TO: **Students Taking First Examinations**
FROM: **Professor Rindskopf**
DATE: **December 1, 2000**

MOST FREQUENT ERRORS ON FIRST EXAMS

While reading first exams, I have noticed that some types of errors are so pervasive that they deserve mention as a warning to future exam takers. While I had the research methods and statistics examinations in mind as I did this, and the examples are all from past exams in those areas, I hope that the ideas presented may be useful more generally.

PROBLEM I: INCOMPLETE STATEMENTS

You should always explain statements as fully as you can. In particular, don't use technical terms (or "buzzwords") without telling what they mean. While we do want you to know the technical vocabulary, we also want to make sure you know what these words and phrases mean. Your best strategy is to pretend you are trying to teach the subject or write a textbook for an intelligent reader, giving full explanations and examples.

Example

Question: What is the standard error of a statistic?

Answer: The standard error is the standard deviation of the sampling distribution of the statistic.

Comment: This is true, but incomplete. Tell what a sampling distribution is; give an example; tell why the SE is an important quantity to know (that is, what it is used for?).

PROBLEM 2: FUZZY THINKING

Fuzzy thinking manifests itself through incomplete definition of terms, inclusion of irrelevant terms, omission of relevant terms, failure to mention critical assumptions necessary for an argument, making assumptions that are questionable, and presenting arguments in which conclusions do not follow logically from the premises.

Fuzzy thinking is especially prevalent in the criticism of research designs. Answers should be relevant to the particular question, and practical for the situation described.

Example: Several people recommended a design in which students would be randomly assigned to attend either public or private school. While random assignment is good research practice, it is impractical in the situation described.

Example: Many people list all possible sources of invalidity to all research designs (e.g. history, maturation) without saying why they would be likely to occur in the study being discussed, or what their effect would be. (see also INCOMPLETE STATEMENTS)

Example: "Factor analysis is an example of validity." No; factor analysis is a method used to investigate validity in a set of measures. (see also FUZZY WORDING)

Example of a gratuitous statement: Internal validity is the sine qua non of a research design.

PROBLEM 3: FUZZY WORDING, INCOMPLETE SENTENCES

This is a close companion to and sometimes indistinguishable from, FUZZY THINKING, of which it is often symptom. In statistics and research design, technical terms have specific meanings; it is not enough to "get the general idea." Ideas must be expressed precisely.

Example: One writer wanted to demonstrate what "variable representativeness" means: "The aggression specified for the suburban child must be the same aggression specified for city."

MORE COMPLETE EXAMPLES

With the above principles in mind, consider the following examples from previous examinations. Following the statement of the question is an actual student response, and then a response that would be considered correct. Notice that answers do not necessarily have to be long.

Question 1: What is meant by the terms internal validity and external validity ?

Answer (wrong)

Internal validity is a statement about all the sources of variance which contribute to the main effects resulting from the independent variables of a research design. It also reflects the reliability of the test or treatment (independent variable); how well the independent variable accomplishes

its purpose in the design: Does it measure what it purports to measure? There are many sources of threats to the internal validity. Among these are: history, maturation, the procedure of testing, the variation in the instrument itself, and in the observers, the effect of the regression toward the mean (especially with extremes in the population), the selection of the subjects (whether there is any bias) and mortality (if people drop out, etc.), and the interaction of selection and maturity (the group we selected maybe i.e. more anemic than normal population; therefore they all tire more easily). It is related to the internal consistency of the measure.

External validity is also a measure of the degree to which the test measures what it purports to measure, but it is more related to the external criterion. In the research design the following are possible threats to the interaction of the independent variables: The interaction effects of the test and the independent variable (carry-over effects, etc.); the interaction of the test and the selection of the students; the effect of the reactive arrangements; the multiple effects of many factors. It may be related to prediction, or other kinds of relationships between the variables.

Answer (right)

The validity of a study is assessed by the degree to which the conclusions are decisive and not subject to plausible alternative explanations. A study has high internal validity if the differences among groups can be attributed to the effect of the independent variable. External validity refers to the ability to generalize the results of a study to other populations of people, other (similar) measures of the outcome, and other (similar) treatments.

Question 2: Define aptitude-treatment interaction.

Answer (wrong)

Aptitude treatment interaction is the graphic representation of the individual differences and the treatment or instructional method that you use where the individual different match with a particular method of instruction and this can be graphed. For example particular age and bilingual education method or low intelligence and particular instructional method.

Answer (right)

An aptitude-treatment interaction occurs in an educational context if certain types of people learn best using one method of instruction, while other types learn best using another method of instruction. For example, people who are less intelligent might learn better when taught using

concrete examples than using abstractions, while those who are more intelligent might learn better when taught using abstractions.

(While the above is sufficient, a more complete answer could include a discussion of the difference between ordinal and disordinal interactions, and a plot showing an aptitude-treatment interaction.]

Question 3: What is the appropriate technique for analyzing a crosstabulation of two categorical variables (e.g. , social class and religion), and what hypothesis is tested by that technique?

Answer (wrong)

The appropriate technique for analyzing a crosstabulation of two categorical variables (e.g. social class and religion) would be to use the nonparametric technique of chi square. This technique tests whether the obtained frequencies significantly differ from the expected frequencies attributable to chance.

Answer (right)

The appropriate technique for analyzing a crosstabulation of two categorical variables is chi square. The hypothesis being tested is that the two categorical variables are independent. In this context, one interpretation of this hypothesis is that the probability of being in a particular social class (for example, the highest) is the same for people of each religion. The chi square test determines whether the observed frequencies differ significantly from those that would be expected to occur if the two variables are independent.