *(If you are stuck on internal validity, a hint: Could there have been other explanations for the difference in memory test scores besides notetaking format?)*

**Internal: One of the design confounds (confounding variables) is that during the experiment, two individual behind us were talking to each other during the experiment (when we watched the film to take notes) this affects participants ability to focus nearby.**

**External: We only included Barnard College students in our sample, which limits generalizability due to the sample's lack of diversity. Conclusions drawn from this specific sample may not be applicable to individuals outside of this group.**

**Another concern is ecological validity; the experiment was conducted in a classroom setting aimed at testing memory comprehension. If this experiment intends to generalize findings across all types of memory comprehension (typing vs writing) , conclusions drawn from this specific environment may not be applicable to individuals in different environments.**

1. **If you wanted to replicate this experiment using a within-subjects design, how would you alter the study’s procedure?**

**Use different videos when testing typing and writing. 50 percent of the population would write and the other would type. Then they would switch (one population would type and the other would write) but for a different video. The two groups would be in two separate rooms so that they do not know that they will be typing after.**

**Experiment 2**

1. **What is the sample size for this experiment? (i.e., how many people participated in total?)**

**268 People Participated**

1. **How many people participated in each level of the independent variable?**

**All, 268 People Participated**

1. **Using the descriptive statistics and/or histograms, briefly describe how participants rated their satisfaction with the research explanations that contained neuroscience content and the explanations that did not contain neuroscience content.** For your reference, the satisfaction scale ranged from -3 (very unsatisfying) to 3 (very satisfying), and 0 meant neither satisfying nor dissatisfying.

**Participants had higher satisfaction scores when there contained more neuroscience content. When there was no neuroscience content then the mean was lower.**

1. **Based on the means, does there appear to be a difference in how satisfied participants were with research explanations that did vs. did not contain neuroscience content?** Briefly describe whether and how satisfaction scores appear to differ by whether neuroscience content was included. Note that we are not talking about whether the difference is significant here – just examining whether the mean satisfaction scores differ by neuroscience content and how (i.e., which type did people find more satisfying on average?).

**Looking at the mean research explanations that contained neuro science content was higher than research explanations that did not contain neuroscience content. 0.765 compared to 0.524**

1. **Was the difference in satisfaction between research explanations that did vs. did not contain neuroscience content statistically significant?** Indicate whether inclusion of neuroscience content significantly affected participants’ satisfaction with research explanations and explain how you know, reporting both the *p*-value and confidence interval.

**It is statistically significant because the p value is less than 0.05 (0.003 < 0.05) and if you look at the confidence interval, 0 is not included in the range.**

1. **Did our class experiment replicate the results of Weisberg et al., 2008; 2015?** Briefly explain your reasoning.

**Given the mean difference of of .241 and that explanations with neuro content having a higher rating compared to without neuro content, it does replicate the results of Weisberg finding that people tend to gravitate towards explanation with neuro content are seen as better**

1. **How many participants correctly guessed the purpose of the study or got close?** Researchers are sometimes concerned that seeing all levels of the IV in a within-subjects experiment will clue participants in to the purpose of the study. Does that seem like more of a concern in this within-subjects study than the previous between-subjects design?

**4 people correctly guessed it while 8 people got close (total 12 participants) This does not seem like a concern because a vast majority of participants did not guess the study (256 people didn’t guess) Fewer people guessed correctly then the between subject design.**

1. Thinking about the experiment’s design and/or your experience participating in the study, **identify one potential threat to the internal validity of the study and one potential threat to the external validity of the study.** *(If you are stuck on internal validity, a hint: Could there have been other differences besides neuroscience content between the research explanations?)*

**Internal: A potential threat of the experiment is order effects, the order in which questions you answer can be a confounding variable. People could “anchor” the first response and use that as the standard and try to compare the next explanations towards that one. This may influence how each individual results to each question.**

**External: Since we only used barnard students who specifically are taking psychology classes, it is not that generalizable because we have a very narrow**

1. **If you wanted to replicate this experiment using a between-subjects design, how would you alter the study’s procedure?**

**I would split them into two groups randomly (group 1 would only read research explanations with neuro and group 2 would only be exposed to research explanations without neuro terminology.)**