Exam 1 PS 306, Fall 2004

1. Define what we mean by the term *operational definition* and then cite a specific example of an operational definition from both Milgram's shock study *and* Harlow's infant monkey study. [5 pts]

Your book talks about two types of operational definitions (measured and experimental). For example, you might want to talk about how obedience is operationally defined in Milgram's study. Or you might want to talk about how attachment is operationally defined in Harlow's study.

2. A study was conducted to compare the effectiveness of different rewards that might be used to teach retarded children how words are spelled. Forty retarded children, chronological ages 5 to 7, were randomly assigned to four groups, with 10 children in each group. Each child was shown five common objects and five cards, each showing the printed name of one of the objects. The child's task was to match each object correctly with its name card. Whenever a correct match was made, the experimenter rewarded the child. The type of reward determined the four levels of the factor:

Candy
Tokens that could be exchanged for candy
Tokens that could be exchanged for attention (playing games, reading to the child, etc.)
Verbal praise

The experimenter recorded the number of trials required before a child could correctly match all five pairs of words and pictures. Thus, smaller numbers indicate faster learning. Analyze and interpret the output below as completely as you can. [15 pts]

ANOVA Table for Trials

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Reward	3	59.875	19.958	2.931	.0465	8.794	.644
Residual	36	245.100	6.808				

Means Table for Column 2
Effect: Reward

	Count	Mean	Std. Dev.	Std. Err.
Candy	10	12.400	2.836	.897
Token/Attent	10	11.300	2.263	.716
Token/Candy	10	14.500	3.171	1.003
Verbal praise	10	13.700	2.003	.633

H_0: $\mu_C = \mu_{TA} = \mu_{TC} = \mu_{VP}$ and H_1: Not H_0

Decision: Reject H_0, because P-Value is ≤ .05. Next, you'd conduct a post hoc test:

$$HSD = 3.81\sqrt{\frac{6.808}{10}} = 3.14$$

***Tokens for Attention* led to faster learning (*M* = 11.3) than *Tokens for Candy* (*M* = 14.5). No other differences were significant.**

3. First of all, define external validity and contrast it with internal validity (be very clear!!). What would you describe as the major point that Mook was making in his article on external invalidity? Mook argues that external validity may not be all that important when we are considering:

a. whether something *can* happen, rather than whether it typically *does* happen;

b. a prediction that something should happen in the lab, based on a theory;

c. showing that a phenomenon is so powerful that it can happen even under unnatural conditions.

Using specific studies as examples, briefly describe the evidence from the article that illustrates *two* of the above conditions under which external invalidity may not be all that problematic. [10 pts]

Answering this question requires familiarity with Mook's article.

4. What is the difference between Type I and Type II Errors? What is power? With which type of error is power most closely associated? Why? [5 pts]

Type I Errors occur when rejecting H_0 (in error). Type II Errors occur when retaining H_0 (in error). Power is the probability of correctly rejecting H_0. Power is more closely associated with Type II Errors, because they are complementary. That is, if the probability of a Type II Error is .15 then power is .85.

5. What is a non-manipulated characteristic of a participant? (Examples help.) Briefly describe a (hypothetical) study using such a non-manipulated characteristic as the "independent variable." Then, tell me explicitly why you would not be comfortable making a causal claim about any results that emerged from such a study. [5 pts]

A non-manipulated characteristic of a participant is a characteristic of that person that is intrinsic, so that it cannot be manipulated. Examples of such characteristics would include gender, IQ, race, etc. Suppose, for instance, that you were interested in studying the differences in gender (the "IV") on memory for odors (such as perfumes). Suppose that you find that females have better odor memory (correctly naming the odor). Would you attribute your finding to gender? You should be cautious in making such a claim, because it could well be that the "third variable" is experience with odors (including perfumes, etc.). Thus, people who have more experience making odor judgments is the "actual" "IV," and it turns out that women have more experience with odors.

6. Moore and McCabe (1989) looked at the relationship between expenditures on alcohol and tobacco in data from the Family Expenditure Survey of the British Department of Employment. The data are as seen below:

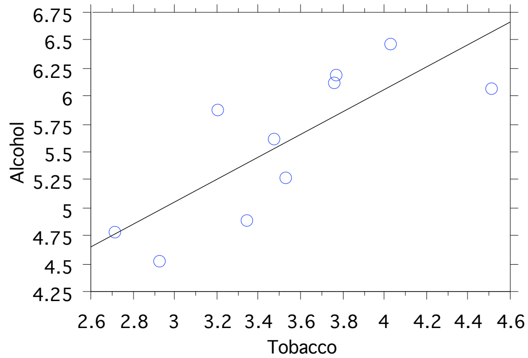
Region	Alcohol	Tobacco
North	6.47	4.03
Yorkshire	6.13	3.76
Northeast	6.19	3.77
East Midlands	4.89	3.34
West Midlands	5.63	3.47
East Anglia	4.52	2.92
Southeast	5.89	3.20
Wales	5.27	3.53
Scotland	6.08	4.51

Below you will find StatView analysis of this data set. Interpret the results as completely as you can. If you knew that expenditure on tobacco was 6.0, what would be your best prediction of expenditure on alcohol? If you knew that expenditure on tobacco was 4.02, what would be your best prediction of expenditure on alcohol? As it turns out, there was also data from Northern Ireland. For that area, the expenditure on alcohol was 4.02 and the expenditure on tobacco was 4.56. What would the inclusion of this area have on the overall analysis? (In other words, if you were to add this data point to the scattergram, what would happen to the correlation?) {Howell} [10 pts]

Regression Summary
Alcohol vs. Tobacco
Row exclusion: Untitled Dataset #1

Count	10
Num. Missing	0
R	.784
R Squared	.615
Adjusted R Squared	.567
RMS Residual	.446

Regression Plot
Row exclusion: Untitled Dataset #1



Y = 2.041 + 1.006 * X; R^2 = .615

ANOVA Table
Alcohol vs. Tobacco
Row exclusion: Untitled Dataset #1

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	2.543	2.543	12.785	.0072
Residual	8	1.591	.199		
Total	9	4.135			

Regression Coefficients
Alcohol vs. Tobacco
Row exclusion: Untitled Dataset #1

	Coefficient	Std. Error	Std. Coeff.	t-Value	P-Value
Intercept	2.041	1.001	2.041	2.038	.0759
Tobacco	1.006	.281	.784	3.576	.0072

There is a significant positive linear relationship between expenditure on alcohol and expenditure on tobacco, *r*(8) = .784, *p* = .007. If the expenditure on tobacco were 6, then the expenditure on alcohol would be 8.1. If the expenditure on tobacco were 4.02, then expenditure on alcohol would be 6.1. You should be able to determine that adding that one additional point (alcohol = 4.02 and tobacco = 4.56) would decrease the correlation (in fact, making it non-significant).

7. A habituation task is often used to test the memory of infants. In the habituation procedure, a stimulus is shown for a brief period, and the researcher records how much time the infant spends looking at the stimulus. This same process is repeated again and again. If the infant begins to lose interest in the stimulus (decreases the time looking at it), the researcher can conclude that the infant "remembers" the earlier presentations and is demonstrating habituation to (boredom with?) an old familiar stimulus. Hypothetical data from a habituation experiment were analyzed to produce the analyses seen below. Infants were exposed to stimuli of Low, Medium, and High Complexity. In the course of the experimental session, infants saw each stimulus 6 times for 1 minute each, with the order of presentation varied in a counterbalanced fashion (e.g., LMH, LHM, MLH, etc.). The dependent variable of interest here is the time spent looking at each type of stimulus on the sixth (and last) presentation of each stimulus. Complete the source table and then analyze and interpret the data as completely as you can. [15 pts]

ANOVA Table for Complexity

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Subject	9	.150	.017				
Category for Complexity	2	.255	.127	56.978	<.0001	113.956	1.000
Category for Complexity * Subject	18	.040	.002				

Means Table for Complexity
Effect: Category for Complexity

	Count	Mean	Std. Dev.	Std. Err.
Low Complexity	10	.744	.116	.037
Medium Complexity	10	.890	.073	.023
High Complexity	10	.966	.048	.015

H_0: $\mu_L = \mu_M = \mu_H$ H_1: Not H_0

Reject H_0, because P-Value ≤ .001. Next, you'd need to compute a post hoc test:

$$HSD = 3.61\sqrt{\frac{0.002}{10}} = 0.05$$

Thus, children spent significantly more time looking at the *High Complexity* items (*M* = .966), compared to the *Medium Complexity* (*M* = .890) and *Low Complexity* items (*M* = .744). Children also spent more time looking at *Medium Complexity* items than at *Low Complexity* items.

8. What is a ceiling effect? How would you know if you were dealing with a ceiling effect in a study? What would be the likely impact of a ceiling effect on $MS_{Treatment}$ and MS_{Error}? [5 pts]

A ceiling effect is present when the dependent variable is too easy, leading to uniformly high scores. Because the ceiling effect would lead to a preponderance of high scores, there would also be less variability, thus lowering MS_{Error}. At the same time, the condition means would be more similar, also lowering $MS_{Treatment}$.