

# A Multidimensional Multiple-Choice Testing System

Ray W. Karras

MOST TEACHERS OF HISTORY will probably agree with The College Board about the purposes of multiple-choice questions. As stated in its booklet, *Advanced Placement Course Description: History* (May 1984, p. 6): "The questions in the multiple-choice section are designed to test the students' factual knowledge, breadth of preparation, and *knowledge-based analytical skills*." I have italicized these last words because here is where means fail to serve ends. Few multiple-choice questions asked by The College Board — or by others — actually require students to think critically about their factual knowledge of history.

Neither The College Board nor the Educational Testing Service should be blamed for faulty multiple-choice testing. As a former member of ETS's Advanced Placement Test Development Committee in American History, I can testify that AP multiple-choice questions are designed to reflect, not direct, accepted practices in the country's classrooms. The College Board's *Course Description* is cited here primarily because it reflects standard testing practices. Nevertheless, while AP multiple-choice test makers have good reason to be cautious about radical changes, they also have a commitment to help raise academic standards. The model multidimensional multiple-choice questions offered below represent an approach toward meeting this commitment, as well as the hope that teachers will share in the process of reform.

The following suggestions require systematic classroom instruction. Knowledge-based analytical (i.e., critical thinking) skills call for the explicit teaching of specific processes that can and should be tested. In my own college preparatory courses all work is structured around six specific thinking processes applied to reading and classroom work and is tested in both essay and multiple-choice forms. Rote memorization and straight narrative reporting are discouraged. This approach, it should be stressed, has not prevented students from doing well on traditional multiple-choice tests, nor does it take longer to teach than the conventional variety. I only suggest that narration and rote memorization are inappropriate processes in history courses designed to promote critical thinking through learning knowledge-based analytical skills. The College Board says as much in its advice on AP history courses:

Although there is little to be gained by rote memorization of names and dates on an encyclopedic basis, a student must be able to draw upon a reservoir of systematic factual knowledge in order to exercise analytic skills intelligently. Striking a balance between obtaining a command of systematic factual knowledge and analyzing that knowledge critically is a demanding but crucial task in the design of a successful AP course in history. [*Course Description*, p. 3]

Striking this balance is an equally crucial task for the test maker. What is needed are multiple-choice questions that will at once impose the discipline of objective testing and require critical thinking.

Traditional multiple-choice questions essentially ask for straight recall from text reading or class work. These questions operate on a single level of generalization, asking students to equate one piece of information with another. For example, a *Course Description* sample question asks:

*Brown v. Board of Education of Topeka* was a Supreme Court decision that

- (a) was a forerunner of the Kansas-Nebraska Act.
- (b) established free public colleges in the United States.
- (c) outlawed racially segregated public schools.
- (d) established free public elementary and secondary schools in the United States.
- (e) provided for federal support of parochial schools.

No critical analysis is needed to recall the fact that *Brown v. Board of Education* outlawed racially segregated public schools. The same criticism applies to a more sophisticated type of question:

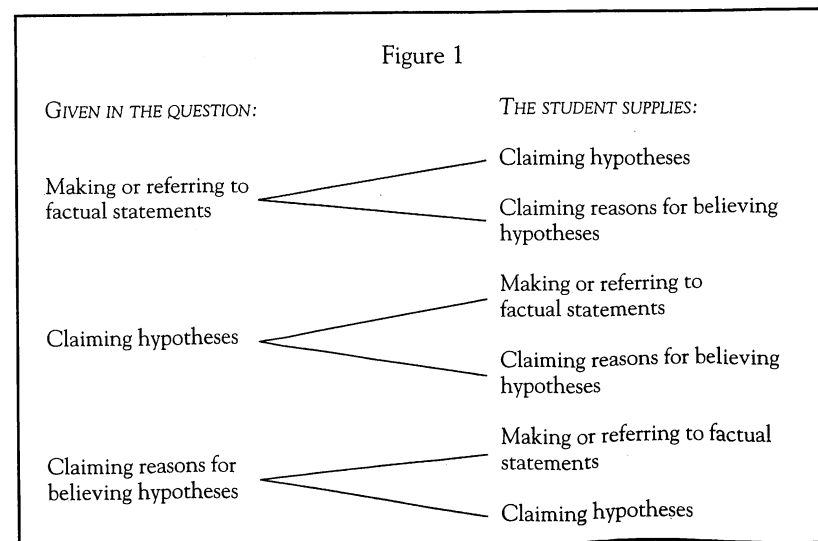
Which of the following was the LEAST important consideration in the United States' decision to drop the atomic bomb on Japan in August, 1945:

- (a) Dropping the bomb would give a new and powerful argument to the peace faction in the Japanese government.
- (b) Dropping the bomb would presumably shorten the war and therefore save the lives of American soldiers that would be lost in the invasion of the Japanese homeland.
- (c) Scientists could propose no acceptable technical demonstration of the bomb likely to convince Japan that further fighting was futile.
- (d) Scientists wished to demonstrate to Congress that the \$2 billion spent, after long debate, on the six-year Manhattan Project had not been wasted.
- (e) The President and the State Department hoped to end the war in the Far East without Russian assistance.

Such a question asks for the recall of a hypothesis (reflected in choice [d]) claiming that the fear of wasting \$2 billion was the least important reason for dropping the bomb. The factual basis for rejecting claim (d) is unsteady: What is "long debate"? Who debated whom? "Recalling" hypotheses seems to be diametrically opposed to critical thinking. This question teaches students that hypotheses should be memorized as facts, that there is one "right" hypothesis explaining the dropping of the bomb, and that whatever teachers may say in classrooms about critical thinking had best be put aside when answering multiple-choice questions. If students are to think critically, they should be expected to test hypotheses against reasoning and evidence; they should not be required — indeed, they should be forbidden — to accept as true what is only claimed. Answer (d) may be the best answer to the question, but this is not the way to ask for it.

How then can multiple-choice questions test knowledge-based critical thinking skills? The answer, I believe, first requires a theory of critical thinking that can be applied to multiple-choice questions. Second, we need to standardize and explicitly teach students certain critical thinking processes. Third, we need to construct appropriate multiple-choice questions.

One aspect of critical thinking is that it is a process of relating elements of learning to each other at different levels of generalization. Critical thinking does not translate into the equation that piece of information "A" equals piece of information "B." Instead, we need questions whose parts move between and among at least three levels of generalization: 1) making factual statements, 2) claiming hypotheses, and 3) claiming reasons for believing hypotheses. Given any one of these elements, students may be asked to supply one or both of the others. This is illustrated in Figure 1.



This theory can be embodied in six specific processes (or steps and operations) which students can be taught to apply to all their reading, writing, and classroom work. These processes ask the student to: 1) raise hypothesis-producing questions, 2) claim controversial hypotheses answering these questions, 3) claim logical reasons for believing the hypotheses, 4) apply factual knowledge as evidence to support the claims, 5) refute opposing claims and evidence, and 6) ask analytical questions for new unknown facts to test the claims the student has made. The student is the active user and developer of these processes. The teacher explains them, provides the historical sources to which they apply, demonstrates and supervises their use in classroom instruction, and evaluates the competence with which they are used.

Defining the terms in the processes is a major teaching task. For example, "claim" — as a verb — is used as "maintaining a position in the face of possible argument" (*American Heritage Dictionary*). As a noun, a claim is a statement of belief about the position taken. Thus, students claim hypotheses and reasons for believing them. Statements of claims are to be sharply distinguished from statements of fact: it is a fact that I arrived late at school; but it is a claim that I could not help arriving late at school. Similar strictures apply to other terms in the processes. Students are guided to these and only to these instructed usages throughout the course.

Arguing historical hypotheses is the centerpiece of this system. By "hypothesis" I mean a value-free controversial claim based on reasons and evidence defended against opposing claims and left open to testing by new and unknown factual knowledge. In this sense, a historical hypothesis is most definitely not a "conclusion." In common usage, "conclusion" may mean "end," yet no historical hypothesis delivers a final end to inquiry. Nor should we accept the meaning of "conclusion" that students bring into our history courses from their mathematics or logic classes. These classes teach, for example, that by deduction, if  $A = B$  and  $B = C$ , then  $A = C$  is a logical conclusion. This is clearly not what historians mean by conclusions. Therefore, it is best to cede the term "conclusion" to other disciplines, and to hold the term "historical hypothesis" for history teaching. We can teach that a hypothesis may claim why historical events happened (e.g., the American Revolution was mainly caused by economic conflict), or what mainly characterizes them (e.g., the Open Door Policy was successful in achieving its aims), or a debatable fact (e.g., Lee Harvey Oswald acted alone in assassinating John F. Kennedy). To be worth studying, historical hypotheses should be controversial.

The following sample questions are taken from final examinations given to college preparatory students in United States history. Each question addresses one of the six processes.

1. **Raising hypothesis-producing questions.** Here the tester wants students to recognize a question whose answer will claim a hypothesis, and at the same time will be knowledge-based.

Which question will produce a response that is a hypothesis?

- (a) Was freedom of speech constitutional between 1776 and 1788?
- (b) Was the Sherman Anti-Trust Act constitutional?
- (c) Was the *Dred Scott* decision favorable to the Missouri Compromise?
- (d) Did the framers of the Constitution intend the Supreme Court to have its present powers of judicial review?

In this question choice (a) requires the factual knowledge that there was no constitutional free speech guarantee between 1776 and 1788. It also requires understanding that what is and what is not constitutional is by definition a fact that courts decide. Choice (a)'s answer will be an objective fact, not a claim of a hypothesis. Answer (b) also requires understanding of the concept "constitutional," as well as the knowledge that the Sherman Anti-Trust Act is to this day in effect. Choice (c) requires a simple factual answer, not a hypothesis: *Dred Scott* did declare the Missouri Compromise unconstitutional. Answer (d) is correct. What the framers intended is controversial, as is the nature of judicial review — to which students should remember the Constitution never specifically refers.

2. **Claiming controversial hypotheses.** In the following question the student must choose a claim that not only meets the requirements of a controversial hypothesis, but also reflects knowledge of late nineteenth-century history. The correct answer may or may not be a claim developed in class preparation; the student is asked, as in all questions, to analyze and synthesize as the question is answered.

Which of the following is the best historical hypothesis?

- (a) Lack of government support mainly caused the relative weakness of the labor movement between 1880 and 1900.
- (b) The Anti-Imperialist League failed to achieve its goals because it opposed Cleveland's policy on Hawaiian annexation.
- (c) The Haymarket Square Riot of 1886 resulted in the conviction of eight agitators.
- (d) The Populist movement shows that farmers cannot compete if they sell in an open market.

The correct answer, (a), presents a hypothesis that competes with alternative claims — for example, that ineffective union leadership, or perhaps oppressive employers, were mainly responsible for union weakness. Choice (b), though stated as a hypothesis, is based on factual errors. Cleveland opposed Hawaiian annexation, and the Anti-Imperialist League, of which Cleveland himself was a member, was formed after his administration. Answer (c) merely states a historical fact. Choice (d), stated in the present eternal tense, is not a historical hypothesis, but an economic theory.

3. **Claiming logical reasons for believing hypotheses.** Here the student chooses a set of claims which state reasons that are logically necessary for believing the hypothesis given. At the same time, the correct choice must be supportable by facts. The following example presents a hypothesis and four sets of possible reasons for believing it.

*Hypothesis:* The United States has consistently maintained the principles of the Monroe Doctrine. Which of the following is the best set of reasons for believing this hypothesis.

- (a) I. Non-interference in Latin American affairs has consistently been upheld in the Roosevelt Corollary.
- II. Non-interference in European affairs has been consistently upheld in the Open Door Policy.
- (b) I. United States reactions to the sinking of the Maine were due mainly to the fear of new European aggression in the Western Hemisphere.
- II. United States policy toward belligerents in Europe in 1916 was due mainly to Wilson's support of the Monroe Doctrine.
- (c) I. The United States rejected the League of Nations.
- II. Lodge's opposition to the League of Nations was due to his desire for a Republican Party victory in the 1920 election.
- (d) I. Non-interference in Latin-American affairs was upheld in Theodore Roosevelt's laissez-faire policy.
- II. Non-interference in Latin-American affairs was upheld in the Roosevelt Corollary.

Choice (a) is incorrect because both its I and II are based on erroneous information. Answer (d) is wrong because the laissez-faire concept is inappropriate in I, and the facts regarding the Roosevelt Corollary eliminate II. Answer (b) is correct. Accurate knowledge of Monroe Doctrine provisions is applied to two episodes. Furthermore, both claims I and II in (b) are controversial at a lower level of generalization than the hypothesis in the question stem; if reasons I and II can be supported, then the hypothesis can be believed. It is not the respectability, but the arguability of hypotheses that students are asked to address.

Erroneous answer (c) illustrates several things about the process. Choice (c)I illustrates the crucial stipulation that a statement of fact is not a claim of belief. It is a fact, not a claim, that the U.S. rejected the League. The statement omits telling why this rejection should be regarded as a reason for believing the hypothesis. Had (c)I said, for example, "United States' rejection of the League of Nations was due mainly to support of the Monroe Doctrine by Wilson's opponents," it would have served as a reason for believing the hypothesis. As it stands, (c)I is only a piece of evidence necessary to support (c)II; taken together, (c)I and (c)II do not give two independent reasons for believing the hypothesis. Even separately, (c)II does not logically support it.

4. **Applying factual knowledge to support claims.** Thus far our questions have explicitly asked students to find relationships between two levels of generalization: between hypotheses and subordinate claims of reasons for believing hypotheses. Implicit has been the third level of generalization, that of knowledge of factual statements. Now we want to make the implicit explicit by testing directly for the relationship between facts and claims. First, however, teachers should clearly teach the characteristics of factual information desired from their students.

When we ask students for facts, we actually ask them for *statements* of facts. We read and write statements of fact; we do not "give" facts — we refer to them. For purposes of instruction it may be said that factual statements tell who, where, when, or what: that they refer to that which can be seen, heard, touched, or counted; and that they include accounts of historical events, statistics, and primary source quotations. For example, it is a statement of fact that Franklin Roosevelt said in 1940: "Your boys are not going to be sent into any foreign wars." But it is a historian's claim, not a fact, that Roosevelt did not intend to send American soldiers into foreign wars in 1940. The nature of statements of fact should be taught explicitly through daily classroom practice.

The question below relates facts to a claim. The factual statements needed are not given in the question; they are instead only referred to.

*Claim:* Pre-Civil War presidents were more supportive of laissez-faire policies than were post-Civil War presidents. Which pair of statements offers facts that support this claim?

- (a) I. What Thomas Jefferson said about Federalists in his First Inaugural Address
- II. What Woodrow Wilson said about intervention in Europe in 1917
- (b) I. What Grover Cleveland did about Coxe's Army
- II. The reason why Franklin Roosevelt ordered a bank moratorium
- (c) I. What Andrew Jackson did about the Bank of the United States
- II. What Woodrow Wilson did about the Federal Reserve System
- (d) I. What Alexander Hamilton did about the assumption of state debts
- II. What Herbert Hoover said about voluntary self-help in 1930

Choice (a) is incorrect; neither Jefferson nor Wilson discussed laissez-faire in the contexts of the statements. Furthermore, some students mistakenly believe Hamilton was a president; that laissez-faire means any kind of government intervention; and that half a comparison, even with Herbert Hoover in (d), is better than none. Notice too that (b) offers a nonfactual statement of FDR's reasons for the moratorium. Answer (c) is correct because it requires the factual statement that Jackson vetoed one national bank and Wilson signed another into law.

5. **Refuting opposing claims and evidence.** Essentially mirror images of the foregoing processes, refutation elements may be used in multidimensional multiple-choice questions just as supporting elements are used.

6. *Asking analytical questions for new, unknown facts to test claims.*

Students should seek evidence unknown to them that might further support, weaken, or even upset given claims. The test maker constructs a claim arising from the assigned reading and then considers facts missing from the assignment that would further test the claim. For the following example, readings in I. Bartlett, E. Fenton, et al., *A New History of the United States: an Inquiry Approach* (New York, 1969), were assigned.

*Claim:* During economic recessions the United States government has characteristically attempted to inflate the currency. Which of the following is the analytical question that tests this claim?

- (a) What did the Populist Party platform say about ownership of railroads?
- (b) Why did Populists want bimetallism?
- (c) What did the New Deal do about unemployment in 1935?
- (d) What did the Federal Reserve do about the money supply in 1920?

Choices (a) and (c) are irrelevant; besides, the assigned reading already contains the information asked for here. Choice (b) will yield a claim, not a fact, for an answer. Correctly choosing (d) shows that the student recognizes an indicator of evidence not contained in the assignment, that the student knows a function of the Federal Reserve System, and that 1920 was a year of recession. Finally, further research into the claim offered would very likely require an answer to the question asked in (d).



Can students answer the multidimensional questions I suggest? Yes, provided processes of knowledge-based analytical skills are explicitly, systematically, and consistently taught and practiced throughout the school year. Ordinarily, twenty-five multidimensional questions fill a test hour, allowing about two minutes to think through each question. I have found that at the beginning of the year I can expect to give passing scaled grades to those who answer at least half correctly; by year's end the scale tends to move upward. A good way to prepare students is to ask them to make up their own multidimensional test questions; making these tests is at least as educative as taking them. For the teacher, constructing a fresh hour test of twenty-five questions is time-consuming. However, subsequent versions of the same questions can be quickly constructed by altering one or more elements.

Doubtless, integrating knowledge and critical thinking in its multiple-choice questions would require a deeper intervention in nationwide history instruction than The College Board or ETS has yet dared. But there is precedent for such a step. ETS itself does a marvelous job with similar testing in its *Law School Aptitude Test Bulletin* — which, oddly, students confront only after they are out of school, indeed, out of college! Readers familiar with the work of Edwin Fenton at Carnegie-Mellon

University will recognize elements of his inquiry methods in this paper. Perhaps the most striking precedent is found in the essay sections of AP history tests themselves, where students are implicitly invited to use the very processes outlined in this paper.

It is probably true that most history students are unfamiliar with the multidimensional learning and testing system I suggest. Perhaps this is a good reason for teaching them about it.