Basic Overview

In the experiment from Lab 1, we had each person in class study a list of 24 word pairs. One word in each pair was lower case and the other upper case. We gave people 2 minutes to study the list. The next two phases were counterbalanced across participants so we could factor out order effects. In one test phase, we gave everyone a list of 96 words. Some of these words were completely new (NOT on the first list) and others were the 24 uppercase words from the first list (but not in uppercase this time). People had 2 minutes to check the ones they recognized from the first phase. In the other test, we gave people the same list of 24 words from the first phase, but this time the uppercase words were missing (just saw the lower case words and a blank). Participants were asked to write in the upper-case
 they saw in the first list in the blank (again had 2 minutes to fill it out). We scored the responses in terms of the following conditions:

> ### Recognition Task (96 items):
> Hit = correctly remembered when you should have
> FA = recognized but actually wasn't on the list (false recognition)
> CR = correct rejections, not on study list and not checked (correctly said old)
> Miss = was on study list but did not check it
>
> ### Recall Task (24 items)
> Hit = correctly recalled
> FA = wrong word recalled
> Miss = blank/not filled out
> (no CR here because no rejections)

In addition, 1/2 of the study pairs were strong associates (e.g., BABY-BOTTLE) while others were weaker (e.g., COMPUTER-GRASS). We separately computed hit rates for strong and weak items to assess the affect that prior knowledge/associations might have had on people's performance (you might be able to guess the other word using prior knowledge for strong associates, but this might be less effective for weak associates.

Psychological Question

What process do people use to remember? There are certainly more than one kind of remembering. Recognition refers to the feeling of remembering you get when you see something that you've seen before. This is like walking on the street and seeing someone you know from high school. We say you recognize them. Another kind of remembering is recall. In this process, you generate the information from your memory. For example, writing down your phone number is recall. Choosing your phone number out of a list of possible phone numbers is recognition. One question addressed in Lab 1 is are these two processes completely different?

One theory described in the Tulving and Thompson (1973) paper is that recall consists of two mental phases. First you generate a list of possible items in your head, then you look at the list you generated in your head and see if you recognize any of them. So everything is really about recognition in the end, but there is a "generation" step where you mentally create possible targets to remember. So to do the recall task from Lab 1 when you see a lower-case cue, you mentally think of a couple possible words and then see which of these you recognize. In the recognition part of the lab, you don't have to generate the list of things to remember because it is already given to you on the paper. You just have to check off each word you recognize. The implication of this theory is that recall performance should always be at least as good as, but usually worse than, recognition. Thus, your performance in recognition should be as good as or better as recall. If we find better performance in the recall phase than in recognition, that is evidence that

recall is not simply recognition applied to a internal "mental list". *Note: re-read the section on page 357 titled "Generation-Recognition Models" in the Tulving & Thompson (1973) paper to review.*

As an alternative to this account, Tulving & Thompson (1973) propose an "encoding specificity principal" whereby what you remember depends on how it was studied. By this account, performance on a remembering task will be better if the study task is similar to the test task. You may have been told by your parents at some point to not study while listening to loud music because when you go take a test, the same loud music won't be playing. Basically you want the study and test situations to be as similar as possible. By this account, recall should be better then recognition partly because you didn't just store in mind the words on the page in the study phase as much as you encoded the relationship between the pairs of words. Thus, it is easier to remember the target words when you get one of the pairs than if you just get a list of words alone. The Mind Hacks hint #12 talks about this a bit as well. *Note: re-read the section on page 359 titled "Encoding specificity principal" in the Tulving & Thompson (1973) paper to review. Also, re-read the section titled "Logic of Experimental Comparison between Theories".* The central idea is you don't even store "words" in memory as much as you register the situation or context in which you studied the words. Remembering cues that are most effective are those that bring you back to the same context in which you studied the items.

Analysis Steps (to be completed in class Monday & Wednesday):

1. You first want to compute descriptive/summary statistics of the hits, false alarms, etc... for both phases of the experiment. **This includes the mean, SD, std. error of the mean, and median values.**

2. Next, we want to know if the HIT RATE in recall and recognition was different (and what direction, was recall or recognition better?). This gives us some basic indication about if people were better in one condition or other. This is accomplished with a t-test on the differences hit rates (i.e., a within-subjects paired t-test). If performance is lower on the recognition test than on the recall test this is a problem for the Generation-Recognition account described above. However, it would seem supportive of the "encoding specificity principal". **Remember this involves computing a paired t-test between recall and recognition hit rates (basically a one sample t-test on the differences between the two conditions). You should compute the t-value by hand using the procedure we described in the previous exercise, and then convert this into a p-value.**

3. Next we want to know if people in the two tasks (recall/recognition) were simply more likely to say yes/ try to remember something (i.e., a criterion shift). For example, if people in the recognition task were simply more willing to say 'yes i remember that' then we should see a higher rate of FALSE ALARMS. **Thus we can compare the false alarm rate between the two conditions (again with a paired t-test).** The hope is that the HIT rate was significantly different between the two testing conditions, but the FALSE ALARM rate was similar. This is t-test #2 for the lab.

4. Next, we want to analyze separately if the hit rate was higher for the SRONG associates compared to the WEAK associates for both recall and recognition. Again, this can be done with a simple paired t-test similar to the one we conducted in Exercise 2 and as you did above. This is t-test #3 and #4 for the lab.

5. Next, we want to know if the order of the test had any effect. To do this we will compare the two halves of the data (group 1 and group 2) to see if there was bias towards higher or lower performance (i.e., hit rates) on the first or second test. This is t-test #5 for the lab and should be a unpaired t-test (we will talk about how to do this in class wednesday).

6. You should make a bar plot (using excel) of the hit rates in recall and recognition, along with the false alarms. The plot should include error bars that show the standard error of the mean.

7. Finally, we will want to conduct a signal-detection analysis of performance in the recognition task. There is not direct analog of signal-detection for recall (since there are no correct rejections). However, we will use this as an exercise in our "Bootcamp" sequence where we learn tools to help you do your final projects. For this you will want to compute a d' and beta estimate (discussed in class) for each subject and report the means and standard deviations of these values in your paper. As mentioned in class (wednesday) signal detection is a method for measuring performance (in perception and memory tasks) where the quality of the method is strong.

This list summarize the data you are expected to present in your paper. You will be working on this for 2-3 days in class along with lectures that help you through new stages of the analysis.