

Teacher-Made Tests in the Social Studies

can play a positive role.

IN order to have tests play a positive role in the educational process they clearly should reflect the same objectives that guide instruction. When instructional activities are planned with the intention of developing an understanding of certain concepts and relationships, tests should be devised which measure these understandings.

All too often, however, tests are afterthoughts for teachers, hastily assembled to provide a tangible basis for grading. Many such teacher tests place a disproportionate emphasis on the recall of miscellaneous information, much of this information unrelated to the major concepts and relationships emphasized in instruction. Since students tend to regard tests as indicators of what the teacher wants them to learn, it is not surprising that they spend their time on the indiscriminate memorization of such information.

Thus, the first order of business for a teacher preparing a test is the formulation of the major concepts and relationships he has attempted to communicate. Even if he has difficulty in devising ques-

tions which call for a thoughtful understanding of these concepts and relationships, such formulation provides needed guidance in the preparation of simple recall questions. For example, if an understanding of the concept of federalism is an objective, this suggests that certain facts about the 1781-1789 period in American history, and not others, are important. Similarly, if understanding the relationship between technology and resource utilization is an objective, this suggests the importance of historical and geographic examples of the effect of inventions on the utilization of resources. In this manner clear objectives help teachers to devise test questions which require the recall of information related to the objectives.

The formulation of objectives, such as the understanding of major concepts and relationships, also contributes to the improvement of tests by focusing attention on questions which require thoughtful understanding rather than a simple recall of information. Clearly, the ability to recall factual data, definitions and even relationships in the same form that they were presented by the teacher or textbook is not an indication that understanding is present. Questions which test for understanding must present a situa-

Dana G. Kurfman is Associate, Test Development Division, Educational Testing Service, Princeton, New Jersey.

tion somewhat unique for the student so that he is called upon to stop and reflect before responding.

One approach to measuring understanding involves the presentation of information in the form of a graph, cartoon, chart or map which has not been used in instruction. It is certainly true that cartoons and graphic materials present problems of reproduction for most teachers. Yet when their role in testing understanding is appreciated, the extra expense involved in their reproduction may become more tolerable. The following cartoon question might be used in a comparative government unit test. It requires a form of understanding, rather than merely recall, in the sense that a student must apply knowledge learned in a purely verbal form to a different mode of communication.

1. In the twentieth century the situation illustrated by the cartoon¹ has pertained most often to which of the following countries?

- ° (a) France
- (b) Spain
- (c) Canada
- (d) India.

A good approach to measuring understanding of a concept is the use of questions which require the classification or identification of an example. A map can be used effectively to determine whether students can identify examples of peninsulas, islands, and many other elementary geographic concepts. In the case of more complex concepts it is sometimes helpful to use a quotation as follows:

2. "The aim of statesmen pursuing this policy was, generally, to preserve their own independence of action to the utmost. Hence, the basic rule was to ally against any state threatening domination. If one state seemed to dictate too much, others

¹ Cartoon by Jim Ivey, originally appeared in *The St. Petersburg Times*, Florida.



"... Had so many children she didn't know what to do..."

would shun alliances with it, unless they were willing to become its puppets."

This policy is generally referred to as

- ° (a) balance of power
- (b) imperialism
- (c) isolationism
- (d) coexistence.

To determine whether students understand cause-effect or means-ends relationships, it is often effective to state certain conditions and ask which of several effects is likely to ensue. The understanding of many geographic and economic principles can be determined in this way. It is also possible to ask which of several means will accomplish a given end. The next question uses the latter approach.

3. "The United States has a trade surplus of 3 billion dollars a year, but that is not enough to cover military expenditures and other payments abroad amounting to 7.5 billion dollars a year."

The imbalance of 4.5 billion dollars could be reduced if there were an increase in

- (a) expenditures of American tourists abroad
- * (b) exports of consumer goods
- (c) imports of raw materials
- (d) income from American investments by foreigners.

Some of the best examples of understanding-type questions are those which require a comparison of two or more factors. These factors may be different persons, ideas, institutions or historical periods. The major characteristic of such questions is the requirement that students see a relationship between the ideas or historical periods compared. Some questions of this type may also require a rather careful analysis of the two or more factors to be compared. Although comparison questions can just as easily ask about differences, the following example calls for an understanding of a major similarity in three historical programs.

4. Which of the following was common to the New Nationalism, the New Freedom, and the New Deal?

- (a) Reduction of the public debt
- (b) Elimination of large-scale business corporations
- * (c) The regulation of the economy
- (d) Elimination of the influence of lobbies on legislation.

Another aspect of understanding can be probed by questions which require judgments of *relative* significance or similarity. In such questions students must choose the least or most important factor from among several likely possibilities. Questions five and six call for this type of judgment or interpretation.

5. Of the following, the most persistent

source of conflict between the western and eastern parts of the United States during the course of the nineteenth century was

- * (a) the issue of currency inflation
- (b) the regulation of monopolies
- (c) internal improvements
- (d) westward expansion.

6. With respect to which of the following are elections of members to the British House of Commons and to the United States House of Representatives most similar?

- (a) The length of the election campaign
- (b) The cost and character of the election campaign
- (c) The number of other officials elected on the same ballot
- * (d) The proportion of the total number of seats to be filled in any one election.

It is worth noting that inserting *most* or *least* in the stem of a question does not in itself make the question a problem of judgment. To ask which was the most important consequence of the Civil War, and then include only one actual consequence among three or four options unrelated to the Civil War does not require a judgment of relative significance.

In addition to measuring understanding by means of single questions, it is possible to formulate a group of questions based on a passage drawn from some source or even prepared by the teacher. Such sets of questions have the advantage of permitting students to concentrate on one topic or problem rather thoroughly before taking up another topic or problem. This approach tends to mitigate the danger of objective testing becoming an exercise in intellectual hopscotch. Questions 8 through 11, for example, require a critical examination of the following concept of separation of powers in American government.

"According to one view of government, there are three types of power involved in

making and applying policies. These have been called legislative, executive, and judicial powers. The United States government, according to this view, is based upon the principle of the separation of these powers. Congress, having the legislative power, determines policies which are formulated as laws. The President, representing the executive power, carries out the laws. And the Supreme Court, having judicial power, interprets the laws in relation to individual cases."

8. Which of the following could be used to argue against the idea that the type of government described in the passage actually describes the government of the United States?

- I. The presidential veto
- II. Judicial review of legislation (by the U. S. Supreme Court)
- III. The necessity for both houses of Congress to agree on legislation.

- (a) I only
- (b) II only
- °(c) I and II only
- (d) I and III only.

9. Of the following, the greatest threat to separation of powers, as defined in the passage has been the

- (a) Senate filibuster
- (b) "necessary and proper" clause of the Constitution dealing with congressional power
- (c) differences between bills authorizing certain actions and bills appropriating the needed funds
- °(d) exercise of the presidential prerogative.

"Many students of American government have concluded that certain changes in congressional procedure would ensure more effective cooperation among the parts of the government. One proposal would eliminate the custom of recognizing congressional seniority."

10. If this action were taken, it would have the greatest immediate effect upon the

- (a) extent of federal patronage available

- (b) time taken up by debate in Congress
- °(c) distribution of committee assignments
- (d) extent of legislative logrolling.

11. Another immediate practical consequence of such a change would be to diminish the influence of most Congressmen from

- °(a) Georgia
- (b) Illinois
- (c) Colorado
- (d) California.

Thus far, it has been suggested that teacher-made tests can be improved by basing them on more carefully derived specific objectives, and by increasing the proportion of questions requiring an understanding rather than a recall response. A third consideration in improving classroom tests is the careful review of questions to assure their clarity and correctness. This is particularly essential when writing questions dealing with probable relationships, comparisons and judgments of relative importance. Such questions are more likely to contain ambiguities and are less likely to withstand scholarly criticism than questions requiring only the recall of specific information. A teacher, after having developed the "perfect question," may be dismayed when fellow teachers can find no correct answer or even two or more correct answers.

In fact, the cooperative review of test questions by teachers who teach the same course is probably the most effective way of achieving clear and correct questions. Several teachers are more likely than one teacher to see ambiguities and errors. Moreover, since test questions are supposed to exemplify course and unit objectives, this sort of cooperation is an excellent means of clarifying teacher objectives. It is especially effective in

(Continued on page 80)

the tests used. The bulk of educational measurement and evaluation is done in the classroom by the teacher, and must be harmonized with findings of "The Testing Program." An important part of this effort should be the continuous provision of specialized assistance to teachers in developing and improving their own instruments.

Planning an evaluation program and creating measurement devices are as much an exercise in educational philosophy as they are a task for the statistician. The validation of instruments and procedures requires coordinated efforts of everyone, since the task involves educational goals, classroom practices, conditions of learning, characteristics of children, as well as criteria of test construction. For too long a period, we have acted as though all this could be done without the participation of those who know curriculum and instruction best. It is clearly evident that this is not the case, and that the improvement of testing and evaluation in schools is interlocked with the improvement of instruction. Perhaps it might be said that the basic job is one of finding ways to quantify and communicate the observations, judgments and experiences of teachers.

—CLIFFORD BEBELL, *Director, Elementary and Secondary Education, State Department of Education, Denver, Colorado.*

Teacher-Made Tests

(Continued from page 19)

making the stated objectives of a course or unit more meaningful for all the teachers involved. Another dividend ensuing from joint work on test questions comes from seeing how others pose questions requiring the understanding of a concept or relationship. The number of unique questions, so necessary for measuring

understanding, can thereby be expanded and made part of a continuous test file available to all teachers.²

Although competent teachers are the best judges of the correctness of a given response to a question and the incorrectness of the remaining responses, any teacher who discusses questions with his students after a test has been given soon becomes aware of ambiguities undreamed of earlier. One way of systematically checking test questions for clarity and accuracy is by means of an item analysis. This involves a comparison of approximately the top quarter of the papers on a test with the bottom quarter of the papers. If fewer or even the same number of students in the upper group compared to the lower group respond incorrectly to a given question, that question may well need revision. By noting which incorrect response is attracting the better students, it is possible to identify the difficulty.³

A teacher who takes the time to clarify the major concepts and relationships he desires his students to understand, follows this up with questions requiring understanding, and then revises questions to improve their clarity and correctness, has made testing a positive rather than a negative factor in the educational process. Then, even those students who study only to do well on a test will also be improving their understanding of the objectives the teacher considers important.

² A file of social studies test questions can be expanded immeasurably by obtaining several bulletins, each containing several hundred questions, published by the National Council for the Social Studies, 1201 16th Street, N.W., Washington 6, D.C. Bulletins now prepared deal with Study Skills and Critical Thinking (#15), World History (#9), American Government (#13) and American History (#6).

³ A clear, brief description of how to perform an item analysis is provided in "Short-Cut Statistics for Teacher-Made Tests," Evaluation and Advisory Service Series No. 5, Educational Testing Service, Princeton, New Jersey.

Copyright © 1962 by the Association for Supervision and Curriculum Development. All rights reserved.