The Skidmore Honor Code is in effect for this exam, as always. You'll be asked to write out the Honor Code statement at the end of the exam. Read each question carefully and answer it completely. Show all your work, so that you may receive partial credit. Keep in mind that a point is roughly equal to a minute, so don't spend too much or too little time on your answer to any question. Good Luck!

1. Dr. I. P. Freeley was interested in replicating the Middlemist et al. study (effect of invasion of personal space on time to urinate), but using a repeated measures design for greater power. For several days, he had a rotating cadre of confederates (so that the students wouldn't think that someone was stalking them) follow a set of students enrolled in an introductory psychology class. Whenever one of these students would enter a restroom to urinate, a confederate would check to ensure that no one else was using a urinal. If the participant were alone at one of the urinals, the confederate would either: 1) go to the urinal immediately next to the student (Near Stall); 2) go to a urinal one urinal away from the student (Distant Stall); or would simply go to the mirror and comb his hair (Alone). The dependent variable, as in the Middlemist study, was the time (in minutes) between when the unwitting participant unzipped his pants and when he began to urinate (micturate). Complete the source table below and interpret these data as completely as you can. [10 pts]

ANOVA Table for Distance

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Subject			.13				
Category for Distance				111.2	<.0001	222.433	1.000
Category for Distance * Subject			.01				

Means Table for Distance
Effect: Category for Distance

	Count	Mean	Std. Dev.	Std. Err.
Alone	10	.550	.227	.072
Distant Stall	10	.560	.196	.062
Near Stall	10	1.140	.259	.082

2. Mook argues that external validity is not always the purpose behind psychological research. For each of the studies below, indicate why the study is not externally valid, then why it's not a concern, given the intentions of the researcher(s). [10 pts]

Study	Why not externally valid	Why lack of EV is not a concern
Argyle (glasses and intelligence)		
Harlow (infant monkeys and drive reduction theory)		
Hecht (dark adaptation)		
Brown & Hanlon (parental role in grammar acquisition)		

3. Correlational designs do not allow you to make casual claims. Why not? Be very explicit about the difficulty of claiming that changes in one of the two variables in a correlational study *causes* the related changes observed in the second variable. We also discussed the shortcomings of using non-manipulated characteristics of the participants as "independent variables" in an experiment. How is this class of variable related to the notion of correlational designs? [10 pts]

4. Two researchers were interested in studying the effects of reward magnitude on performance. Both researchers used introductory psychology students as participants, the same total number of participants (21), the same type of reward and reward magnitudes ($1, $5, $20), the same apparatus, the same task, and the same performance measure (DV). One researcher used an independent groups design and, on the basis of the results, cannot reject the null hypothesis (that reward has no effect on performance). The other researcher used a repeated measures design and found a statistically significant effect of reward magnitude — larger rewards lead to better performance. Assume that neither study has a major flaw (e.g., repeated measures design is properly counterbalanced, random assignment to conditions). There are two fundamental reasons why the two researchers might have reached different conclusions. One reason concerns the sensitivity of the test of the null hypothesis. The other reason concerns the nature of the participant's experience in the two studies. Provide me with a clear explanation of the two reasons for the different results that the two researchers obtained. Would you trust the results of one study more than the other? Why? Finally, complete the source tables for the two experimenters seen below. [10 pts]

Independent Groups Design (F_{Crit} = 3.55):

Source	*df*	*SS*	*MS*	*F*
Treatment		28		
Error				
Total		100		

Repeated Measures Design (F_{Crit} = 3.23):

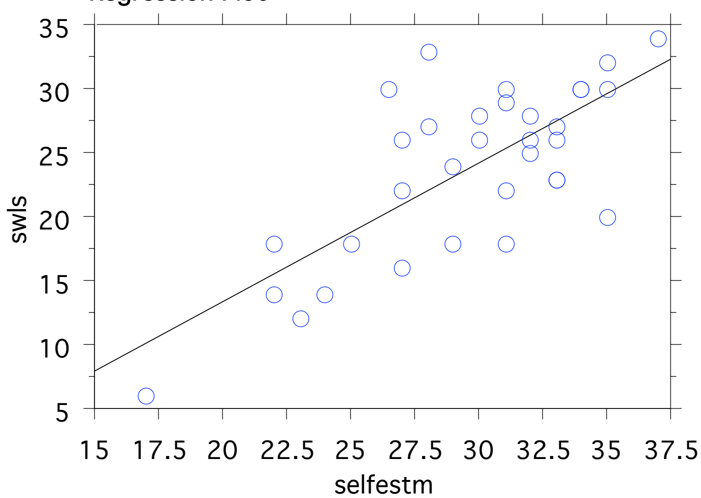
Source	*df*	*SS*	*MS*	*F*
Subject		100		
Treatment		20		
Error (Subj x Treat)				
Total		200		

5. In your first lab, there were a number of different personality measures. One was the Rosenberg Self-Esteem Scale (*selfestm*) and another was the Satisfaction with Life Scale (*swls*). Had you correlated those two measures, you would have seen an output like the one below. Interpret the output below as completely as you can. If a person had a self-esteem score of 30, what would you predict that person's SWLS score to be? What proportion of variance do these two measures share? If you were to talk about this result in a Discussion, what might you say about the relationship? [10 pts]

Regression Summary
swls vs. selfestm

Count	33
Num. Missing	0
R	.737
R Squared	.543
Adjusted R Squared	.528
RMS Residual	4.571

Regression Plot



Y = -8.202 + 1.081 * X; R^2 = .543

ANOVA Table
swls vs. selfestm

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	769.876	769.876	36.851	<.0001
Residual	31	647.639	20.892		
Total	32	1417.515			

Regression Coefficients
swls vs. selfestm

	Coefficient	Std. Error	Std. Coeff.	t-Value	P-Value
Intercept	-8.202	5.329	-8.202	-1.539	.1339
selfestm	1.081	.178	.737	6.071	<.0001

6a. First of all, imagine a repeated measures design with seven levels. Can you tell me *why* you'd need to counterbalance such a design, what kind of counterbalancing you'd use, and how many participants you'd need? What is the impact of counterbalancing on order and carry-over effects? [3 pts]

6b. OK, now let's assume that there is a particular order effect — a practice effect. That means that scores on the DV will improve over time as a result of practice. What is the impact on your error term (MS_{Error}) of counterbalancing? [2 pts]

7. In Lab 2, you saw a set of photo-arrays. As you know, each participant in that study rated each of the faces on the extent to which that face was a match for the eyewitness description (1 = "Poor Match" to 7 = "Great Match"). If the photo-array were unbiased, then the ratings of the six faces would be similar. To the extent that some faces were rated as less similar to the eyewitness description, then they were not really fair alternatives. As you know, each participant rated each of the six faces (making it a repeated measures design). However, it would be possible to use an independent groups design. Suppose that we used *n* = 64 in an independent groups design. Thus, there would have been 64 people rating the extent to which Face 1 (F1) matched the eyewitness description. Another 64 people would rate the extent to which Face 2 (F2) matched the eyewitness description. Etc. Complete the analysis below and then interpret these results as completely as you can. [10 pts]

ANOVA Table for Lineup1

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Face				49.4	<.0001	247.171	1.000
Residual			2.0				

Means Table for Lineup1
Effect: Face

	Count	Mean	Std. Dev.	Std. Err.
F1	64	3.078	1.462	.183
F2	64	4.172	1.486	.186
F3	64	4.125	1.363	.170
F4	64	2.234	1.342	.168
F5	64	5.984	1.327	.166
F6	64	3.734	1.606	.201

8. Below you'll find a repeated measures analysis of the exact same data (as in Problem 7). Focusing solely on the source table, and comparing it with the source table found in the prior problem, what can you tell me about the nature of the two *F*-ratios? How and why do they differ? Would you have expected such a difference? [5 pts]

ANOVA Table for Lineup1

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Subject	63	279.018	4.429				
Category for Lineup1	5	508.513	101.703	64.246	<.0001	321.228	1.000
Category for Lineup1 * Subject	315	498.654	1.583				

Means Table for Lineup1
Effect: Category for Lineup1

	Count	Mean	Std. Dev.	Std. Err.
L1.1	64	3.078	1.462	.183
L1.2	64	4.172	1.486	.186
L1.3	64	4.125	1.363	.170
L1.4	64	2.234	1.342	.168
L1.5	64	5.984	1.327	.166
L1.6	64	3.734	1.606	.201