

ID# \_\_\_\_\_

Final Exam

PS 306, Spring 2001

First of all, be sure to use your actual student ID#. I have it on my class lists if you're not sure of your number. Second, as on earlier exams, be sure to adhere to the Skidmore Honor Code. Third, remember that I think of a point as a minute, so you should allocate your time appropriately. Fourth, read each question carefully and answer it completely, showing all your work. Finally, let me extend to you my thanks for a pleasant semester. Have a great summer. And, if you're not a senior, then I look forward to seeing you next year.

P.S. In deference to Emily's attention span, the exam is shorter than normal. ☺

1. Historically, psychology may be characterized as the study of the behavior of white males. A while ago, there was a book titled *Even the Rat was White*, illustrating the bias present in psychological research. Of course, such bias is a two-edged sword. What is *good* about collecting data from white males only? What is *bad* about collecting data from white males only? [5 pts]

2a. Two researchers were interested in studying the effects of reward magnitude on performance. Both researchers draw from the same pool of participants, use the same total number of participants (24), 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 uses 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 uses a repeated measures design and finds 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). There are three fundamental reasons why the two researchers might have reached different conclusions. One reason concerns the sensitivity of the test of the null hypothesis. Another reason concerns the nature of the participant's experience in the two studies. A final reason has to do with the tentativeness of hypothesis testing, regardless of whether or not a researcher rejects the null hypothesis. Provide me with a clear explanation of the reasons that the two researchers may have reached such different conclusions. Would you trust the results of one study more than the other? Why? [15 pts]

2b. Finally, complete the source tables for the two experiments, as seen below. {Remember, the RM design is more efficient, so participants generate more than one piece of data.} [10 pts]

**Independent Groups Design:**

Source	SS	df	MS	F
Between Treatments	24			$F_{\text{Crit}} = 3.47$
Within Treatments (Error)				
Total	98			

**Repeated Measures Design:**

Source	SS	df	MS	F
Between Treatments	14			$F_{\text{Crit}} = 3.20$
Between Subjects	88			
Error (Residual)				
Total	194			

3. Suppose that you were interested in conducting a 4x6 mixed design, with the first variable (with 4 levels) an independent groups factor and the second variable a repeated measures variable. Suppose, further, that you needed to obtain a minimum of 35 pieces of data per condition (due to power considerations). What's the minimum number of participants that you would need for your study? [5 pts]

4. We discussed the study that Doob & Wood (1972) conducted, reported in "Catharsis and aggression: Effects of annoyance and retaliation on aggressive behavior." As they state, "The catharsis hypothesis of aggression usually refers to a decrease in aggression after the expression of aggression. The assumption as stated by Freud is that there is a certain amount of aggression that has to be expressed, and that once this has happened, there is less left to be expressed later on."

As you may recall, half the participants were first annoyed by the confederate (who made all sorts of nasty personal comments), and the other half were not annoyed by the confederate. Then one-third of each group: (1) shocked the confederate, who had been moved into an adjoining room, for errors in a learning task, (2) watched the experimenter shock the confederate, who had been moved into an adjoining room, for errors in a learning task, or (3) saw nothing, as the experimenter took the confederate into an adjoining room to administer the "learning test." In the final phase of the experiment, participants judged the "creativity" of associations made by the confederate to words presented by the subject. If they thought that a response was uncreative they gave the confederate a shock. They were told to increase the duration of the shock for increasingly uncreative responses. Confederates (who did not really receive any shocks) were cautioned by the experimenter not to comment if a shock was delivered, and then the experimenter left the room.

Doob & Wood were interested in testing the hypothesis that participants would experience a cathartic effect from observing the boorish confederate being shocked, or shocking her themselves. This would lead them to be less likely to shock the confederate in the final (judging creativity) phase of the experiment.

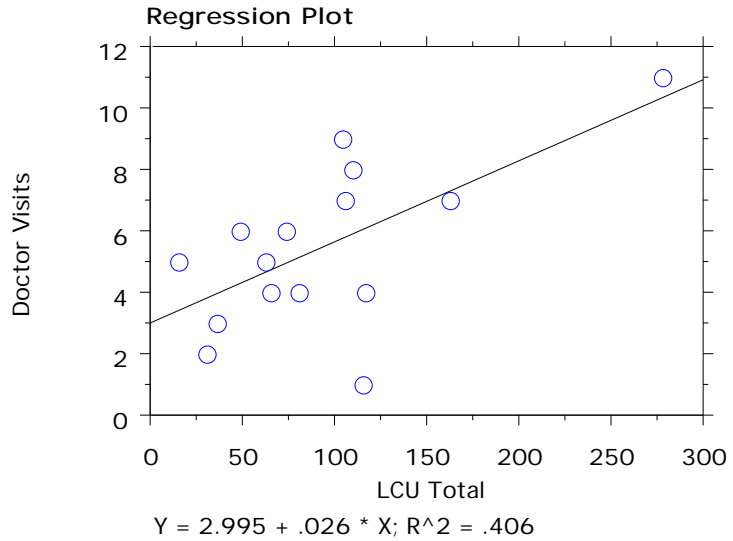
Participants were given an informed consent form that told them only that they were in a learning experiment and that they might be given electric shocks. They were told that they could withdraw from the experiment, etc. Finally, the participants were debriefed about the fact that the other "participant" was actually a confederate of the experimenter and that they had not really shocked the confederate. If you were on a participant review board presented with a proposal for this experiment, what would be your reaction, and why? Be sure to draw specifically on the APA guidelines. [5 pts]

5. Briefly, but clearly, describe the design and major results of the two studies by Ross, Lepper & Hubbard (1975) on perseverance in self-perception and the effectiveness of debriefing. What are the implications for ethical treatment of human subjects? [10 pts]

6. Studies have suggested that the stress of major life changes is related to subsequent physical illness. Holmes and Rahe (1967) devised the Social Readjustment Rating Scale (SRRS) to measure the amount of stressful change in one's life. Each event is assigned a point value, which measures its severity. For example, at the top of the list, death of a spouse is assigned 100 life change units (LCU). Divorce is 73 LCUs, retirement is 45, change of career is 36, the beginning or end of school is 26, and so on. The more life change units one has accumulated in the past year, the more likely he or she is to have an illness. The following StatView analyses show the results from a hypothetical set of data. Interpret these results as completely as you can. For these data, if a person had accumulated 100 LCUs, how many doctor visits would you predict? Provide three possible interpretations for the observed relationship. [15 pts]

**Regression Summary**  
**Doctor Visits vs. LCU Total**

Count	15
Num. Missing	0
R	.637
R Squared	.406
Adjusted R Squared	.360
RMS Residual	2.135



**ANOVA Table**  
**Doctor Visits vs. LCU Total**

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	40.469	40.469	8.877	.0107
Residual	13	59.264	4.559		
Total	14	99.733			

**Regression Coefficients**  
**Doctor Visits vs. LCU Total**

	Coefficient	Std. Error	Std. Coeff.	t-Value	P-Value
Intercept	2.995	.996	2.995	3.006	.0101
LCU Total	.026	.009	.637	2.979	.0107

7. Dr. Noah Weisser was interested in investigating the effects of sleep deprivation and alcohol on driving ability. To that end, he conducted a 3x3 independent groups study. Participants were given identical 32 oz. drinks that could contain 2, 4, or 8 ozs. of vodka (with the remainder of the glass filled with orange juice). Within each level of alcohol, one-third of the participants were deprived of sleep for 2 days, one-third were deprived of sleep for 3 days, and the final third were deprived of sleep for 4 days. Each of the participants operated a driving simulator for 30 minutes, during which the number of driving errors (objects hit, crossing into the oncoming lane, etc.) were recorded. Complete the analysis shown below and interpret the results as completely as you can. [20 pts]

**ANOVA Table for Errors**

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Alcohol			344.264		<.0001	335.932	1.000
Sleep Dep			230.889		<.0001	225.301	1.000
Alcohol * Sleep Dep			23.306		<.0001	45.483	1.000
Residual			2.050				

**Means Table for Errors**  
**Effect: Alcohol \* Sleep Dep**

	Count	Mean	Std. Dev.	Std. Err.
2 oz., 2 days	8	1.375	1.061	.375
2 oz., 3 days	8	3.250	1.035	.366
2 oz., 4 days	8	6.375	1.188	.420
4 oz., 2 days	8	3.500	.926	.327
4 oz., 3 days	8	6.875	.991	.350
4 oz., 4 days	8	7.750	.707	.250
8 oz., 2 days	8	7.250	1.282	.453
8 oz., 3 days	8	9.500	1.195	.423
8 oz., 4 days	8	16.500	3.071	1.086

8. Briefly define experimenter expectancy effects and demand characteristics. Then, tell me why experimenter expectancy effects might be considered demand characteristics. Use Rosenthal's study on the "early data returns" effect to illustrate your point. [15 pts]