

"Making Multiple Choice Tests More Effective"

With

Linda Suskie

Monday, April 25, 2017

Handouts

Schreyer Institute for Teaching Excellence
<http://www.schreyerinstitute.psu.edu/>

Introduction to Statistics
Test Blueprint for Final Examination

- 4 points Determine the value of t needed to find a confidence interval of a given size.
- 4 points Given a proportion and a sample size, decide if the normal distribution can be used instead of the binomial.
- 4 points Calculate a “sample error” or “error margin” for a proportion.
- 4 points Understand the effect of p on the standard error of a proportion.
- Given a research problem...
- 24 points —Choose the appropriate statistical analysis from those studied in this course.
- 16 points —Decide on the appropriate null and alternative hypotheses and statement them correctly.
- 8 points —Identify the critical value(s) bordering the critical region (this involves deciding if the test is one- or two-tailed and perhaps computing the degrees of freedom).
- 24 points —Choose the appropriate standard error formula.
- 12 points —Choose the appropriate test statistic.
- 8 points —Decide whether to retain or reject the null hypothesis.
- 16 points —Write a paragraph summarizing the study, your findings, and your recommendations.
- BONUS**
- 4 points Understand the effect of n and s on your chances of getting statistically significant results.

© Linda Suskie

Educational Research Methods
Final Exam Outline

The final exam will consist of 25 multiple-choice items, each worth 4 points. The items will cover most of the concepts listed below. If we do not get to the last unit or so, those concepts will not be on the final; there will instead be a few more items on the concepts in the other units.

Validity and Reliability (16 points)

Understand what reliability and validity are.

Correctly identify the type of reliability and validity evidence being provided by given information on an instrument.

Explain the meaning and implications of measurement error.

Recognize examples of measurement error in a given situation.

Identify the general principles for ensuring validity.

Inferential Statistics (16 points)

Articulate a null hypothesis.

Select the most appropriate inferential statistics (t, F, or χ^2) for a given research situation

Know the most common “cut-off” point that statisticians use in deciding whether two means differ *statistically significantly* from one another.

Correctly interpret the results of t, F, and χ^2 tests as presented in research articles.

Recognize the effect of standard deviation and sample size on the results of a statistical test.

Experimental Research (12 points)

Interpret correctly the symbolic representations of experimental designs.

Explain the benefits and limitations of each experimental and quasi-experimental design covered in class.

Identify the appropriate research design for a given research situation.

Correlational Research (12 points)

Summarize what a regression equation is and what it is used for.

Explain what r^2 , R, R^2 , and partial correlations are and what they tell us.

Summarize what multiple regression analysis is used for and what it tells us.

Qualitative Research (Observation, Interviews, and Ethnographic Research) (16 points)

Explain what qualitative research is and its key characteristics.

Articulate the pros and cons of qualitative research.

Explain what are focus groups are.

Identify the pros and cons of focus group research.

Identify the key principles in conducting focus groups.

Explain what ethnographic research is and identify examples of it.

Historical Research (12 points)

Articulate the need for historical research.

Identify kinds of historical research sources.

Recognize examples of primary and secondary resources.

Explain how to evaluate historical research.

Content Analysis (12 points)

Explain what content analysis research is.

Identify the pros and cons of content analysis.

Recognize examples of different kinds of content analysis.

Explain what a readability index is.

Explain how to analyze content analysis data.

Multiple Units (4 points)

Identify the most appropriate research method for a given situation.

Tips on Writing Good Multiple Choice Questions

All the suggestions that follow stem from two basic precepts:

- I. Remove all barriers that will keep a knowledgeable student from getting the item right.
- II. Remove all clues that will help a less-than-knowledgeable student get the item right.

General Tips

1. Plan your multiple choice test by first writing a list of the goals you want to assess and the number of items you want to write for each goal.
2. Keep each item as *short* and concise as possible. Avoid irrelevant material, digressions, and qualifying information unless you are specifically teaching the skill of identifying needed information. Don't repeat the same words over and over in the options; put them in the stem.
3. Define all terms carefully (e.g., largest in terms of area or population? What do you mean by sometimes, usually, or regularly?). Many terms are easily misinterpreted.
4. Don't make the vocabulary unnecessarily difficult.
5. Watch out for "interlocking" items, where a student can pick up the answer to one question from another.

Writing a Good Stem

6. The stem should ask a complete question, even if it is phrased as an incomplete statement. The student shouldn't have to read the alternatives to determine the problem. Often such questions are really four or five true-false statements strung together.
7. The best items can be answered without reviewing all the options, as one must do in a "Which of the following" question.
8. Don't ask questions on trivia.
9. Don't ask questions that can be answered from common knowledge. Someone who hasn't studied the material shouldn't be able to answer the questions correctly.
10. Don't lift statements directly from the text. They will only measure simple recognition and recall and promote superficial learning.
11. Avoid negative items. If you must have them, underline and/or boldface **NOT** or **EXCEPT** and/or put it in all caps.
12. Avoid grammatical clues to the right answer by using expressions like "a/an", "is/are", or "cause(s)". Testwise students know grammatically incorrect options are wrong.

13. Avoid “specific determiners” like “always” or “never,” especially in true/false questions.

Writing Good Options

14. Use capital letters, not lower case, to label options.

15. Avoid “None of these.” A student may correctly recognize wrong answers without knowing the right answer. Use this option only when it is important that the student know what *not* to do. If you use “none of these,” use it more than once, both as a correct answer and an incorrect answer.

16. Avoid “All of the above.” It requires the student to read every option. Some students may not be aware of this (or forget the directions); if they recognize Option A as correct, they will choose it without reading further. Others will recognize A and B as correct and recognize “All of the above” as correct even though they would not have recognized C.

17. You needn’t have the same number of options for every question. Some may have only 3, others 4 or 5.

18. Order responses logically; numerically if numbers, alphabetically if single words. If there is no logical order, insert the correct answer into the responses randomly.

19. Line up responses *vertically*. If your responses are so short that this seems to waste paper, arrange your test in two columns.

20. Make all alternatives roughly the same length. Testwise students know the longest option is often the properly-qualified, correct one.

21. Avoid interlocking options, in which the student can deduce the correct answer through the process of elimination.

22. Avoid repeated words or other verbal associations between the stem and the correct response. Testwise students will pick up this clue.

The Special Problems of “Best Answer” Questions

23. Use this kind of question *only* when experts generally agree on what is best. If more than one option exists, specify that you mean “best” in one particular expert’s opinion (e.g., “According to Freud, what is the best way to..”). If neither of these situations apply, avoid using a best-answer question, since you are asking for an opinion and more than one answer may be easily justified. Avoid using “According to the textbook” or “According to your teacher”; if you need to use these qualifiers, you’re probably asking about a trivial point.

24. Underline and/or boldface **BEST** and/or put it in all caps.

25. Define in what sense you mean “best”: easiest, most useful, etc.

26. Never use “all of these” or “none of these,” since these answers make no sense in a best-answer situation. For the same reason, never use such answers as “A and C only” in best-answer problems.

Writing Good Distracters

27. The best distracters help diagnose where each student went wrong in his or her thinking. Identify each task a student needs to do to answer a question correctly and create a distracter for the answer a student would arrive at if he completed each task correctly.
28. Use intrinsically true or at least plausible statements. Testwise students recognize ridiculous statements as wrong regardless of the stem.
29. Use common misconceptions or stereotypes.
30. Repeat key words from the stem or use other verbal associations between the stem and incorrect options. Testwise students think these are clues to the correct answer.

© *Linda Suskie*

Exam Analysis Report

Form A {LINKED}

Instructor:

Course: XXXX XXX - Quiz2.4 SP17 - Section: ALL

Exam ID: XXX

Number of Students Tested: 260

High: 98 (100.0%) **Low:** 28 (28.6%) **Mean:** 75.01 (76.5%) **Median:** 77.00 (78.6%) **Standard Deviation:** 15.45 **Cronbach's Alpha** .58

ITEM	A		B		C		D		E		F		G		H		I		J		PBS
	TTL	R	TTL	R	TTL	R	TTL	R	TTL	R	TTL	R	TTL	R	TTL	R	TTL	R	TTL	R	
1 - 97.3%	7	.60/.22	253	.77/.15																	.17
2 - 85.0%	38	.60/.17	221	.79/.14	1	.64/.00															.43
3 - 91.5%	22	.59/.17	238	.78/.15																	.34
4 - 75.8%	14	.67/.14	14	.63/.16	35	.63/.15	197	.81/.14													.45
5 - 71.2%	185	.81/.13	21	.60/.14	46	.67/.14	8	.61/.19													.49
6 - 57.3%	58	.73/.12	149	.83/.13	35	.65/.16	18	.56/.14													.47
7 - 83.1%	39	.69/.17	179	.78/.16	4	.63/.11	37	.81/.12													.22
8 - 48.8%	28	.59/.14	47	.72/.15	57	.70/.14	127	.85/.12													.51
9 - 60.8%	35	.65/.17	31	.71/.15	158	.81/.14	35	.71/.13													.39
10 - 65.8%	33	.71/.11	55	.63/.17	171	.82/.13															.47
11 - 78.5%	39	.66/.16	17	.65/.15	204	.79/.14															.36
12 - 91.5%	238	.79/.14	22	.53/.15																	.46
13 - 73.5%	191	.80/.14	55	.68/.15	7	.74/.12	7	.53/.16													.36
14 - 91.5%	7	.63/.15	12	.55/.21	238	.78/.14	3	.52/.12													.37

Understanding Your Report: Next to each item number is the % of students that answered the item correctly. The TTL column indicates the total number of students that selected the represented option. The R value is the mean score (%) and Std. Deviation of scores for a particular distractor. You would expect that the students that did well on the exam selected the correct response, thus generating a higher mean score and a higher PBS (Point Biserial Correlation). In cases where an incorrect response distracted students that did well on the exam, exhibited by a high R value, should result in a lower PBS score. PBS score ranges from -1.0 to 1.0, with a minimum desired score greater than 0.15. Refer to the color legend to identify problem items.