CONDUCTING A FORMAL EXPERIMENT AND RELATED STATISTICAL ANALYSES

INFO 470 Research Methods in Informatics

OBJECTIVE

The goal of this assignment is to have you gain experience conducting a formal experiment and the inferential statistical analyses that go with it. The type of experiment involves participants typing text in two different body postures (sitting and standing) to see how these postures might affect human performance. You will analyze the data from this experiment using the R statistical programming language and report your results in a formal experiment write-up.

FAIR WARNING!

This is not an assignment you can do right before the deadline. You must start immediately!

GROUP SIZE

This assignment is to be done in **groups of two (2)**. Groups should be members of the same TA section. If you are unable to achieve this, please see your TA or the Instructor.

RELATED READINGS

Be sure you have read and understood the following readings from the course. If you have not, your chances at succeeding on this assignment are significantly diminished.

- Lazar, J., Feng, J.H. and Hochheiser, H. (2010). Experimental research. Chapter 2 in *Research Methods in Human-Computer Interaction*. West Sussex, U.K.: John Wiley & Sons, pp. 19-40.
- Lazar, J., Feng, J.H. and Hochheiser, H. (2010). Experimental design. Chapter 3 in *Research Methods in Human-Computer Interaction*. West Sussex, U.K.: John Wiley & Sons, pp. 41-68.
- Lazar, J., Feng, J.H. and Hochheiser, H. (2010). Statistical analysis. Chapter 4 in *Research Methods in Human-Computer Interaction*. West Sussex, U.K.: John Wiley & Sons, pp. 69-98.
- MacKenzie, I.S. (2013). Hypothesis testing. Chapter 6 in *Human-Computer Interaction: An Empirical Research Perspective*. Waltham, MA: Morgan Kaufmann, pp. 191-209.
- MacKenzie, I.S. (2013). Sections 8.2.5 – 8.2.7 in *Human-Computer Interaction: An Empirical Research Perspective*. Waltham, MA: Morgan Kaufmann, pp. 299-303.

WHAT TO DO

You will recruit 10 participants to take part in your study. Half should have been considered male and half female at birth.[1] You will need access to a Microsoft Windows machine to run the *TextTest* software available from the course website. TextTest enables you to run text entry studies and to get a CSV file as a data table for your results. Ideally, you will test your participants all in the same (quiet) location so as to avoid introducing confounds and unwanted variance into the study.

Your test should have each participant enter 30 phrases while sitting and 30 phrases while standing. The first 5 phrases in each posture should be practice phrases, and the remaining 25 phrases should be test phrases. Be sure to **counterbalance** the order of standing and sitting so as to avoid carryover effects. For both postures, be sure that your participants are comfortable and able to type easily. For the standing posture, you may need to find a desk that you can raise, or use the top of a file cabinet, or use a stable box set atop a desk. Whatever way you achieve the standing posture is fine, so long as you are not disadvantaging that condition (e.g., by having a wobbly laptop, an unstable surface, a keyboard at not-the-right-height, etc.). Your goal is to give both sitting and standing a "fair shot"

[1] This criterion is called sex-assigned-at-birth and is purely a biological distinction. The more descriptive and complex construct of gender could be used in more sophisticated research projects, but for our purposes, we'll keep this simple for the assignment.

at showing their optimal performance. Only then can you equitably compare these postures and draw any conclusions about human performance in them.

Alongside your posture factor with levels *sitting* and *standing*, you have a sex-assigned-at-birth[2] factor with levels *male* and *female*. (Think about whether your two factors are between-subjects or within-subjects.) By testing an even mix of males and females, you will be able to analyze whether there is any effect of sex on human typing performance, and whether different sexes perform differentially across postures. (Note: Of course I expect there to be no difference in sexes, but we're going to examine this for fun anyway!)

Be sure you can fill in the blanks in the following description of this study. The small blanks require numbers. The large blanks require words. (Hint: such a description should be part of your experiment write-up in the end.)

> This experiment is a __ × __ mixed factorial design with a _____-subjects factor for _____ and a _____-subjects factor for _____. The _____ factor has __ levels, _____ and _____. The _____ factor has __ levels, _____ and _____.

INFORMED CONSENT

Before conducting any experiment sessions, be sure you have your participants' permission. Explain to them what you are doing and why, and give them an opportunity to opt-out or express discomfort at any aspect. At the end of your experiment session, give them the opportunity to talk to you about the study and any concerns they might have. Obviously, do not use your participants' names or specific identifying information so as to protect their privacy.

USING THE TEXTEST SOFTWARE

TextTest is a software program that I wrote for conducting text entry studies. Download the TextTest zip file[3] from the course website. Unzip the file and locate TextTest.exe. (There is no installer for this program; you just run the executable directly.) Go to the Tools > Options menu and on the Auto tab, check "Stop the test after trial" and enter 30 trials. Also check "Switch to test mode after trial" and enter 5 trials. Click OK. Then do Test > Start New. Before clicking Start, it is important to use File > Rename Log and give the log name something like "P01_F_sit.xml" for participant #1, female, sitting. That way, you encode the conditions of the session in the filename itself. (More on this later.) Then click Start and have your participant begin the test.

Your participant can hit Enter after each trial to advance to the next trial, so he or she should not need to use the mouse at all. Timing is *not* from the clicking of the Start or Next button, but from the timestamp associated with the first and last keypresses. Arrow keys and clicking into the text with the mouse are disallowed. Backspace is the only error correction mechanism. A general rule is to tell your participants to "proceed quickly and accurately," having them balance speed and accuracy, both. Fixing errors is good, but if participants leave a few errors in the final text, that is fine. (The text entry analysis algorithms take errors into account and yield error rates in your measures.) Certainly, if your participants notice an error "way back" in their text, they should not feel compelled to backspace pathologically through a bunch of correct letters to get to their error. Participants should not talk while doing the test, and should remain focused for the duration of the study session. It is your job as the experimenter to ensure that your study sessions are conducted professionally and ethically.

WRANGLING YOUR DATA

In the end, you will have 20 XML log files produced by TextTest, two logs for each participant, one for sitting and one for standing. TextTest can now parse these XML files for you to produce corresponding CSV files of text entry measures for statistical analyses. With all of your XML files in one directory, run TextTest and select Analyze >

[2] For brevity, "sex-assigned-at-birth" will be termed just "sex" hereafter, but the full phrase is intended throughout.

[3] There are three other files in the zip included for completeness, but you don't need to worry about these. For those of you who are curious, the TextTest.html file allows the use of TextTest from a web page that acts as a client sending keystrokes to TextTest.exe acting as a server. The TcpSendText.exe program is similar but this client is a desktop executable. The WobbrockLib.dll is a library of standard routines I use in many of my programs, including TextTest. It must be in the same directory as TextTest.exe for TextTest to run.

Logs. You can then multi-select all of your XML log files at once for analysis. When you see the "Output Options" dialog, just check the first option, "Main measures for each trial (*.csv)." (The other checkboxes produce more specialized text entry analyses, which you're free to play with but don't need.)

You will now see that TextTest has made a CSV file for every XML file you selected. You can open the CSV files in Microsoft Excel to see how each file has a column header row and a row for every text entry trial. (The topmost row is a credit citation link, which you will remove.) Your job now is to combine these separate CSV files into a single CSV file with one column header row and all your testing data. There are five things that you need to do:

- Remove the credit citation link from atop each CSV file. You can do this in Microsoft Excel by clicking on row number 1 in the left margin and then clicking the Delete button on the Home tab of the ribbon. The new first line of each file should be the column header names.

- Each CSV file was for a single participant of a given sex in a given posture, but TextTest knows nothing of these details. So you need to create three new columns at the left of every CSV file. The first column should be *Participant*, the next should be *Sex*, and the third should be *Posture*. If you named your files well, as suggested above (e.g., "P01_F_sit.xml"), it will be easy to know which files represented which participants, sexes, and postures. Creating these columns is easy in Excel and filling in their values is easy with either copy+paste or by dragging. Use "F" or "M" for the values of *Sex*, and use "sit" or "stand" for the values of *Posture*.

- You do not want to analyze practice trials. These trials were just for your participants to warm up and find their rhythm. Therefore, delete all rows that have the *Testing?* Column equal to "0", which means "false" and therefore were practice trials. They should be the first 5 rows of every CSV file you have. (In Microsoft Excel, you can highlight these five rows by dragging down the left of your table and clicking the Delete button on the Home tab in the ribbon.)
- You now need to remove the column header names row from each CSV file because as you go into <u>one</u> unified CSV file, you don't want the column header names repeated in your table. However, you do need to keep *one* header row for your new unified file that will contain all your data. Set up that master file with a single header row, and remove the header rows from each of the individual CSV files.
- Finally, copy and paste all of your CSV data into your new master CSV file. This file should now have every testing trial you conducted, with columns at the left encoding participant number, sex, and posture. With 10 participants doing 2 postures and 25 test trials in each posture, plus a single header row, you should have $10 \times 2 \times 25 + 1 = 501$ total rows in your CSV file. If you do not have 501 rows, something went wrong!

Data wrangling is part of any empirical research process. It's always a pain. The good news is that now your data is ready for statistical analysis!

ANALYZING YOUR DATA

Remember again what you're trying to do: You now have all of your data in a single CSV file and you want to know whether posture and/or sex made a difference in typing performance. So *Posture* and *Sex* are your independent variables. (By convention, we capitalize and italicize independent variable names, also known as factors.) What difference specifically do we care about? For this assignment, we'll be examining a single dependent variable, also known as a "measure" or a "response," which is *AdjWPM*, for "adjusted words per minute." You will see this as a column in your CSV file. This response is a combined speed-accuracy measure that is like the speed measure of words per minute (WPM) but reduces raw WPM in proportion to the number of uncorrected errors left in the final transcribed text for each trial. Errors that are made and fixed <u>during</u> entry take time, but errors that <u>remain</u> in the transcribed text are at odds with speed, and thus we reduce WPM accordingly as AdjWPM.

Your job now is to bring your single unified CSV file into the RStudio environment using the R statistical programming language and to depict and analyze your data. Specifically, your analysis should include: (1) descriptive statistics (means and standard deviations) for all combinations of *Posture* × *Sex*; (2) an omnibus (overall) ANOVA taking into account the study design; and (3) appropriate Holm-corrected *post hoc* pairwise comparisons.

In addition to your analyses, you will make one or more histograms, a boxplot, and an interaction plot.

We don't expect you at this point to be experts in R or in statistical analyses. To make conducting your analyses and producing your plots easier, we have given you R code as recipes on the course website. We will also walk through these recipes in class and possibly in section. Get the info470.R file and the info470.zip CSV data files that go with it **and be sure you understand how these analyses work**. Understanding these files will serve as a guide for your own analyses.

Hint: Your main analysis code will look something like this:

```
# read in the data file and set factors and contrasts for factors
df = read.csv("data.csv")
df$Participant = factor(df$Participant)
df$Trial = factor(df$Trial)
contrasts(df$Sex) <- "contr.sum"
contrasts(df$Posture) <- "contr.sum"
contrasts(df$Trial) <- "contr.sum"

# conduct the mixed-factorial ANOVA
m = aov(AdjWPM ~ Sex * Posture + Error(Participant/Posture), data=df)
summary(m)

# if the interaction was significant, this would be the pairwise comparison code
summary(as.glht(pairs(lsmeans(m, pairwise ~ Sex * Posture))), test=adjusted(type="holm"))
```

Alternatively, or in addition, you could use the ezANOVA function like we have done in class and shown in info470.R, and conduct your pairwise comparisons manually. Note that it may be easiest to first reshape your data table from long-format to wide-format to conduct paired-samples *t*-tests. See the `dcast` function in the code we have given you.

WRITING UP YOUR REPORT

Writing up a formal experiment requires rigor and attention to detail. It is also somewhat formulaic—not in a bad way, but experiments have been written up for hundreds of years, and best practices are firmly established. You should follow these practices. Fortunately, they are described for you in Chapter 8 of the MacKenzie (2013) book, specifically sections 8.2.5 – 8.2.7 (pp. 299-303). You should structure your paper with an Intro, followed by a Method section with MacKenzie's four subsections, followed by Results, Discussion, and finally, Conclusion.

Consider using tables to report means and standard deviations. When you report your statistical tests, be sure you use the proper annotation and formulation for these results. (Hint: see Figure 6.16 on page 205 of the MacKenzie (2013) book for an example of how a two-way ANOVA might be reported.)

You should include figures in your report. For example, a non-identifying photo of a participant at the study setup, especially in the standing condition, would be illustrative. Graphs of your data are also essential. Include reasonably-sized graphs that you can make with R or Excel, especially an interaction plot, an example of which can be seen on page 56 of the Lazar et al. (2010) book.

WHAT TO TURN IN

At the start of class on the due date, turn in your polished experiment write-up. Be sure both authors' names are on it. Turn in only one copy per group. In addition, print out and turn in your R code file that you used to explore, plot, and analyze your data.

Your write-up should be no longer than 2,000 words.