Psych 610, HW2

Due Wednesday, September 20, 2023, 1:30pm on Canvas

* Submit **both** an HTML file and a .Rmd file
* Email Ben, LiChen, or Nick with any questions

**R Markdown**

We did not provide you with an R Markdown template this week. For the rest of the class, you will create your own RMarkdown to complete your assignments. Feel free to use the first week’s template as a guide when creating your own.

**Part 1. Reading Questions**

Answer in your R Markdown file. Ideal answer length appears in parentheses after the question.

1. Sometimes, there can be tradeoffs between different types of validity. Describe a situation in which different types of validity may stand in conflict and explain the nature of the tradeoff. (2 sentences)
2. Can a single measure be both valid and *not* reliable? Support your argument. (2 sentences)
3. A classic finding in language development is that the number of words a child is exposed to correlates with the size of that child’s vocabulary. That is, the more words children hear, the more words they know. List and briefly describe four causal pathways that could explain the association between number of words a child is exposed to (X) and the size of a child’s vocabulary (Y). Hint: only *one* of your four should include any variable other than X and Y. (4 sentences)

**Part 2. Experimental Design**

1. In your R Markdown file, identify 4 problems with the study below (there are more than 4 possibilities, just report your favorites), and identify which type of validity (or reliability) each violates.

An industrial and organizational psychologist wanted to examine how the workers' capacity to concentrate was affected by the auditory nuisances in a factory (such as beeping machines, slamming doors, honking delivery trucks). She conducted a study in the laboratory in which she asked 20 workers to learn a list of 30 words in two minutes. She used an experimental manipulation (variable "nuisance"). Ten participants (i.e., workers) learned the 30 words while hearing a continuous white noise of 20 decibels ("silence condition") whereas the other ten participants heard a continuous white noise of 40 decibels during the learning phase ("noise condition"). The psychologist had participants come to the laboratory in groups of two. The first participant to arrive was always in the "silence condition" (cubicle 1) and the second participant to arrive was always in the "noise condition" (cubicle 2). Then, she measured how many words the participants could remember correctly. She analyzed her data with a simple regression analysis in which she regressed the recall scores on the experimental condition (coded 1 and 2). The regression coefficient b1 was not reliably different from 0, *F*(1,18) = .82, *p* = .377. She concluded that auditory nuisances have no effect on concentration capacity.

**Part 3. Data Analysis**

The city of Madison has asked you to conduct a study examining if people who feel personally involved in the community have more favorable attitudes toward protecting the environment. The city also wanted to know if people who feel involved in the community know more about the opportunities to protect the environment the city makes available to them (e.g., recycling center, tax subsidies for solar panels).

To measure attitudes toward protecting the environment, you used a 10-item scale measuring the importance that participants attribute to different environment-related behaviors (e.g., recycling, using public transportation). You also measured feelings of personal involvement in the community (with 3 items). These items used 0-6 response Likert scales, with higher scores indicating higher importance and involvement, respectively. Finally, you had participants respond to 6 yes/no questions measuring their awareness of opportunities to protect the environment. You recruited 60 Madisonians to participate in the study.

Because positive environmental attitudes are considered socially desirable in Madison, you would like to ensure that your environmental attitudes scale actually measures environmental attitudes, rather than the tendency to present oneself in socially desirable ways. To guard against this possibility, you contacted one third of your sample and obtained additional information from them. First, you gave them a questionnaire measuring their tendency to present themselves in socially desirable ways (note that some items are reverse-coded!). Second, you obtained a measure of pro-environmental behavior by asking participants to give you their last electricity bill and the footprint (i.e., surface area) of their place of residence in square meters. Third, you asked for the participants’ permission to observe their water consumption for one month. You obtained the water consumption of these participants during the four months prior to the testing month. By comparing the water consumption during the previous four months to the water consumption during the testing month, you obtain a behavioral indicator of the participants’ tendency to present themselves in socially desirable ways.

You will find the data in a file called “environmentalism\_data.csv”. The codebook for this data file is at the bottom of this document. Consult the codebook to determine what different variables in the dataset mean.

**Complete the following tasks:**

Answer in your R Markdown file.

1. Read the data file into an R dataframe. Run two commands to give you a sense of what’s going on in these data.
2. Test the reliability of the environmental attitudes scale (att1 to att10). Are there any problematic items? Which? How do you know?
3. If you think any items could be removed, check the reliability of the scale again without the problem item(s). Do you think the change in alpha warrants removing this item (assuming you specified this approach in your preregistration)?
4. Compute an average attitude score (“att\_m”) that corresponds to the mean of all 10 of the attitude items (regardless of what you discovered in your reliability analysis). This score is your measure of pro-environmental attitudes.
5. Compute an average involvement score (“inv\_m") that corresponds to the mean of the three personal involvement items.
6. Compute an average knowledge score (“know\_m”) that corresponds to the mean of the five knowledge items.
7. Produce means and standard deviations for the variables you created in questions 4-6 (do it using only 1 line of code).
8. Compute the correlations between the variables you created in questions 4-6. Which variables are most closely related? Does this result make sense?
9. Check the reliability of the social desirability scale (sd1 to sd8). Is the scale reliable?
10. Compute an average social desirability score (“sd\_m”) that corresponds to the mean of all 8 items. This score is your measure of social desirability.
11. Create a new variable (“elec\_per\_foot”) that reflects participants' electricity consumption per square meter of living space. This score is your measure of pro-environmental behavior.
12. For each participant, average the water consumption across the four pre-test months and save it as a new variable (“water\_pre”). Then, compute the difference between this average score and water consumption during the testing month (“water\_test”) and save this as a new variable (“water\_change”). Make sure positive scores mean someone *increased* their water consumption during the testing month). This variable represents a behavioral measure of social desirability.
13. Generate a correlation table consistent with the principle of the multitrait-multimethod matrix (two traits: pro-environmental attitudes and social desirability tendency; two methods: questionnaire measure and behavioral measure). What is the strongest relationship? Does this result make sense? Does the pro-environmental attitudes scale appear to have high construct validity?
14. At the bottom of your R Markdown file, tell us how many hours you spent working on this homework assignment.

