ID# _____ Exam 2 PS 306, Fall 2004

As always, the Skidmore Honor Code is in effect and you will sign your name to attest to your adherence to the code at the end of the exam. Read each question carefully and answer it completely. If you spot a problem with a design, be sure to write about the problem in your answer to the question. Good Luck!

1a. How seriously do people take product reviews? Chaiken and Maheswaran (1992) conducted an interesting experiment in which they varied the credibility of the review source and the general message of the review. They asked college students to read a review of a new telephone answering machine. The researchers told half the participants that the review came from a flyer printed by the discount store Kmart (low credibility) or from the magazine Consumer Reports (high credibility). Each participant then read one of three types of review, an unambiguous strong review, an ambiguous review (the answering machine was better than some machines but not as good as others), or an unambiguous weak review. The researchers then asked the participants to rate on a 10-point scale their willingness to buy the answering machine for $50 (10 = very willing to buy). The results of their study are replicated below. Complete the analysis and interpret the results of this study as completely as you can (as in a Discussion section). [15 pts] [Pittenger]

ANOVA Table for Rating

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Source Credibility				.05	.8143	.056	.056
Endorsement				33.6	<.0001	67.235	1.000
Source Credibility * Endorsement				24.3	<.0001	48.678	1.000
Residual			1.2				

Means Table for Rating
Effect: Source Credibility * Endorsement

	Count	Mean	Std. Dev.	Std. Err.
High, Ambiguous	10	8.000	1.155	.365
High, Strong	10	5.900	1.101	.348
High, Weak	10	3.000	1.155	.365
Low, Ambiguous	10	5.400	.699	.221
Low, Strong	10	6.600	1.174	.371
Low, Weak	10	5.100	1.197	.379

1b. The F-Value (and Power) are quite low for Source Credibility. Would you then conclude, given the lack of a main effect for Source Credibility, that Source Credibility had no impact on the dependent variable (willingness to buy)? Why? [3 pts]

1c. Suppose that these same 60 data points had been analyzed in a single-factor ANOVA on Endorsement. What would the source table look like for that single-factor ANOVA? [7 pts]

Source	df	SS	MS	F

2. How many participants would you need for these designs with a minimum n = 30? [10 pts]

A completely independent groups (between) 3x5 design	
A completely independent groups (between) 2x9 design	
A completely repeated measures (within) 2x9 design	
A mixed 3x7 design, with the second factor repeated measures (within)	
A mixed 2x9 design, with the first factor repeated measures (within)	
A mixed 2x9 design, with the second factor repeated measures (within)	

3. False memories emerge when people study a list of words, all of which are related to a specific "lure" word. For instance, you might be presented with a list of words such as: garbage, waste, can, refuse, sewage, bag, junk, rubbish, sweep, scraps, pile, dump, landfill, debris, and litter. Although the word "trash" never appeared in the list, people will often report having seen "trash." This paradigm was developed by Deese and later revived by Roediger and McDermott, so it is often referred to as the DRM paradigm. Dewhurst and Robinson (2004) were interested in the extent to which such false memories (or intrusions) vary with age. To that end, they had 5-, 8-, and 11-year olds study lists of words (such as the one above) anticipating a memory test. Their errors were examined (as one dependent variable) to see if they varied with age. The type of error a child might make was classified as Phonological (e.g., a rhyming intrusion, such as "bunk" instead of "junk"), Semantic (e.g., reporting "trash" even though it never occurred, but was semantically related to the list), or Unrelated (e.g., reporting that they had seen "shovel"). Dewhurst and Robinson got results such as those seen below. On the basis of these results, tell me what you would expect them to find in their analysis of the data. Then tell me what sense you would make of these data, as in a discussion section. [20 pts]

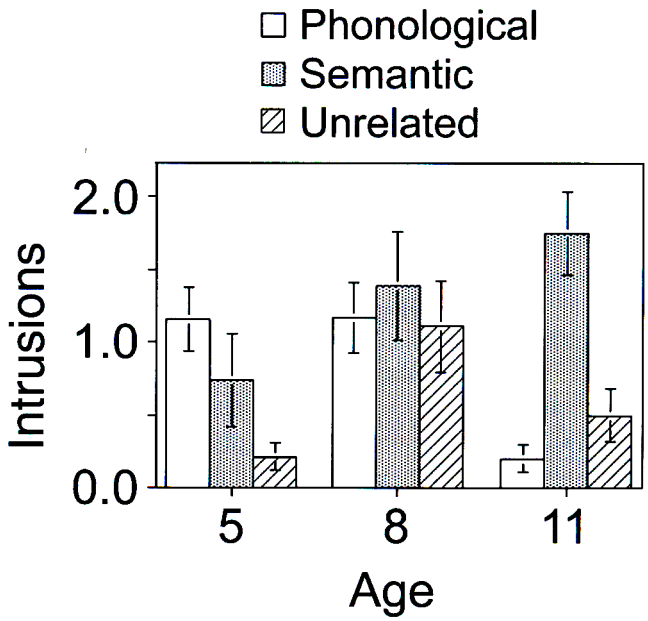

Fig. 2. Mean numbers of phonological, semantic, and unrelated intrusions per child (with standard errors) as a function of age group.

4. We have discussed both experimenter effects and participant effects.
a. Describe the Rosenthal study that showed the problems that might emerge from the experimenter learning about how the participants were performing on a memory experiment *or* the Rosenthal study that showed the impact of telling the experimenter what to expect in terms of the ratings of pictures of people's faces.
b. What antidote(s) would you suggest to deal with such expectancy effects?
c. Define demand characteristics and provide a clear example. [10 pts]

5. Dr. Harold Hedd was interested in the impact of men's hair on women's ratings of the men's attractiveness. He was also interested in the extent to which women's age might have an impact on their ratings. He took thirty pictures of men with full heads of hair and then used computer software to give the men in the pictures a receding hairline and also to make them completely bald. He then presented thirty pictures to women and asked them to rate the attractiveness of the men in the pictures on a 7-pt scale (1 = unattractive and 7 = attractive). One-third of the pictures showed the men with a full head of hair, one-third of the pictures showed the men with receding hairlines, and one-third of the pictures showed the men as bald. To ensure that the effects were not specific to particular men, the men's pictures were rotated through all three options, such that some women saw a man with a full head of hair, other women saw that man with a receding hairline, and other women say that man as bald-headed. The women in the study were young (median age of 26), middle-aged (median age of 48), or older (median age of 62). Complete the source table and interpret the results as completely as you can. [10 pts]

ANOVA Table for Rating

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Age			3.3		.0488	6.578	.581
Hair			60.7		<.0001	121.378	1.000
Age * Hair			.1		.9854	.356	.066
Residual		36.0					

Means Table for Rating
Effect: Age * Hair

	Count	Mean	Std. Dev.	Std. Err.
Middle-Age, Bald	5	5.600	1.140	.510
Middle-Age, Full Hair	5	6.000	1.000	.447
Middle-Age, Receding	5	2.200	.837	.374
Older, Bald	5	5.000	1.225	.548
Older, Full Hair	5	5.600	1.140	.510
Older, Receding	5	2.000	1.000	.447
Young, Bald	5	6.200	.837	.374
Young, Full Hair	5	6.400	.894	.400
Young, Receding	5	2.800	.837	.374