**Psychology of Perception Psychology 4165, Fall 2015 Laboratory 3 (week 2):**

**Stroop Word-Color Task**



From Stroop (1935)

**Lab Overview**

**Previously, in Perception Lab:**

|  |
| --- |
| The purpose of Lab 3 is to pull these tasks together into a simple experiment where you build the entire PsychoPy script, collect data, then perform individual and group analyses by assembling the basic data analysis elements R (import, re-organize, analyze, plot).  **By the end of this lab, you should be able to:**   * **Build a PsychoPy experiment from scratch** * **Assemble R script components to analyze individual and group data** * **Test a hypothesis from response times**   **Week 1:** Define the Stroop Effect, build a script in PsychoPy to test it.  **Week 2:** Collect data, build a script in R to analyze individual data, analyze individual data.  **Week 3:** Build a script to analyze group data, analyze group data. |

Last week you built an experiment in PsychoPy to demonstrate the so-called “Stroop-Effect,” where naming the color of color words tends to be faster when the word and color are congruent (semantic facilitation), and slower when the word and color are incongruent (semantic interference). This went extremely well, and as far as I can tell, every person has a working experiment (which is good, because you’ll need it for today).

This week, you will use your own experiment to collect the responses of you and your partner, and use R to analyze those responses. You have already have extensive experince analyzing data with R, including:

1. Running scripts.
2. Storing data in R objects: mydata <- c(1,2,3,4,5)
3. Working with dataframes
4. Performing basic computation: mean(), sd()
5. Transforming data: qnorm(), pnorm(), log()
6. Fitting linear models: lm(), lme()
7. Bootstrap analysis: bootES()
8. Visualizing data: plot()
9. Using functions: function() {}

Today’s lab will have 4 phases: (a) preparation for data collection, (b) data collection & upload, (c) individual data analysis, (d) group data analysis (maybe)

**LAB INSTRUCTIONS**

**Preparation for data collection**

*“Measure twice, cut once.”* said every grandpa, ever.

We need to double check a couple things before we collect data so that your analysis goes more smoothly. Remember: settings are spelling & case-sensitive!

|  |  |
| --- | --- |
| **3.1** | **Download your experiment to this computer (the ZIP file from last week).**   1. This works faster if each member of the team downloads and runs the experiment on their own computer. Make this happen. 2. Verify that the following files and folders are present, and in this basic configuration (your filenames may vary, *but the file structure should not*!):   Macintosh HD:Users:stevenparker:Desktop:Screen Shot 2015-09-29 at 11.45.52 AM.png |
| **3.2** | **Open the conditions file (\*.xlsx), verify that it has the changes you made last week**   1. You added values to the columns: corrAns, and congruent 2. Verify that these values are correct for each row!!!    1. Are the responses in the corrAns column (left, down, right) correct, given the values in the letterColor column (red, green, blue)?    2. Are the values in the congruent column correct, given both the values in the text and letter Color column? (same=congruent, not same=incongruent) 3. After you’ve inspected your own file and made any necessary changes, have a neighbor from another team triple-check your conditions file, and you check *their* conditions file. 4. Save, close Excel. |
| **3.3** | **Open PsychoPy, then open (⌘O)** **your experiment file (\*.psyexp).**   1. Remember!!! Open PsychoPy first, *then* open the .psyexp file using either:    1. The menu (File > Open).    2. The keyboard shortcut (⌘O).    3. Dragging the .psyexp from the Finder to the PsychoPy icon. |
| **3.4** | **Verify that your experiment script still has all its parts (macro inspection).**   1. A trial routine 2. A practice routine that is similar to the trial routine for your subjects to get used to the buttons. 3. 3 routines containing text & keyboard components, that function to provide: (a) welcome & instructions, (b) a warning that the real trials are coming up, and (c) a warm heartfelt thank you at the end of the experiment. 4. Save. |
| **3.5** | **Verify the settings in each of the components (micro inspection). Most settings don’t matter much, but do make sure that:**   1. trial routine:    1. start times are correct    2. durations are blank    3. text component named “word”    4. keyboard component named “resp”    5. variables are preceded with a “$” ($letterColor, $text, $corrAns) 2. practice routine:    1. same as trial routine, except the names of text and keyboard components are different. 3. Save. |
| **3.6** | **In the trials Loop (the one around the trial routine):**   1. Using the “Browse…” button, reload the conditions.xlsx file (this is a preemptive step, intended to prevent a script crash) 2. set nReps to **10** 3. Save. |
| **3.7** | **In the practice Loop (the one around the p\_trials routine):**   1. Using the “Browse…” button, reload the conditions.xlsx file (this is a preemptive step, intended to prevent a script crash) 2. set nReps to **1** 3. Save. |
| **3.8** | **Briefly test your script on yourself.**   1. Run the script (⌘R), making a mix of correct and incorrect responses. 2. You don’t need to test the entire experiment, but do test at least through, p\_trial and a few of the actual trials. 3. Press escape to end the script.    1. Did the experiment work & look the way it is supposed to? If not, make necessary changes. 4. Open the CSV file, and verify that the responses were recorded, the responses were scored correctly (resp.corr column: 1=correct, 0=incorrect). If not, check the settings in keyboard object “resp” 5. Save. |

**Data collection**

|  |  |
| --- | --- |
| **3.9** | **Run the experiment to collect responses for all members of your team.** |
| **3.10** | **Inspect CSV file(s), verify the data, then upload CSV file(s) to D2L.** |
| **3.11** | **If you have more than one team member, make sure each team member gets their own data file for the analysis.**   1. Share file, flashdrive, email |
| **3.12** | **UPLOAD YOUR CSV TO D2L!!!1!** |

**Individual Data Analysis**

*“There's something that happens with the collection of a large amount of data when it's dumped into an Excel spreadsheet or put into a pie chart. You run the risk of completely missing what it's about.”* Aaron Koblin (digital media artist).

Time to make meaning from your personal pile of responses. Before you start, you need to make an analysis plan. The most common mistake people can make when analyzing data is also the worst: starting the analysis without a plan[[1]](#footnote-1). Let’s avoid this frustrating and time-consuming error by making an analysis plan! We need to explicitly state (a) the variables in our dataset (independent, dependent); (b) how these data tell us anything about our phenomenon of interest, in this case color-word interference effects.

Variables

Every experiment has at least 2 types of variable:

**Independent variable (IV):** In experimental settings, the stimulus condition whose values are free to vary independently of any other variable in the situation ([APA, 2015](http://www.apa.org/research/action/glossary.aspx?tab=9)). (in non-experimental studies, called predictor variable)

**Dependent variable (DV):** In an experimental setting, any variable whose values are the results of changes in one or more independent variables ([APA, 2015](http://www.apa.org/research/action/glossary.aspx?tab=4)). (sometimes called measurement variable)

*Not all IVs are created equal*, and it sometimes can get pretty confusing figuring out which variable is important to our research questions. I’ve invented a couple terms to help us think about independent variables in an experiment:

**Stimuli-level variable:** The way(s) that stimuli are systematically varied that allow measurement by varied responses[[2]](#footnote-2).

**Hypothesis-level variable:** The variable(s) in an experiment that allow the testing of a hypothesis[[3]](#footnote-3).

In perception research, hypothesis-level variables are the contrasting conditions used to test a hypothesis. The levels in a hypothesis-level variable tend to be few in number (2 or 3 levels), but the differences between those levels tend to be large. On the other hand, stimuli-level variables tend to have many levels, and those differences also tend to be small. While variability of stimuli is important to any hypothesis, the results of a stimuli-level variable tend to be uninformative to the bigger picture of the experiment.

|  |  |
| --- | --- |
| **3.13** | **In our experiment, what were the independent variables (IV)? (HINT: 3)**  >>> |
| **3.14** | **Of these IVs, which was/were so-called stimuli-variable(s)[[4]](#footnote-4) ?**  >>> |
| **3.15** | **Of these IVs, which was/were the so-called hypothesis-variable(s)[[5]](#footnote-5) ?**  >>> |
| **3.16** | **What was/were the dependent variable(s) (DV) in this experiment? (HINT: 2)**  >>> |

Expected Results

We’ve already seen the tilde math symbol ( ~ ) used to represent “dependent on.” For example:

*alertness ~ coffee\_consumption* (1)

According to Equation 1, mental alertness *is dependent on* an amount of coffee consumed (a testable hypothesis). Can you think up other silly examples that use the tilde (~)?

|  |  |
| --- | --- |
| **3.17** | **In our experiment, given the hypothesis-level variable and the dependent variable, use the tilde symbol to represent the relationship we’re testing.**  >>> |
| **3.18** | **Based on what you know about the Stroop Effect, what is the expected result of testing the above relationship?**  >>> |

Taking the time to explicitly (a) identify the function of variables, (b) define a mathematical relationship between those variables, and (c) the result of testing that relationship saves you time in the long run, but also ensures you understand what it is that your doing at every step during the analysis. For most people this type of reasoning takes time to develop, but is crucial to the research process. Time to compute some results! So… what do you want to compute?

Summary Statistics

|  |  |
| --- | --- |
| **3.19** | **Double-click** *Lab3\_Individual\_Data\_analysis.R* **to Open the data analysis script.** |
| **3.20** | **Run Part 1: Data Preparation.**   1. We’ve asked OIT to permanently include the bootES package, but in case that’s not working, you *may* have to reinstall bootES. You know what to do. 2. Load YOUR personal CSV |
| **3.21** | **Everything you need is stored in the dataframe** df.   1. Inspect the dataframe and identify the columns relevant to our analysis. Which are the independent variables? The dependent variables? 2. In the space below, make a little “data dictionary” that helps you remember what the data in each column mean. (HINT: resp.rt = response time)   >>> |
| **3.22** | **For this part, you’ll need to perform a basic computations using the values in the columns. (HINT: mean())**   1. What was your overall percent correct?   >>> |
| **3.23** | **On line 33 is a command that removes all the incorrect responses from the dataframe df. Right now it is “commented out” with a # so you could compute the % correct. You have a choice:**   1. Leave df *as is*, and include all responses (even the incorrect ones) in the subsequent analysis. 2. Remove the incorrect responses from df, and continue the analysis with only correct responses. 3. There are pros and cons to either route; which do you choose? Justify your decision!!!   >>> |
| **3.24** | **What summary statistics should you compute to figure out how you performed on this task? (HINT: M, SD)**  >>> |
| **3.25** | **Show those summary statistics below:**  >>> |
| **3.26** | **What do these results mean?**  >>> |
| **3.27** | **Response times typically have highly skewed distributions, so a common practice is to use the logarithm of the RTs in analyses. The log transform was already performed on line 29, but we need to report summary statistics of those data.**   1. Compute the same summary statistics you did before on the log-transformed data. 2. Show your results below:   >>> |

We’re now going to perform inferential statistical tests on your individual data to determine whether your responses resemble chance performance or not. We’ve never done this step at the level of individual data analysis, but the logic is similar as to when we perform these tests on the group data.

These tests are in Part 2: Compute the bootstrap confidence intervals and linear models. There are 6 tests we need to run (bs.lin, bs.lin.d, bs.log, bs.log.d, mod.lin, mod.log). I’ve broken the scripts, which you’ll need to fix. Luckily for you, the “breaks” are very similar to Lab 2. Instructions to fix the scripts are in the script itself. bs.lin is still intact, so use it as a model for the subsequent bootstrap models. After you have fixed the models, run them and copy-paste the outputs in the boxes below.

|  |  |
| --- | --- |
| **3.28** | **bs.lin**  >>>  What does this result mean?  >>> |
| **3.29** | **bs.lin.d**  >>>  What does this result mean?  >>> |
| **3.30** | **bs.log**  >>>  What does this result mean?  >>> |
| **3.31** | **bs.log.d**  >>>  What does this result mean?  >>> |
| **3.32** | **mod.lin**  >>>  What does this result mean?  >>> |
| **3.33** | **mod.log**  >>>  What does this result mean?  >>> |

Plotting Results

Part 3: Plotting Section prepares 8 separate plots of the data, but the final 10 lines combines all the plots to display as a single plot. Run all of Part 3, and put this combined plot below:

|  |  |
| --- | --- |
| **3.34** | **Copy-paste the combined plot here:**  >>> |
| **3.35** | **Is there anything shown in these plot that is unexpected, or does not agree with the statistical tests we’ve run?**  >>> |

By now, you should be able to state conclusively whether your responses are consistent with the Stroop Effect, and substantiate that claim with (a) summary statistics and (b) statistical tests. To get to this point, you built a computer program from scratch (PsychoPy experiment), made some analytical decisions, and wrote code in R to test experimental hypotheses... not too shabs!

Next week, we’ll analyze the results from both lab sections and see how generalizable these responses were to the broader unobserved population.

1. The second most common occurs once the analysis is started: analyzing too fast. [↑](#footnote-ref-1)
2. Orientation exp: test orientations; Loudness exp: tone loudness’ [↑](#footnote-ref-2)
3. Orientation exp: standard orientations; Loudness exp: waveform. [↑](#footnote-ref-3)
4. Ways that stimuli varied, but are/were not most relevant to our hypothesis. Said another way, the variables were common to the entire experiment. [↑](#footnote-ref-4)
5. The variable(s) central to testing a hypothesis in an experiment. [↑](#footnote-ref-5)