

ID# _____

Exam 2

PS 306, Spring 2004

As always, the Skidmore Honor Code is in effect for this exam, so keep your eyes glued to your own exam. Read each question carefully and answer it completely (though not necessarily in perfect prose). Show all your work. Keep the point value of each question in mind (1 pt = 1 minute) so that you have sufficient time to complete the exam. Good Luck!

1. Briefly define the term *confound*. Then, using a *very explicit* example of practice effects (maybe even with numbers?), illustrate why conducting a repeated measures experiment without counterbalancing makes the study confounded. How does counterbalancing eliminate the confound? [10 pts]

2. OK Jeff, here's an interesting study. Gangestad, Simpson, Cousins, Garver-Apgar, and Christensen (2004) studied women over the course of their menstrual cycles to determine if they had a preference for male behavioral displays. I'll reconstruct their study as a two-factor independent groups design, while retaining the basic message of their article. Women watched a videotape of a male being interviewed. Half of the women saw the male respond to a question about himself ("Please tell me about yourself, including who you are, what you like to do, etc."). The other half of the women watched a videotape in which a male responded to a competitor for a date with a young woman (detailing why she should prefer to go on a date with him). For each video, one-third of the women responded on Day 3 of their menstrual cycle (a low fertility day). One-third of the women responded on Day 11 of their menstrual cycle (a high fertility day). Another third of the women responded on Day 21 of their menstrual cycle (a low fertility day). The dependent variable is a rating (on a 5-pt scale) by the women of the attractiveness of the male on a short-term basis. High scores indicate that the males were judged to be attractive for short-term sexual affairs. Complete the source table below and analyze these data as completely as you can. [15 pts]

ANOVA Table for Attractiveness

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Day of Cycle			7.6		<.0001	20.954	.992
Video			10.6		.0002	14.656	.979
Day of Cycle * Video		9.3			.0025	12.841	.906
Residual		61.2					

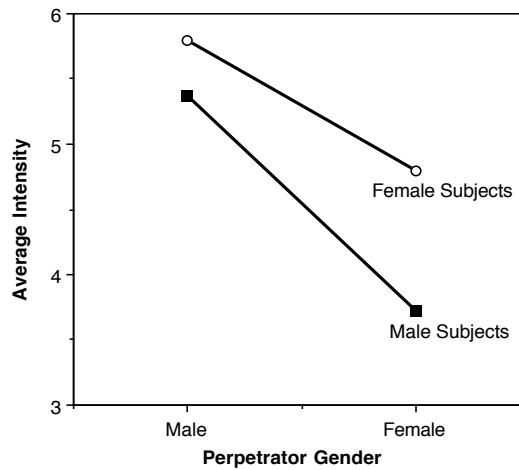
Means Table for Attractiveness

Effect: Day of Cycle * Video

	Count	Mean	Std. Dev.	Std. Err.
Day 11, Competitive	15	4.200	.941	.243
Day 11, Pers Info	15	2.600	.828	.214
Day 21, Competitive	15	2.533	1.060	.274
Day 21, Pers Info	15	2.333	.617	.159
Day 3, Competitive	15	2.800	.862	.223
Day 3, Pers Info	15	2.533	.743	.192

3. In a study by Baron, Burgess, and Kao (1991), male and female participants read accounts of stories that included a description of a sexist act perpetrated by either a male or a female against a female. The 193 participants described the perpetrator in a way that could be scored for intensity of sexist behavior. The displayed sexist behavior was rated 1 for *slightly displayed* to 7 for *extremely displayed*. Part of their *Results* section reads:

Perpetrator gender and participant gender main effects were both significant. Female participants, compared with male participants, gave more intense ratings to both male and female perpetrators...: $F(1,189) = 5.06, p < .03$Furthermore, male perpetrators were seen as displaying more intense gender bias than female perpetrators: $F(1,189) = 15.97, p < .001$. The interaction between participant gender and perpetrator gender was nonsignificant... $p < .34$...These results can be seen in the figure below:



Briefly interpret the meaning of these results, as you would in a discussion section. [10 pts]

4. Hmm. There's an article with the intriguing title, "Why people fail to recognize their own incompetence" by Dunning, Johnson, Ehrlinger, and Kruger (2003). According to Confucius, "real knowledge is to know the extent of one's ignorance." So, how well do you think that you'll do on this exam? Dunning, et al. (2003) asked students who were leaving an exam to judge how well they'd done on the exam. It turned out that students who performed the worst on the exam actually overestimated their performance and students who did the best on the exam were fairly accurate in their self-assessment (with a slight underestimation among the students with the best performance).

In one study, Kruger and Dunning (1999) gave additional information to some students, and that information had an impact on their judgments. Let's imagine a set of results that are consistent with their report. The dependent variable is the percent overestimation of a person's performance on an exam. So a score of zero is an accurate judgment. A positive score indicates overestimation and a negative score is an underestimation of one's performance. The students were divided into four groups based on their actual performance (Bottom Quartile, Second Quartile, Third Quartile, and Top Quartile). In addition, half of the students in each quartile were given a mini-lecture about the material after completing the exam (Add Info), but before making their judgments. The other half of each quartile was not given any additional information (No Info). Complete the source table below and interpret the results of this study as completely as you can. [15 pts]

ANOVA Table for Estimate

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Quartile		6184.6			<.0001	1251.531	1.000
Add Info		1008.2			<.0001	204.020	1.000
Quartile * Add Info		1308.1			<.0001	264.708	1.000
Residual		355.8					

Means Table for Estimate

Effect: Quartile * Add Info

	Count	Mean	Std. Dev.	Std. Err.
Bottom, Add Info	10	12.300	3.164	1.001
Bottom, No Info	10	30.400	4.300	1.360
Second, Add Info	10	9.300	1.418	.448
Second, No Info	10	20.900	2.079	.657
Third, Add Info	10	4.000	1.155	.365
Third, No Info	10	3.700	1.337	.423
Top, Add Info	10	-.300	.823	.260
Top, No Info	10	-1.300	.949	.300

5. Suppose that you are interested in studying the impact of a drug on maze learning in rats. Because you are unsure of the level at which the drug might be most effective, you decide to use 4 different levels of the drug. First of all, tell me (in very general terms) how you would determine the 4 levels that you would use in your experiment. You want to avoid carry-over effects of the drug, so each rat will be exposed to only one level of drug. Because you think that the drug may lead to better performance on some mazes than on other mazes, you want to run each of your rats through three different mazes (Easy, Moderate, and Difficult). Thus, this experiment would be a 4x3 mixed design. In *very explicit fashion*, tell me how you would run this experiment, including the minimum number of rats you'd need for your study and how many you'd actually use, the procedure you'd use, etc. [20 pts]

6. The state superintendent of instruction asks the director of educational research to investigate differences in scores on a standardized teacher examination for senior education students majoring in the following subject areas: English, Mathematics, Physical Education, and Vocational Education. The following results are from a random sample of 32 graduating seniors (16 males and 16 females). Complete the source table and interpret the results as completely as you can. [10 pts]

ANOVA Table for Score

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Major		647.34			<.0001	63.478	1.000
Gender		3.78			.5483	.371	.088
Major * Gender		37.34			.3238	3.662	.278
Residual		244.75					

Means Table for Score

Effect: Major * Gender

	Count	Mean	Std. Dev.	Std. Err.
English, Female	4	37.750	1.708	.854
English, Male	4	37.750	2.630	1.315
Math, Female	4	43.000	5.598	2.799
Math, Male	4	39.000	5.292	2.646
Phys-Ed, Female	4	35.500	2.082	1.041
Phys-Ed, Male	4	34.750	.957	.479
Voc-Ed, Female	4	27.750	1.708	.854
Voc-Ed, Male	4	29.750	2.062	1.031