Exam 3 PS 217, Spring 2010

1. Bäuml, K.-H. T. and Samenieh, A. (2010) had an article published in *Psychological Science* titled *The Two Faces of Memory Retrieval*. They write:

Does retrieval of a specific memory affect recollection of related memories? For instance, does selective remembering of part of a traumatic experience, or part of an incidentally observed crime, affect memory for other details of the event? Casual subjective experience suggests that memory retrieval can improve recollection of related memories. When one is talking with a friend about a common, long-forgotten vacation, remembering a first piece of the event often initiates a chain of retrieval processes, along which more and more of the seemingly forgotten memory is recollected. However, this subjective experience contrasts with scientific experiments that have demonstrated that selective remembering typically impairs recollection of related material.

In their experiment, subjects studied a list of 16 words, one at a time, for 5 sec each. Of the 16 words, four were targets (each beginning with a unique letter, e.g., apple, carrot, mouse, table) and the remaining 12 words were nontargets (each beginning with different unique initial letters, e.g., banana, sandwich). After studying the list, half of the 80 total participants were told to remember the items in the list for a later test and half were told to forget the items in the list, because they weren't needed for a subsequent test.

The other factor was the number of nontarget items in an initial cued recall task. In each group of 40 participants (Remember or Forget), ten were first asked to recall 0 of the nontarget words before recalling the target words, ten were first asked to recall 4 of the nontarget words before recalling the target words, ten were first asked to recall 8 of the nontarget words before recalling the target words, and ten were asked to first recall all 12 of the nontarget words before recalling the target words. For the nontarget words, the participants were given a prompt of a couple of initial letters of the word (e.g., ba---- for *banana* and sa------ for *sandwich*). For the target words, the participants were given a prompt of the initial letter of the word (e.g., a---- for *apple* and m---- for *mouse*).

The dependent variable was the percentage of the target words correctly recalled. Complete the SPSS table below and then interpret the results of the study as completely as you can. [15 pts]

Descriptive Statistics

Dependent Variable:Target Items Recalled

Task	Number of Items	Mean	Std. Deviation	N
Remember	0	48.9000	4.12176	10
	4	44.6000	2.98887	10
	8	35.3000	2.49666	10
	12	20.8000	2.74064	10
	Total	37.4000	11.32005	40
Forget	0	23.8000	2.74064	10
	4	29.8000	3.67575	10
	8	35.3000	2.49666	10
	12	39.2000	3.91010	10
	Total	32.0250	6.66213	40
Total	0	36.3500	13.31906	20
	4	37.2000	8.26279	20
	8	35.3000	2.43007	20
	12	30.0000	9.99474	20
	Total	34.7125	9.61696	80

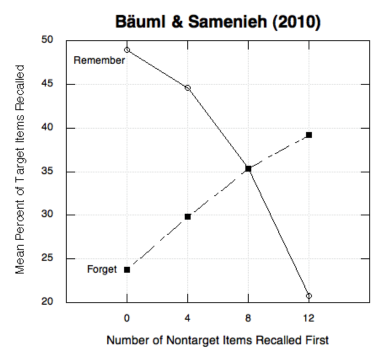
Levene's Test of Equality of Error Variances[a]

Dependent Variable:Target Items Recalled

F	df1	df2	Sig.
.788	7	72	.600

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Task + Items + Task * Items



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power[b]
Task	577.8	1	577.8	56.09	.000	.438	1.000
Items	628.4	3	209.5	20.34	.000	.459	1.000
Task * Items	5360.2	3	1786.7	173.5	.000	.879	1.000
Error	739.9	72	10.3				
Corrected Total	7306.3	79					

The Levene test is not significant (p = .6), so there's no concern about heterogeneity of variance. Moreover, because each F is significant with p < .01, each effect would be significant even with a conservative alpha level of .01.

To assess the interaction, you'll need a post hoc test:

$$HSD = 4.42\sqrt{\frac{10.3}{10}} = 4.49$$

There was a main effect for Task, $F(1,72) = 56.09$, $MSE = 10.3$, $p < .001$, $\eta^2 = .438$. There was a main effect for Number of Items, $F(3,72) = 209.5$, $p < .001$, $\eta^2 = .459$. There was also an interaction between Task and Number of Items, $F(3,72) = 1786.7$, $p < .001$, $\eta^2 = .879$. Post hoc tests using Tukey's HSD indicate that after first recalling 0 or 4 non-target items, people in the Remember condition recalled more target items ($M = 48.9$ and $M = 44.6$) than people in the Forget condition ($M = 23.8$ and $M = 29.8$). However, after first recalling 8 non-target items, there was no difference between the Remember ($M = 35.3$) and the Forget ($M = 35.3$) conditions. Moreover, after first recalling 12 non-target items, people in the Remember condition recalled significantly fewer targets ($M = 20.8$) than those in the Forget condition ($M = 39.2$).

2. If the above ANOVA had been computed as a one-way ANOVA on Number of Nontarget Items Recalled (Items), how would the source table change (illustrate below)? [5 pts]

Source	*SS*	*df*	*MS*	*F*
Item (Between)	628.4	3	209.5	2.38
Within (Error)	6677.9	76	87.87	
Total	7306.3	79		

3. In the two-way ANOVA in Problem 1 and the one-way ANOVA in Problem 2, you should be able to estimate the population variance (σ^2).

What would be your best numerical estimate of the population variance in Problem 1?	10.3
What would be your best numerical estimate of the population variance in Problem 2?	87.87

Given that the data are identical in the two problems, why might your estimate differ? [5 pts]

You're actually estimating two different populations. For the original problem, you're estimating a population in which the type of task is dealt with separately (each sample variance used to estimate the population variance is based on variability in scores from people who all received the same task (either Remember or Forget) and the same number of non-target items recalled (either 0, 4, 8, or 12). For the single-factor problem, you're estimating a population in which the type of task in which people are engaged (Remember or Forget) are combined. Thus, if there's any variability in recall of target items due to the task, you'd expect greater variability in this population. It's also the case that the sample variances being used to estimate the population variance are all from people who had either the Remember or Forget task—all thrown in together.

4. Under which conditions might it make more sense to compute Spearman's *rho* than Pearson's *r*? [2 pts]

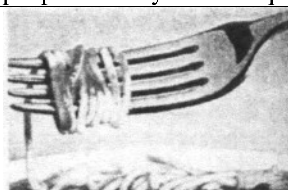
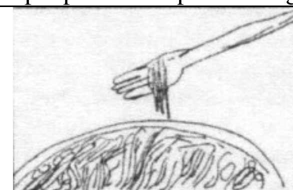
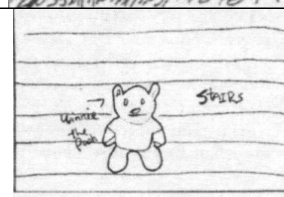
When your data are ordinal, or when one or both variables may be non-linear.

5. The full name is Pearson product-moment correlation coefficient. You can compute the statistic (*r*) as the product of two "moments." Well, actually, the average (mean) of the product of two moments. What two values might you

multiply, then take an average of the products in order to compute *r*? (Of course, you would never actually compute *r* this way.) [2 pts]

It's the product of two *z*-scores (one for X and one for Y).

6. Helene Intraub (University of Delaware) has conducted lots of research on a phenomenon she calls *boundary extension*. In study after study, she and her students have shown that people will report having seen more of a picture than had actually been presented, as seen below:

What people actually see at acquisition	What people later report having seen
	
	

This year, Ben Glicksberg conducted his senior thesis research with me. He examined possible limitations on the boundary extension effect. Essentially, Ben had three conditions:

Condition	First phase	Second phase	Test phase
Control	briefly look at each picture (as in typical Intraub)	distractor task	draw each picture
Simultaneous Drawing	*while* looking at each picture, construct a rough drawing	distractor task	draw each picture
Post Viewing Drawing	*after* looking at each picture, construct a rough drawing	distractor task	draw each picture

As you might intuit, Ben reasoned that if people drew the picture either while it was being presented (Simultaneous) or immediately after seeing the picture (Post), then they might later exhibit less boundary extension than Intraub typically finds. And, of course, the Control condition should exhibit the typical boundary extension effect.

Ben looked at his data in a number of different ways, but here's one type of data he analyzed. The dependent variable was the percentage of the original picture represented in the final drawing. Thus, if the original object took up 68% of the original frame and a participant drew the object so that it took up 68% of the frame at test, then that person had a score of 1.0 and exhibited no boundary extension. However, if the participant drew the object so that it took up 34% of the frame at test, then that person had a score of .5 and exhibited boundary extension. (That is, as seen in the examples above, the object takes up a smaller proportion of the total picture than was true in the original picture.)

Below is an SPSS output from Ben's data. Complete the source table and then interpret his results as completely as you can. Do Ben's results confirm his expectations? [15 pts]

Descriptive Statistics

Dependent Variable:Rat

Condition	Mean	Std. Deviation	N
Control	.6806	.06059	10
Post	.9507	.08107	10
Simult	.9561	.04608	10
Total	.8625	.14473	30

Levene's Test of Equality of Error Variances[a]

Dependent Variable:Rat

F	df1	df2	Sig.
1.579	2	27	.225

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Condition

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power[b]
Condition	.496	2	.248	62	.000	.817	1.000
Error	.108	27	.004				
Corrected Total	.604	29					

The Levene test indicates that there's no reason to be concerned about heterogeneity of variance (*p* = .225). It's also the case that your *F* is large and would be significant even with a conservative alpha level.

You need to compute a post hoc test: $HSD = 3.5\sqrt{\frac{.004}{10}} = .07$ $HSD = 3.5\sqrt{\frac{.004}{10}} = .07$

There is a significant effect of Condition, $F(2,27) = 62$, $MSE = .004$, $p < .001$, $\eta^2 = .817$. Post hoc tests using Tukey's HSD indicate that people in the Post Presentation Drawing (*M* = .951) and Simultaneous Drawing (*M* = .956) later drew the object as closer to the original percentage (i.e., 1.0) than people in the Control condition (*M* = .681).

7. Of course, Ben actually had each subject view a number of different pictures of objects. Looking just at the ten people in the Post Condition (drew a rough picture after briefly seeing the picture of an object), the data below represent their percent of the original object size (as above, that is values of 1.0 mean no boundary extension, values less than 1.0 mean boundary extension, and values greater than 1.0 mean boundary restriction) for the eight objects they saw (BananaRat means the ratio for drawings of the bananas). Complete the analysis below and interpret the results as completely as you can. Did participants vary in their drawings of particular objects? [15 pts]

Descriptive Statistics

	Mean	Std. Deviation	N
BananaRat	1.2183	.32420	10
FlowerRat	1.0526	.10675	10
PanRat	.8699	.13860	10
TireRat	1.0144	.15152	10
ForkRat	.7679	.17453	10
PailRat	.9007	.07988	10
BootRat	.9081	.11380	10
HatRat	.8736	.08407	10

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power[a]
Objects	Sphericity Assumed	1.36	7	.194	8.31	.000	.482	1.000
Error(Objects)	Sphericity Assumed	1.47	63	.023				

You'd need to compute a post hoc test: $HSD = 4.44\sqrt{\frac{.023}{10}} = .21$

There was a significant effect of the type of object, $F(7,63) = 8.31$, $MSE = .023$, $p < .001$, $\eta^2 = .483$. Post hoc tests, using Tukey's HSD, indicate that banana (*M* = 1.218), flower (*M* = 1.052), and tie (*M* = 1.014) exhibit significantly less boundary extension than fork (*M* =

.768). Banana exhibits significantly less boundary extension than pan (*M* = .870), pail (*M* = .901), boot (*M* = .908), and hat (*M* = .874).

8. Eating disorders are often common occurrences on college campuses, particularly for females. One hypothesis is that problems with same sex peer relationships may be associated with bulimic behaviors (Schultz & Paxton, 2007). A psychologist wanted to test this hypothesis by using the friend-trust subscale from the Inventory of Parent and Peer Attachment (IPPA) Scale (Armsden & Greenberg, 1987) as a measure of quality of peer relationships. This 10-item summated rating scale includes statements similar to, "My friends respect my feelings" with responses on a 5-point rating scale ranging from 1 (never) to 5 (always). Thus, scores on the scale may range from 10 to 50. Bulimic symptoms were measured using the bulimic symptoms subscale from the Eating Disorder Inventory (Garner, Olmsted, & Polivy, 1983). This 7-item summated rating scale assesses bulimic behaviors using a 6-point rating scale with scores ranging from 1 (not at all symptomatic) to 6 (most symptomatic). Scores on this scale may thus range from 7 to 42. {Kiess & Green}

Below is an SPSS analysis of a set of data. Interpret the results as completely as you can. If a person had a score of 30 on the Attachment (IPPA) scale, what would you predict that person's Bulimic Symptom score to be? If a person had a score of 50 on the Attachment (IPPA) scale, what would you predict that person's Bulimic Symptom score to be? Does it appear that creating more attachments with peers is a good way to combat bulimia? [10 pts]

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.667[a]	.444	.394	6.65268

a. Predictors: (Constant), Attachment

ANOVA[b]

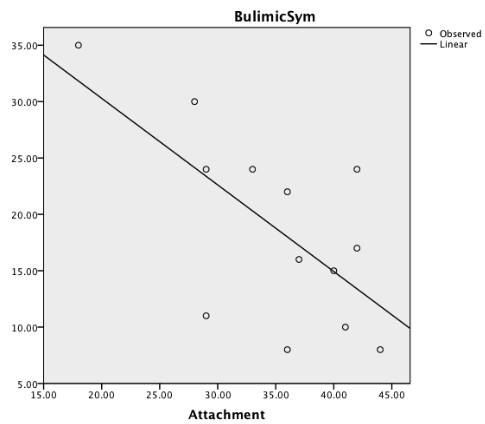
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	389.468	1	389.468	8.800	.013[a]
	Residual	486.840	11	44.258		
	Total	876.308	12			

a. Predictors: (Constant), Attachment
b. Dependent Variable: BulimicSym

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	45.656	9.249		4.936	.000
	Attachment	-.768	.259	-.667	-2.966	.013

a. Dependent Variable: BulimicSym



There is a significant negative linear relationship, *r*(11) = -.667, *p* = .013. The effect size, using the coefficient or determination is r^2 = .444.

The regression equation is: $\hat{Y} = -.768X + 45.656$ $\hat{Y} = -.768X + 45.656$

Thus, if Attachment = 30, you would predict a Bulimic Symptom score of 22.62.
If Attachment = 50, you could predict a Bulimic Symptom score of 7.256 if the linear trend continued. Otherwise, you would simply decline to make a prediction, given that you didn't observe a person with an Attachment score of more than 45.

Attachments may be a good way to combat bulimia, but you can't make that claim based on this correlational study. It may, in fact, be that having greater bulimic symptoms lead you to create fewer attachments. Or it could be that there is some other variable (e.g. self esteem) that leads to the observed relationship.

9. You rub on the brass statistical lamp and a genie pops out. If the genie will grant you only one wish as you conduct research, what would you wish for? [1 pt]

I'd ask for power!