ID# _____ Exam 2 PS306, Spring 1998

Read each question carefully and answer it completely. Once again, think of a point as a minute, so do not spend too much (or too little) time on any answer. Keep your eyes focused on your own exam.

1. In a study of hyperactivity among elementary school boys, nine groups of participants were randomly selected from a school population of ADHD, 7-year-old boys. (ADHD is Attention Deficits with Hyperactivity, and left untreated, it can prevent a child from attending to incoming learning stimuli and may also create major disruptions in the classroom.) The researcher wanted to study the classroom effects on the activity levels of the participants. Both the drug Ritalin as well as a behavior modification program served as factors. The drug was varied from 20 mg of Ritalin to 10 mg of Ritalin to no dosage (in the form of a placebo). The behavior modification program consisted of giving the child ten tokens to start the day and then taking away a token for each hyperactive infraction. The tokens that were saved could then be exchanged for some valued prize. The behavior mod program was varied from no program, to the program being used every other day, to the program being in force every day. After 4 weeks, all the children were evaluated for hyperactivity and were assigned scale scores ranging from a possible low of 0 (no indication of hyperactivity) to a high of 40 (extreme hyperactivity). Complete the source table below and interpret the data from this study as completely as you can. [25 points]

ANOVA Table for Hyperactivity Score

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Treatment	2	1283.143	641.571	26.292	<.0001	52.583	1.000
Drug	2	657.238	328.619	13.467	<.0001	26.934	.999
Treatment * Drug	4	128.762	32.190	1.319	.2746	5.277	.376
Residual	54	1317.714	24.402				

Means Table for Hyperactivity Score
Effect: Treatment * Drug

	Count	Mean	Std. Dev.	Std. Err.
BM Ev Oth Day, 10 mg	7	19.000	5.477	2.070
BM Ev Oth Day, 20 mg	7	17.143	5.429	2.052
BM Ev Oth Day, Placebo	7	22.429	7.413	2.802
BM Every Day, 10 mg	7	12.286	3.352	1.267
BM Every Day, 20 mg	7	8.714	3.592	1.358
BM Every Day, Placebo	7	14.429	3.645	1.378
No Beh Mod, 10 mg	7	22.000	3.958	1.496
No Beh Mod, 20 mg	7	16.429	4.276	1.616
No Beh Mod, Placebo	7	29.143	5.815	2.198

[Note that the output differs due to the differences between older and newer versions of StatView.]
Because there is no interaction, I would focus my attention on determining the source of the two main effects. For this design, I would actually use the same HSD for both main effects. That is, q = 3.42 (with 3 treatments and 54 df_{error}), so HSD = 3.69. (Note that n = 21 for each of the 3 means being compared.) The means for Treatment would be 19.5, 11.8, and 22.5 for Every Other Day,

Every Day, and No Beh Mod, respectively. The means for Drug would be 17.8, 14.1, and 22 for 10mg, 20 mg, and Placebo, respectively. Thus, for Treatment I would conclude that hyperactivity scores were lowest for the Every Day group compared to the other two group, but those two groups did not differ. I would also conclude that the 20 mg group was significantly less hyperactive than the 10 mg group, which was less hyperactive than the Placebo group.

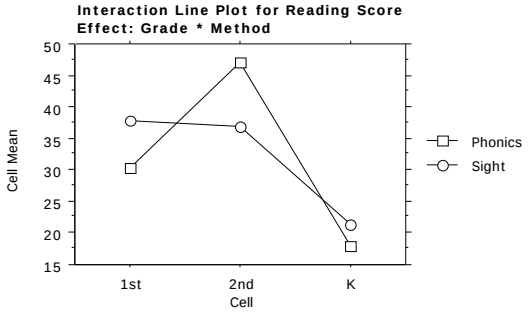
2. A researcher is interested in whether the phonics method of teaching reading is more or less effective than the sight method, depending on what grade the child is in. Twenty children were randomly selected from each of three grades: kindergarten, first grade, and second grade. Achievement was measured in terms of reading comprehension, with higher scores indicating more comprehension. Within each grade, 10 children were assigned to each of the two methods of teaching reading, Phonics or Sight. Analyze the results of this study as completely as you can. [25 points]

ANOVA Table for Reading Score

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Method	1	1.667	1.667	.142	.7080	.142	.065
Grade	2	5113.733	2556.867	217.503	<.0001	435.006	1.000
Method * Grade	2	872.133	436.067	37.095	<.0001	74.189	1.000
Residual	54	634.800	11.756				

Means Table for Reading Score
Effect: Method * Grade

	Count	Mean	Std. Dev.	Std. Err.
Phonics, 1st	10	30.200	3.490	1.104
Phonics, 2nd	10	47.000	3.091	.978
Phonics, K	10	17.800	3.521	1.114
Sight, 1st	10	37.800	3.490	1.104
Sight, 2nd	10	36.800	3.490	1.104
Sight, K	10	21.400	3.471	1.097



Because the interaction is significant, that's where I'd focus my attention. The figure showing the interaction leads me to predict that there will be no difference in K, then Sight might be better in 1st and Phonics better in 2nd. For my HSD, I would use q = 4.2, so HSD = 4.55. Thus, I would conclude that in K there is no difference between P and S,. However, in 1st grade, S > P. Furthermore, in 2nd grade, P > S. Thus, the impact of Method differs substantially as a child progresses from K to 1st to 2nd grade.

3. Suppose that you are interested in conducting a two-factor experiment investigating memory for words (and, to a degree, the word-frequency effect ☺). One factor is word frequency, which varies over 4 levels (Low Frequency, Moderately Low Frequency, Moderately High Frequency, and High Frequency). The other factor is the type of instruction that you provide the participants, which varies over 4 levels (No Instruction, Use Imagery to Learn the Words, Use Repetition to Learn the Words, and Form Anagrams of the Words). The dependent variable that you choose to use is pure recall of the words (percent of words recalled).

a. First, consider this as a completely between (independent groups design). If you want to have 25 scores per condition, how many total participants would you need? [2 pts]

Because this is a 4x4 independent groups design, you would have 16 cells. Thus, with 25 scores per cell you would need 400 participants.

b. Next, tell me how many participants you would need if you were to run this study as a completely within (repeated measures) design (keeping a minimum of 25 scores per condition)? Although you can tell me the number of participants needed, you should also tell me something about why doing the study as a completely within design may not make a great deal of sense, right? [4 pts]

With 16 conditions, you wouldn't even consider complete counterbalancing. Thus, you'd use incomplete counterbalancing, which means that you'd need to run multiples of 16 orders. To get above 25, you'd need to run 32 participants.

However, the way that participants "learn" the words doesn't lend itself to a repeated measures design. It would be cumbersome at best. More problematic is the possibility that the participants won't use the "method" that you tell them to use. It would be best to run this study as a mixed design with Instruction as a between (independent groups) factor, as seen below.

c. Finally, you should be able to conceive of this study as a mixed design. Tell me *briefly* how you would do so (which factor between, etc.), with a focus on the implications of the number of participants needed to provide you with a minimum of 25 scores per condition. [4 pts]

As above, Instruction would be an independent groups factor and Word Frequency would be a repeated factor. Thus, you would need to counterbalance a factor with 4 levels, which means that you could use complete counterbalancing (24 orders), but you'd need 48 participants for each of your levels of Instructions (to get above 25 scores/condition). Thus, you'd need 192 total participants.

Your 24 orders would come from every possible order of the 4 word frequencies:
L, ML, MH, H
L, ML, H, MH
L, MH, ML, H
L, MH, H, ML
L, H, ML, MH
L, H, MH, ML
ML, L, MH, H
Etc.

You would assign a participant randomly to one of the 4 levels of Instruction, and then (after giving them the appropriate instructions) you would give them the words in one of the 24 orders. The next participant would be randomly assigned to one of the remaining 3 levels of Instruction, and then that person would get the words in one of the 24 orders. Etc. [Note that I would be running the participants in randomized replications of the between factor and then using one of each of the counterbalancing orders.

4. In this (paunchy?) midsection of the course, we've been discussing confounds that may arise in the design of experiments. List 4 different confounds (your favorite ones?) with a brief definition and example of each. [15 pts]

As long as the definitions and examples were clear, I accepted a fairly wide range of answers to this question.