



The Principals' Partnership
<http://www.principalspartnership.com/>
A Program of Union Pacific Foundation

Research Brief

Teacher Made Exam Designs

Question: How should test items be effectively designed on teacher made exams?

Summary of Findings:

What are the purposes of exams? Would teachers give exams if they were not grading students? According to Wilen, Bosse, Hutchison & Kindsvatter, using one test for many purposes is "...rather like trying to use one tool-say a screwdriver or a hammer-for jobs ranging from brain surgery to pile diving" (p. 339). Exams also provide information that should inform the instructional program, let the students know their strengths as well as areas for growth, and tell teachers what information their students know and what they still need to know. "Assessment is the strongest medium educators have for telling students, parents, and community what they care about in education" (Saphier & Gower, p. 511).

When determining how students will be assessed, it was advised in the literature that the following questions be asked: What is being assessed?, Why is the information being assessed?, Why are the students being assessed?, and How will the results be used? Once these queries have been answered, the type of test/exam and its questions should be designed, reviewed and revised. The most common tests are objective, primarily because they are easier to grade than essays.

The following table taken from Kellough and Kellough (p. 358), provides approximate testing times for the average student on the most common items:

Type of test item	Time needed per item
Matching	30 seconds per matching item
Completion	30 seconds per item
Multiple-choice	1 minute per item
Completion drawing	2-3 minutes
Arrangement	2-3 minutes
Short explanation	2-3 minutes
Essay and performance	10 or more minutes

Exam design is more challenging than it might initially appear. Some considerations:

- include several different types of items
- insure that all objectives have been included and are reliable
- use language that all students can understand
- keep track of the levels of questions so that higher as well as medium and lower levels are included
- type it and make certain it is error free
- provide ample space to write answers

Each type of question has its own characteristics. The most common are described below.

- **Matching:** This tends to be used to judge relationships, select differences between fact and relationships, identify definitions and choose concepts.

The literature universally recommended that there should be more options than questions, as this helps prevent students from guessing by default. i.e.

- | | | | |
|----|------------------|-------|-------------------------|
| A. | $3 \times 4 =$ | _____ | 1. 28 |
| | | | 2. 21 |
| B. | $7/6 =$ | _____ | 3. 120 |
| | | | 4. 168 |
| C. | $12 \times 14 =$ | _____ | 5. 22 |
| | | | 6. 12 |
| D. | $7 \times 4 =$ | _____ | 7. 1.28 |
| | | | 8. 114 (Marzano, p. 88) |



The Principals' Partnership
<http://www.principalspartnership.com/>
A Program of Union Pacific Foundation

Research Brief

- **Completion or fill in the blank:** Rote memory is often the skill in this type of question as there is usually one right word or phrase that fits in the space. These are best used for review or preassessment. If using on an exam, provide a word bank. i.e.
To test their hypotheses, scientists and social scientists conduct _____. (Kellough & Kellough, p. 361).
- **Multiple choice:** There is often more than one choice. These can require higher levels of thinking, however, it often measures lower levels of cognition. If there's more than one response, have students identify all correct responses. Make the distractors plausible and related to the concept. Three or fewer choices is not recommended as it makes it much easier to guess. Avoid words such as "all," "none," and "always" as they tend to be trick questions and usually do not test the course's objectives. Use charts and graphs as they can be helpful, especially for students who are limited in their English skills and/or who are at lower academic levels. Provide space so that students can write rationale for their answers. i.e.
Of four cylinders with the following dimensions, the one that would cause the highest-pitched sound would be:
(a) 4 inches long and 3 inches in diameter
(b) 8 inches long and 3 inches in diameter
(c) 4 inches long and 1 inch in diameter
(d) 8 inches long and 1 inch in diameter (Kellough & Kellough, p. 364)
Explain your response.
- **Completion drawing:** The students complete an incomplete drawing. It takes less time for the student to do this and is easy to grade. It measures conceptual knowledge and is more complete if students write a short explanation of the drawing.
- **Arrangement:** Students put items in order. It can be good for discussion, but be careful in grading. It is recommended that students include a rationale for their response. i.e.
Arrange the following list of events on a timeline in order of their occurrence: The assortment of balls on the table represents the planets in our solar system. Arrange the balls in their proper order around the sun. Explain your response (Kellough & Kellough, p. 361).
- **True/False:** The items are either correct or incorrect. Write each one as true, then make the appropriate ones false by changing a word or phrase. Use one idea only and reword concepts from the book, do not take a phrase directly from the text. Have the students either write out true or false then there will not be confusion about the what was written or have them circle the correct answer. Have the students write out the rationale for their responses. i.e.
When a farmer saw eight crows sitting on the fence surrounding his cornfield, he shot three of them. Five were left on the fence. True or False? _____ Explain (Kellough & Kellough, p. 369)
- **Short answer:** These answers tend to be several sentences to a paragraph long or short essays. A greater amount of information can be covered in these responses. In order for students to experience success on these answers, provide practice questions in an appropriate timeframe. i.e.
Briefly describe the major impact of each of the following events:
1. Battle of Gettysburg
2. Battle of Atlanta (Marzano, p. 95)
- **Essay:** Students respond in writing to a problem or question. Prepare a list of key phrases/thoughts that should/could be in the responses. Make certain to allow ample time for students to respond. Practice time is important so that students will be successful in this venue. i.e.
In the story just read, does the author elaborate the setting in great detail or barely sketch it? Explain your response. (Kellough & Kellough p. 362).



The Principals' Partnership
<http://www.principalspartnership.com/>
A Program of Union Pacific Foundation

Research Brief

Magazine:

Ceocaris, C. and Ross, M. (September, 1999). A test worth taking. *Educational Leadership* 57 (1), 29-33.

A test structure that is aligned with expecting students to meaningfully demonstrate their knowledge in a subject is described in this article.

Books

Kellough, R. d. and Kellough N. G. (2003). *Secondary School Teaching*. Pearson Education: Upper Saddle River, NJ.
Chapter 11, p. 341-380 include thorough descriptions of a myriad of different test items along with examples of each.

Marzano, R. J. (2000). *Transforming classroom grading*. ASCD: Alexandria, VA.

Chapter 6, p. 86-105, describes different components of tests and assessments and includes examples of different types of test questions and structures.

Ronis, D. (2000). *Brain compatible assessments*. Skylight Professional Development: Arlington Heights, IL.
This book examines different forms of assessment and their compatibility with the brain and student learning.

Saphier, J. and Gower, R. (1997). *The skillful teacher*. Research for better teaching: Acton, MA.

Chapter 16, p. 459-514, deals mostly with authentic assessment and provides a variety of rubric designs.

Wilens, W., Bosse, M. I., Hutchison, J. and Kindsvatter, R. (2004). *Dynamics of effective secondary teaching*. Allyn and Bacon: Boston, MA.

Chapter 10 Evaluation of Classroom Performance, p. 332-366, provide a broad spectrum of different types of assessments along with specific examples of how to write different types of test items.

Submitted

Date: 10/23/06

By: Dr. Karen Walker Lebanon Valley College

<http://www.principalspartnership.com/>

This is provided as a service to educators by The Principals Partnership and Union Pacific Foundation, neither of which assumes any responsibility for the content of the brief or the positions taken by the authors or the Web sites or other authors whose works are included. This research brief reflects information currently available and is not the official position of The Principals Partnership or Union Pacific Foundation.

Disclaimer: All URLs listed in this site have been tested for accuracy, and contents of Web sites examined for quality, at the time of addition. Content accuracy and appropriateness, however, cannot be guaranteed over time as Web sites and their contents change constantly. The author takes no responsibility for difficulties which may result from the use of any Web site listed herein. Please notify the [Webmaster](#) if you find any dead links or inappropriate material.

Permission: You may use or download content for research or educational purposes, or for your personal, noncommercial purposes, provided you keep unchanged all copyright and other notices with them. No other use of any



The Principals' Partnership
<http://www.principalspartnership.com/>
A Program of Union Pacific Foundation

Research Brief

content is permitted. You agree that you will make only lawful use of this research brief, and will only use these briefs in compliance with all federal, state and local laws and regulations. You agree that you will make no use of the research that violates anyone else's rights, including copyright, trademark, trade secret, right of privacy, right of publicity or other rights.