Final Exam PS 306, Fall 2004

1. Dr. Janet Oh spoke to our class as though she were delivering a lecture in another course. [That's not uncommon as job candidates come through the department.] However, her lecture did have a methodological focus. Briefly describe the lecture and tell me what you learned about the methodologies described in the lecture. [5 pts]

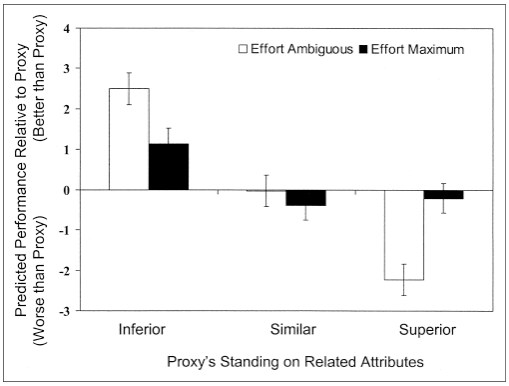
Idiosyncratic to this particular semester.

2. In order to assess the extent to which children would be credible eyewitnesses, psychologists have conducted a variety of studies in which they work to implant false memories in children. Using the APA Ethical Guidelines, and acting as a member of an IRB, what would you recommend about the following proposal for research. [10 pts]

At the outset, parental approval would be obtained, but (to keep the situation as natural as possible) the children are kept unaware that they are participating in an experiment. Half of the children will be going to their physicians for a routine check-up. The other half of the children are quite ill, which requires that their physicians cause the children quite a bit of pain as a result of their treatment on their office visit. Thus, although this variable is not manipulated, it allows the researcher to examine the extent to which the pain (and anxiety) that the children undergo has any effect on their susceptibility to false memories. Within the office visits for all children, they are touched by the physician and the nursing staff, but never in the genital area. (All the interactions are videotaped, so that what actually happens is recorded.) After the visit, adult experimenters interview all the children. Throughout the interview, for half the children in each group, the experimenters continually suggest that the physician had touched the children in the genital area (in an effort to implant a false memory). For the other children, the experimenters conduct an equally lengthy interview, but never mention that the physician had touched the children inappropriately. After assessing the extent to which they were able to implant false memories, the experimenters will debrief the children.

Answer this question using the APA guidelines.

3. People often compare themselves to other people, which has led social psychologists to study the phenomenon. In the proxy model, one estimates one's own likelihood of success by comparing oneself to a proxy. For example, suppose that you were trying to decide whether or not to go to graduate school. If you knew someone (a proxy) who was very much like you (in perceived intelligence, motivation, grades, etc.) and that person had failed in her attempt at graduate school, that knowledge might lead you to question your own likelihood of success in graduate school. The type of tasks studied varies quite a bit, but one area studied by Martin, Suls, and Wheeler (2002) involves grip strength (though they study other kinds of tasks as well). Participants begin by squeezing an exercise hand grip as many times as they could in a 30-sec period (Task1). Their performance was recorded and related to the participants. Then participants were shown a hand dynamometer, which measures hand grip force in kg/force exerted. They were asked to predict how much pressure they could exert on the dynamometer (Task 2). To aid them in their predictions, participants were shown (fictitious) results from a participant "from last semester." Participants were randomly assigned to one of three proxy conditions: Inferior (the proxy hadn't squeezed as many times on Task 1 as the participant), Similar (the proxy had squeezed about as often on Task 1 as the participant), or Superior (the proxy had squeezed more often on Task 1 than the participant). In addition, based on written notes, the effort invoked by the proxy is either Ambiguous (participants were told that the experimenter did not know whether the proxy had exerted maximal effort or not) or Maximal (participants were told that the proxy had exerted maximal effort). The study was conducted as a 3x2 independent groups design. The dependent variable is a prediction difference score (PDS). Thus, if the participant expected to perform better than the proxy, the PDS would be positive. Solely from the figure below, predict what the authors found in their analyses of these data (treat any difference as significant) and interpret the results as best you can (as you would in a Discussion section). [15 pts]



Main Effect for Effort: Yes/No Why? **If there is a ME, it's small. [Ambig =0, Max = .13]**

Main Effect for Proxy: Yes/No Why? **Yes. [Inf = 1.75, Sim = -.2, Sup = -1.35]**

Interaction: Yes/No Why? **Yes, the effects of Effort are not the same at all levels of Proxy.**

Interpretation **When Proxy doesn't do as well on the first task, participants think that they will do better on the second task, but especially if the proxy's effort was ambiguous. When proxy was similar on the first task, participants thought they might do a bit worse than proxies who had exerted maximum effort on the first task. If the proxy's performance was superior, participants thought that they'd do a lot worse if the proxy exerted ambiguous effort, but less poorly if the proxy had exerted maximum effort. It appears that when the participants know that the proxy exerted maximum effort, the participants think that they will be able to produce their own maximum effort and perform similarly. That is, they'll do better than the proxy when they had done better on the first task and then a bit worse when they did the same or worse than the proxy on the first task. When the participant doesn't know about the proxy's effort, however, the participants presume that they can do a lot better if they did better than the proxy on the first task (probably because they plan to exert maximum effort). When the participant doesn't know about the proxy's effort and they know that the proxy did a lot better than they did on the first task, the participants think that they will not do nearly as well, probably because they figure that even if they do the best they can do, the proxy would be able to do better on the second task by exerting more effort.**

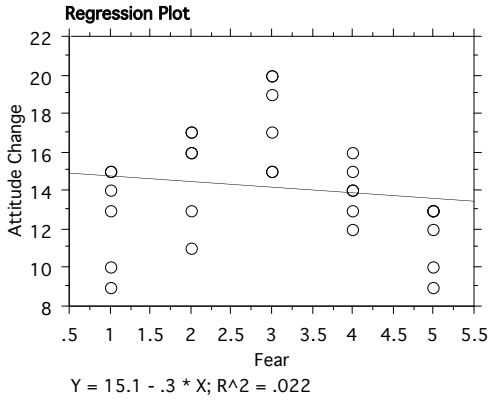
4. You are interested in doing a senior thesis with Professor X, but you have no idea what you might investigate. You talk to Professor X and, after reading a number of articles recommended by Professor X, work with Professor X to determine an area of research. You then work with Professor X to derive a research design. You collect your data and then approach Professor X for help with the data analysis and interpretation. When you have completed your thesis, Professor X informs you that your thesis may well become a part of a journal article, but that your contributions were insufficient to merit authorship. Using the principles articulated by Fine and Kurdek, either argue that Professor X has made a reasonable decision or that Professor X has not appropriately acknowledged your contributions. [10 pts]

Answer this question using the principles discussed by Fine & Kurdek.

5. Some researchers, such as McGuire (1968), have studied the relationship between the amount of fear invoked in a persuasive message and the extent of attitude change. Suppose that you observed a set of results such as those seen below. Interpret the results as completely as you can.
If a person had a Fear Score of 3, what would be your best estimate of that person's Attitude Change score? [5 pts]

Regression Summary
Attitude Change vs. Fear

Count	30
Num. Missing	0
R	.149
R Squared	.022
Adjusted R Squared	•
RMS Residual	2.924



ANOVA Table
Attitude Change vs. Fear

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	5.400	5.400	.632	.4335
Residual	28	239.400	8.550		
Total	29	244.800			

Regression Coefficients
Attitude Change vs. Fear

	Coefficient	Std. Error	Std. Coeff.	t-Value	P-Value
Intercept	15.100	1.252	15.100	12.061	<.0001
Fear	-.300	.377	-.149	-.795	.4335

First of all, you should notice that there is not a significant linear relationship between the two variables ($r(28) = .149$, $p = .4335$). That said, in looking at the scattergram, you should see that there is a nice relationship between the two variables, but that it isn't linear. In order to analyze these data, you might consider computing two separate linear analyses, with Fear scores ≤ 3 and then with Fear scores ≥ 3. With a score of 3, you would probably obtain an Attitude Change score of approximately 18 (but you couldn't use the regression equation to make that prediction).

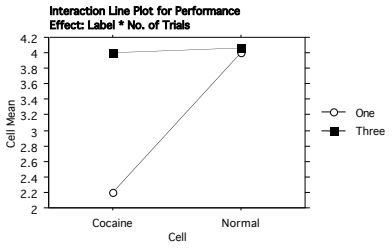
6. As you know, Thurman, Brobeil, Ducette, and Hurt (1994) conducted research on the effects of the label *prenatally exposed to cocaine*. As in our lab, participants were either told that the children they would be viewing (who were completing a delayed-response task) had been prenatally exposed to cocaine or were normal, healthy children. Moreover, the participants saw the children perform in the task *either* once or three times. (That is, this experiment is a completely independent groups design.) The dependent variable is a rating of overall performance on the task (1 = Very Poor and 5 = Very Good). Suppose that the results had turned out as seen below. Complete the source table and then interpret the results of the study as completely as you can. [10 pts]

ANOVA Table for Performance

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Label	1	13.067	13.067	20.709	<.0001	20.709	.998
No. of Trials	1	13.067	13.067	20.709	<.0001	20.709	.998
Label * No. of Trials	1	11.267	11.267	17.857	<.0001	17.857	.993
Residual	56	35.333	.631				

Means Table for Performance
Effect: Label * No. of Trials

	Count	Mean	Std. Dev.	Std. Err.
Cocaine, One	15	2.200	.862	.223
Cocaine, Three	15	4.000	.756	.195
Normal, One	15	4.000	.756	.195
Normal, Three	15	4.067	.799	.206



There is a significant main effect for Label ($F(1,56) = 20.709$, $p < .001$) and a significant main effect for Number of Trials ($F(1,56) = 20.709$, $p < .001$). There is also a significant interaction between Label and Number of Trials, $F(1,56) = 17.857$, $p < .001$. To interpret the interaction, you need to compute a post hoc test (Tukey's HSD):

$$HSD = q\sqrt{\frac{MS_{Error}}{n}} = 3.75\sqrt{\frac{.631}{15}} = .77$$

Thus, it appears that when viewed for three trials, participants rate the performance of the "Cocaine" child and the "Normal" child as equivalent. However, when rating the children after observing only one trial, the performance of the "Normal" child is rated as significantly higher than the "Cocaine" child. This interaction might be explained by the role of experience/familiarity, whereby prejudices/biases may be eroded by exposure. When viewing the children for only one trial, essentially little experience, participants may rely on heuristics and think that the child prenatally exposed to cocaine performs more poorly. However, after observing children for three trials, participants may be more willing to ignore their preconceived notions and judge the children on the basis of their actual performance.

7. Many tasks require people to be vigilant over a long period, such as an air traffic controller monitoring a radar screen to maintain the proper distance between planes. Thus, psychologists have studied various factors affecting vigilance. As an example, Pfendler and Widdel (1986) had people monitor a simulated display of the dials in the control room of a ship for a 2.5-hour period to see if people became increasingly likely to miss changes in the dials over the 2.5-hour period. Suppose that we were to extend the experiment to a 4-hour period. The dependent variable is the amount of time it takes a person to detect a change in a dial. (Higher numbers mean poorer performance.) People's performance is measured after 1, 2, 3, and 4 hours. Below is a StatView analysis of the data from this experiment. Interpret the results as completely as possible. Be very explicit about the basis for your decisions. [10 pts.]

ANOVA Table for Duration

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Subject	12	81.188	6.77				
Category for Duration	3	18.985	6.33	63.3	<.0001	183.367	1.000
Category for Duration * Subject	36	3.727	0.10				

Means Table for Duration
Effect: Category for Duration

	Count	Mean	Std. Dev.	Std. Err.
1 hr	13	6.300	1.242	.344
2 hrs	13	6.377	1.339	.371
3 hrs	13	7.077	1.448	.402
4 hrs	13	7.792	1.283	.356

There is a significant effect for Duration, $F(3,36) = 63.3$, $p < .001$. To assess the impact of particular durations, a post hoc test would be essential:

$$HSD = q\sqrt{\frac{MS_{Error}}{n}} = 3.8\sqrt{\frac{.10}{13}} = .33$$

People take significantly longer to detect the change after 4 hours compared to 3, 2, and 1 hours. People also take significantly longer to detect the change after 3 hours compared to 2 and 1 hours. There is no difference between 1 and 2 hours. Note that this study is not appropriately counterbalanced. That is, the participants are working for 4 hours, and are measured on the hour. Thus, the effect could well be a fatigue effect or carryover effects, but you can't determine what's going on because of the lack of counterbalancing. At the same time, it would be very difficult to conduct this study as an appropriately counterbalanced repeated measures design.

8. An eclectic collection of questions for the grand finale. [5 pts]

a. In a 4x5 mixed design, with 4 levels of the independent groups factor and 5 levels of the repeated factor, how many participants would you need if you wanted at least 40 pieces of data in each cell?

With complete counterbalancing, you'd need 480 participants. With incomplete counterbalancing, you'd need 160 participants.

b. What is the definition of power?

Power is the probability of correctly rejecting the H_0.

c. What is a Type I Error?

Type I Error arises when you incorrectly reject H_0.

d. What do Sir Cyril Burt, Samuel George Morton, Stephen Breuning, and Sharon Zeitlin have in common?

They all fabricated data.

e. If the effect size is large, what should you observe in a comparison of the means of your study?

The means should be fairly different (certainly in comparison to the error term).