

Education Policy Analysis Archives

Volume 9 Number 39

October 3, 2001

ISSN 1068-2341

A peer-reviewed scholarly journal
Editor: Gene V Glass, College of Education
Arizona State University

Copyright 2001, the **EDUCATION POLICY ANALYSIS ARCHIVES** .
Permission is hereby granted to copy any article
if **EPAA** is credited and copies are not sold.

Articles appearing in **EPAA** are abstracted in the *Current Index to Journals in Education* by the [ERIC Clearinghouse on Assessment and Evaluation](#) and are permanently archived in *Resources in Education*.

When An "A" Is Not Enough: Analyzing the New York State Global History and Geography Exam

**S. G. Grant
University at Buffalo**

Citation: Grant, S.G. (2001, October 3). An "A" is not enough: Analyzing the New York State Global History and Geography Exam. *Education Policy Analysis Archives*, 9(39). Retrieved [date] from <http://epaa.asu.edu/epaa/v9n39.html>.

Abstract

Education Week's report "Quality Counts" judges New York State's curriculum and assessment policy efforts to be an "A." Surface-level reviews such as "Quality Counts" tell something about the workings of state policy, but they are more useful as snapshots than as well-developed portraits of curriculum and assessment change. In this article, I analyze the new New York State Global History and Geography standards and tests using a set of social studies-specific criteria which inquire deeply into the implications for real instructional change. From that vantage, I argue that New York's policy efforts, while seemingly well-intentioned and reflective of surface-level change, fail to promote powerful teaching and learning in social studies. Teachers intent on producing ambitious teaching and learning will find little to interfere

with their efforts. But as a set of reforms intended to encourage substantive change, the new global history test falls short.

By some reports, New York state has made considerable strides in redesigning its state standards and assessment programs. For example, the authors of Education Week's report, *Quality Counts* (see <http://www.edweek.org/sreports/qc00>) judge New York's efforts to be an "A." In that report, New York scored points for having new content standards in all school subjects and at elementary, middle, and high school levels; for having tests which employ multiple-choice, short answer, and extended response questions; for requiring passing state assessments for high school graduation; and for using a range of policy tools such as report cards, ratings, financial assistance, and state sanctions to encourage improved test performance.

Surface-level reviews such as *Quality Counts* tell us something about the workings of state policy, but they are more useful as snapshots than as well-developed portraits of curriculum and assessment change. Attempts to construct such portraits demand more rigorous criteria than whether a type of test item appears or not. When such criteria are applied in the context of the new New York state global history exam, it is hard to justify Education Week's lofty grade. In short, an A from Education Week isn't enough.

In this article, I do a document analysis of the new NYS Global History and Geography standards and tests using a set of social studies-specific criteria which inquire deeply into the implications for real instructional change. From that vantage, I argue that New York's curriculum and assessment efforts, while seemingly well-intentioned and reflective of surface-level change, fail to promote powerful teaching and learning in social studies. Teachers intent on producing ambitious teaching and learning will find little to interfere with their efforts. But as a set of reforms intended to encourage substantive change, new global history test falls short.

Design of the Study

Led by Patricia Avery from the University of Minnesota, several colleagues and I from universities around the U.S. developed a set of criteria by which to analyze the new state curriculum and assessments emerging in our respective states. (Note 1) Drawing on the current thinking in our field, especially as it is reflected in national standards documents (e.g., National Center for History in the Schools, 1994; National Council for the Social Studies, 1994) and state-level standards (e.g., New York State Education Department, 1998), we constructed criteria that ask to what extent the new state tests ask students to:

- demonstrate knowledge of significant concepts and issues in history and the social sciences?
- consider multiple perspectives on issues and events?
- manipulate and interpret social science data?
- engage in higher order thinking about significant social studies concepts and issues?

I operationalize these criteria in the sections which follow. Note, however, that these measures really pose two questions. The first inquires about the simple existence of each criteria listed, e.g., is there *any evidence* to suggest that students much

demonstrate knowledge of significant concepts? The second question implies a quality measure, e.g., to *what degree* must students demonstrate knowledge of significant concepts? The first kind of question is not unlike those asked by Education Week, where the singular appearance or absence of a criteria is deemed important. The second kind of question pushes deeper, asking about the importance or meaningfulness of the measure. Evidence of a measure is interesting, but the extent to which that measure is meaningful seems ultimately more useful.

The prevailing pattern that emerged from my analysis can be termed, "yes, but... ." *Yes*, there is evidence of attention to the subject-specific criteria we developed, *but* inquiries into that evidence suggest that the new global exam comes up far short of a substantive change.

Background on New York State Curriculum and Assessment

In New York state, the belief that tests drive change is alive and well. But while the notion that tests matter is widely held, little empirical evidence supports a robust connection between tests and learning. In fact, Stake and Rugg (1991) argue that "in sixty years of vast international research on school testing, the policy of emphasizing test performance in order to improve education has never been validated" (p. xx). If true, it is no surprise to learn that the available research suggests that the relationship between testing and teachers' practices is complicated at best (Cimbricz, in review; Cohen & Barnes, 1993; Firestone, Mayrowetz, Fairman, 1998; Grant, in press). Tests matter to teachers (see, for example, Smith, 1991a, 1991b), but how teachers interpret and act on the import of new tests is largely uncharted ground.

That little is known about if and how tests influence teaching and learning has yet to inhibit state-level policymakers in New York (and most other states) from using them. To understand the recent changes in the state assessment program, however, one needs to consider the long history of state involvement in curriculum and testing.

New York state policymakers draw on a long history of attempts to influence classroom teaching and learning. Administered for over 100 years, the Regents testing program tests high school students on standardized, criterion-referenced exams that are tied to state-developed course syllabi in all academic subjects. In social studies, students take the Global Studies test at the end of a two-year Global Studies course sequence in ninth and tenth grades; eleventh graders take the U. S. History and Government test after completing a course of the same name. State curricula and tests also exist for elementary and middle school teachers and students.

A Mix of Old and New in New York State Standards and Assessments

The most recent changes in the state curriculum and assessments began in the early 1990s under the previous education commissioner, Thomas Sobol. Richard Mills, commissioner since 1994, continued that effort. Interestingly enough, Mills came to New York intending to decrease the traditional emphasis on standardized testing. The education reform movement Mills led in Vermont resulted in a state-level assessment program based on student portfolios rather than on tests. Mills abandoned this approach in New York, however. Sensing that the state's draft curriculum frameworks were being largely ignored, Mills reportedly asked a teacher to explain. "'You don't get it,' the teacher said, with what Mr. Mills remembers as almost a sneer. 'If the standards are not on the test, they're not real'" (Hartocollis, 1999, B1). (Note 2)

That comment apparently proved key for Mills is now an unabashed supporter of

standards-based tests as a vehicle for classroom change. The *Learning Standards for Social Studies* (New York State Education Department, 1998) represent the state's latest K-12 curriculum; new tests in grades 5, 8, 10, and 11 are emerging over the next two years.

Compared with the previous round of curricular revisions in the mid-to-late 1980s, the new standards documents represent a mix of old and new. Virtually no change appears in the K-5 grades curricula, which continue to follow an expanding horizons model. There are also no discernible changes in the seventh and eighth grade U.S. and New York State history sequence, or in the twelfth grade Participation in Government and Economics courses. A modest change is evident in the eleventh grade U.S. history and government course in that a emphasis on geography surfaces. Major changes seem localized at sixth grade, where the course of study expands from Western and Eastern Europe and the Middle East to the entire Eastern hemisphere, and at ninth and tenth grades, where the emphasis has changed from a cultural approach as represented in Global Studies to a chronological, history-based study expressed as Global History and Geography.

The state-level testing program also reflects a mix of old and new. Compared to the tests in most other subjects, the new social studies assessments seem the least changed. Whereas the new mathematics, science, and English-language arts tests make liberal use of open-ended and extended tasks, the social studies exams continue to rely largely on multiple choice questions. Moreover, compared to the tests in sister subject matters, the multiple choice questions posed on the social studies exams seem directed toward lower levels of understanding.

The multiple choice questions notwithstanding, the new state social studies exams do look different from the old ones. The principal change is in the writing portion of the exam. Unlike many minimum competency tests, New York students have always had to write essays on state exams. The new tests are different primarily in the fact that a) students will no longer have a range of essay prompts to choose from, and b) a new kind of essay question, a document-based question (DBQ), is being introduced on each of the fifth, eighth, tenth, and eleventh grade tests. (Note 3) A DBQ asks students to write an essay synthesizing information from as many as eight primary source documents (e.g., short quotes from government documents and famous individuals, political cartoons, poems, charts and graphs). The DBQ from the Global History and Geography exam administered in June, 2000 is as follows:

Historical context : Economic systems attempt to meet the needs of the people. Capitalism and communism represent two different ways to meet people's economic needs.

Task: Using information from the documents and your knowledge of global history, answer the questions that follow each document in Part A. Your answers to the questions will help you write the Part B essay, in which you will be asked to:

Describe how these two economic systems attempt to meet the needs of the people

Evaluate how successful each system has been at meeting the economic needs of the people

This task is followed by eight documents, seven quotations (e.g., R.W. Emerson, Adam Smith, Friedrich Engels) and one political cartoon, which present contrasting views of communism and capitalism. One or two main idea questions accompany each document. An example of a document and the attendant question follows:

...masses of laborers...crowded into factories. They are slaves of the machine and the manufacturer. Instead of rising as industry progresses, they sink deeper and deeper into poverty... Karl Marx and Friedrich Engels, *The Communist Manifesto*

The attendant main-idea question is: "According to Marx and Engels, what was the effect of the capitalist factory system?"

After responding to short answer questions such as this, students are directed to:

- write a well-organized essay that includes an introduction, several paragraphs, and a conclusion
- use evidence from at least *four* documents to support your response
- include additional related information.

High school students will also write a second, "thematic" essay based on a single prompt. The thematic essay from the June, 2000 Global History and Geography exam is:

Write a well-organized essay that includes an introduction, several paragraphs addressing the task below, and a conclusion.

Theme: Justice and Human Rights--Through history, the human rights of certain groups of people have been violated. Efforts have been made to address these violations.

Task:

- Define the term "human rights"
- Identify two examples of human rights violations that have occurred in a specific time and place
- Describe the causes of these human rights violations
- For one of the violations identified, discuss one specific effort that was made or is being made to deal with the violation.

Students are then advised:

You may use any example from your study of global history. Do not use the United States in your answer. Some suggestions you might wish to consider include: Christians in the early Roman Empire, native peoples in Spain's American colonies, untouchables

in India, blacks in South Africa, Jews in Nazi Germany, Muslims in Bosnia, Kurds in Iraq or Turkey, or Tibetans in China.

Each of the two essays is scored by two classroom teachers on a 6 point rubric, from 0-5. On the DBQ above, a score of 0 "fails to address the task or theme, is illegible, or is a blank paper." By contrast, a score of 5:

- addresses all aspects of the task by accurately analyzing and interpreting at least four documents
- thoroughly describes and evaluates capitalism and communism
- incorporates information from the documents in the body of the essay and may cite from the document in an appropriate fashion, but does not copy the entire document
- incorporates relevant outside information such as the early British factory system, Stalin's five-year plans, collapse of communist system in the Soviet Union
- takes into account the point of view of the authors in the description and evaluation of capitalism and communism
- is a well-developed essay, consistently demonstrating a logical and clear plan of organization
- introduces the theme by establishing a framework that is beyond a simple restatement of the task or historical context and concludes with a summation of the theme

Scores between a 5 and a 0 reflect lesser attention to each of the points above. For example, under a score of 3, the third point states, "incorporates limited or no relevant outside information."

Once the tests have been corrected, teachers are directed to a conversion table on the back cover of their manuals. There, they total a student's multiple choice and short answer scores (total of 61 possible points) and then look across a series of columns from 0-10, which represent the least to most possible points on the two essays. At the cross-section of these two scores is a converted score which ranges from 0-100. In the past, students had to score a 65 in order to pass the exam. A 65 is still the targeted state score, but districts are allowed to lower the required passing score to 55 for the next couple of years.

If the new tests themselves are only modestly revised, two other changes seem more dramatic. One is that the new fifth and eighth grade tests will mirror the high school exams in form and will produce individual student scores. Previously, tests at those levels, termed "Program Evaluation Tests," were general knowledge exams aimed at helping teachers understand the effectiveness of their content and pedagogical decisions. The shift to Regents-like tests and individual student scores at lower grades seems intended to raise the stakes of these tests by tying them more directly to the high school Regents exams. The second change concerns the function of the Regents test. In the past, passing Regents tests in all academic subjects meant that a student earned a "Regents" diploma, a distinction of some note. Students who desired to could opt to take the less rigorous Regents Competency Exam (RCT) and earn a local diploma. Beginning in 2001, ninth graders will no longer have these options. The RCT is being phased out, and all students will have to pass five Regents examinations (English, mathematics, global history, U.S. history, and science) in order to graduate.

"Yes, But... .": Analyzing the NYS Global History Exam

By most any measure, NYS policymakers deserve credit for the curriculum and assessment revisions they have made. They might have taken a less ambitious route by leaving the state curriculum and tests largely unchanged or by reverting to a minimal competency exam. Since they did not, Education Week's grade of A may well be justified. But if the criteria applied are more rigorous and more specific to the subject matter of social studies than those used by Education Week, then other interpretations of the new standards and assessments seem valid.

Recall that I analyzed the NYS Global History and Geography exam by asking to what extent the new state tests ask students to:

- demonstrate knowledge of significant concepts and issues in history and the social sciences?
- consider multiple perspectives on issues and events?
- manipulate and interpret social science data?
- engage in higher order thinking about significant social studies concepts and issues?

Recall also that I split this question in two. First, I looked for the mere existence of each criteria. Second, I inquired about the quality of the evidence for each criteria. My analysis suggests that while evidence for each of the criteria can be found, in no case is the quality or meaningfulness of that evidence strong. In short, the answer to each question is, "yes, but... ."

Knowledge of Significant Concepts and Issues

To be sure, there is a whole lot of knowledge represented on the new global history exam. This claim prompts little surprise, however, given the scope of the course title (i.e., "*Global* history and geography"), the 27 single-spaced pages of the state curriculum, and the fact that the course is taught over two school years. A quick review of the curriculum and test suggests apparent attention to significant concepts and issues: Geographic influences, religious beliefs, economic systems, political forces, cultural practices, and international relations map across an array of developed and developing, ancient and modern civilizations.

Yet even a surface-level analysis begins to yield some problems. For while the testmakers develop items for a wide range of concepts and issues, a quick count of the multiple-choice questions offers some troubling patterns. One pattern is that questions related to western nations (i.e., Europe, including Russia/USSR) dominate the test: Twenty-four questions assess issues relevant to the west, while only 10 questions each are assigned to India/Asia and to the rest of the world (Africa, Latin America, Caribbean, South America, and the Middle East). (Note 4) A second pattern is that the numbers of questions related to early civilizations (8 questions) and the middle ages (9 questions) are notably subservient to those attached to the modern era (31 questions). (Note 5)

This latter pattern could be predicted for two reasons. One is that historians and social scientists know more about modern times than the past, so to see that truism reflected in the apportionment of test questions is no surprise. The second reason is that the state curriculum gives preference to the modern era (18 pages) over early (6 pages) and middle (3 pages) periods. Since the test is reputed to reflect the state social studies

standards, it makes sense that the ratio of questions would reflect the chronological preferences established in the curriculum.

The first pattern is harder to understand, however. First, the clear preference for western-based questions flies in the face of the national movement to be more inclusive of other cultures. While the debate over multiculturalism has been contentious, it is hard to understand why the testmakers would so clearly privilege western history. This action is also hard to understand from a curricular point of view. While New York policymakers' efforts at creating a multicultural curriculum have been variously praised (Cornbleth & Waugh, 1995) and excoriated (MacDonald, 1992), the rhetoric in the social studies standards appears to support a strong endorsement of a *global* perspective:

This curriculum provides students with the opportunity to explore what is happening in various regions and civilizations at a given time. In addition, it enables them to investigate issues and themes from multiple perspectives and make global connections and linkages that lead to in-depth understanding. (New York State Education Department, 1998, p. 71)

The decision to emphasize questions related to the west is especially difficult to defend when one realizes that within each of the chronological units described in the state standards is attention to western *and* non-western people, events, and issues. For example, the unit entitled, "Global Interactions (1200-1650)," is divided into four sections, two of which--Early Japanese History and Feudalism and The Rise and Fall of the Mongols and Their Impact on Eurasia--are explicitly non-western. (Note 6) European issues and events do dominate the later units as world and cold wars get heavy play. That said, on the relationship between the west and the rest of the world, the disparity between the state standards and the state test is stark.

The disparities between the nations and eras represented and between the state curriculum and exam are interesting, but really do not help us understand whether the concepts and issues portrayed are *significant*. But then what constitutes a significant event turns out to be a pretty thorny issue, both for historians (see, for example, Carr, 1961) and for students (Barton & Levstik, 1997, 1998; Grant, 2001; Seixas, 1994, 1997). One might debate the relative merits of questions related to Karl Marx v. Ho Chi Minh, but it seems that with few exceptions the test addresses the big ticket items of a standard account of global history.

And that's part of the problem. The disparity in questions between western and non-western nations notwithstanding, the real issue related to significance is the type of questions asked rather than the content. In short, test makers aimed at low-level knowledge questions rather than at higher-order thinking questions. As a case in point, consider this multiple-choice question:

The Magna Carta, the Glorious Revolution, and the writings of John Locke all contributed to Great Britain's development of:

1. absolute monarchy
2. ethnic rivalries
3. parliamentary democracy
4. imperialist policies

Typical of the multiple-choice section, this question reflects an emphasis on generally expected, and clearly western constructs and events. But while the significance

of these elements to global history is undeniable, the question merely asks students to identify and label them. I address this notion of insignificant questions about significant events more directly in succeeding sections. For now then, my analysis suggests that, yes, the new state global exam does demonstrate attention to important concepts and issues, but does so in a way that may not push students' thinking.

Multiple Perspectives

The criterion of the extent to which the new state test addresses multiple perspectives is another case of "yes, but... ." While the inclusion of the DBQ indicates a move toward multiple views, that move is less apparent in the multiple choice section than one might expect. Moreover, the heavy tilt toward western themes undercuts the range of perspectives possible.

Several multiple-choice questions appear to reflect diverse perspectives because they give students multiple pieces of information. On closer inspection, however, all but two questions present compatible rather than diverse viewpoints. Typical of this kind of question is the following:

Base your answer to question 10 on the statements below and on your knowledge of social studies.

- *Statement A*: The might of a country consists of gaining surpluses of gold and silver.
- *Statement B*
: A nation's strength is found in economic independence and the maintenance of a favorable balance of trade.
- *Statement C*
: We need to gain colonies both a sources for raw materials and as markets for our manufactured goods.

Which economic system is being described by these statements?

1. traditional
2. feudal
3. command
4. mercantile

Students read three different statements, but each statement is necessarily tied to the others as a vehicle for defining mercantilism. Rather than dealing with multiple perspectives, then, students must only deal with multiple pieces of information. (Note 7)

The two multiple choice questions which do ask students to untangle multiple views employ the same stem:

Base your answers to questions 46 and 47 on the speakers' statements below and on your knowledge of social studies:

- *Speaker A*: The gods approached Vishnu, the lord of

creatures, and said: "Indicate to us that one person among mortals who alone is worthy of the highest rank... " Vishnu reflected, and brought forth a glorious son who became the first king.

- *Speaker B*: The traditional African society, whether it had a chief or not, was a society of equals and it conducted its business through discussion.
- *Speaker C*: Ideally, the best form of government is one where every citizen not only has a voice, but also, at least occasionally, is called on to take actual part.
- *Speaker D*: A monarch's authority comes directly from God, and this is how the leadership and power in a society should be determined.

46. Which speakers would support the theory explaining the power of France's Louis XIV, Spain's Philip II, and England's Elizabeth I?

1. A and D
2. B and C
3. A and C
4. B and D

47. Which speakers would agree with the idea that some form of democracy is the best way to govern a society?

1. A and D
2. B and C
3. A and C
4. B and D

One could quibble with the fact that the four statements represent only two views of government, but that really is a quibble. The questions might have been worded more clearly (especially #46), but the point remains: Students must be able to sort through differing views of government in order to make sense of the questions posed.

What seems like a similar quibble above rises to the level of critique in the DBQ. The eight documents divide cleanly into four categories: those that support capitalism (an excerpt from an unidentified work by Ralph Waldo Emerson; an excerpt from Adam Smith's *Wealth of Nations*), those that support communism (a quote attributed to "Katia," a 16-year-old ninth grader from Moscow in the 1980s; an excerpt from Friedrich Engels, *Principles of Communism*; an excerpt from Harry Schwartz in *The New York Times*, 1952), those that critique communism (an excerpt from, "The Peasant Wars on the Kremlin," by T. P. Whitney; a political cartoon from the *Providence Journal Bulletin*), and those that critique capitalism (an excerpt from *The Communist Manifesto* by Marx and Engels). (Note 8) Those who would question the documents selected would rightfully emphasize the clean lines of support for and critique of each system. There is no gray area here, for depending on the source, capitalism and communism are either portrayed as sin or salvation. In the first part of their essays, students are asked merely to describe how each system attempts to meet its citizens' needs. Since five of the eight documents provide clear fodder for this task, it hardly seems a significant challenge. The

second task, to evaluate how successful each system has been in meeting its citizenry's needs, seems more cognitively provocative. Here, students would presumably draw on the documents which critique each system. But notice what is missing: Students are provided only with partisan critiques. No data appear, for example, on how citizen-workers have fared under the respective systems. Presumably, students will draw on their knowledge that some countries like the former Soviet Union have renounced communism. But without more and better data, and especially data that offers direct comparisons, it is difficult to see how students can do much with this task.

Those who would defend this DBQ might counter that even a weakly constructed DBQ offers a profoundly different task than students normally undertake on a standardized test. That so much of the testing in social studies relies on multiple-choice questions has long been a sore spot among social studies educators. Clearly, this DBQ offers a new opportunity for assessing students' knowledge and skills.

Taken together, these points underscore the "yes, but..." argument about the new global exam. Including a DBQ ratchets up the substance of the test and begins to promote the notion of multiple perspectives. But the strength of that claim is undercut by what, with seemingly little effort, could have been a more powerful experience. Substituting documents that presented more nuanced views of capitalism and communism and that presented some comparative data would have gone considerable distance in beefing up a fledgling effort.

Manipulating and Interpreting Social Science Data

As noted above, the authors of the DBQ could have enhanced the student tasks by including some comparative data. Doing so would have contributed greatly to the generally weak way that social science data are handled on the global exam.

The types of questions represented on the new test generally call for definitions of terms (e.g., limited monarchy, totalitarianism, NAFTA) and identification of people, events, and social trends (e.g., Napoleon Bonaparte, French Revolution, democracy in Latin America). Few questions probe much below a surface-level knowledge of global history. And of those questions, a mere handful deal with social science data. To be sure, there are questions which employ illustrations, political cartoons, and maps. None of these, however, qualifies as data in the sense that students are presented with information that they must manipulate and interpret in order to answer the attendant questions. Of the 50 multiple choice questions, then, only three call upon students to use data. One question presents students with two circle graphs of the world population by region. The first graph shows the distribution for Europe, China, Latin America, North America, India, and four other areas in mid-1992; the second graph projects the distribution for the same regions in 2025. Two questions follow:

Which factor best explains the projected change in China's population by 2025?

1. increased immigration to China
2. religious doctrines discouraging birth control
3. government limits on family size
4. increased agricultural production in China

Which conclusion about world population in the next 25 years is

best supported by the information in these charts?

1. Technological improvements will cause a population decline throughout Asia.
2. Developed nations will be home to a majority of the world's population.
3. Efforts to curb population growth in developing nations will be successful.
4. Africa may experience problems with overpopulation.

A few questions later, students encounter a chart describing Internet usage in countries across the world. Three categories of usage along with representative countries are listed. For example, "heavy usage" countries include Canada, Norway, and the United States; "medium usage" countries include Chile, Britain, and Argentina; and "little use" countries include Mexico, Columbia, and Saudi Arabia. One question follows:

Which conclusion about Internet usage can be drawn from this chart?

1. Developing nations have easier access to the Internet than developed nations do.
2. A high standard of living in a nation is linked to high Internet usage.
3. Internet usage limits international cooperation.
4. Eastern Hemisphere nations use Internet connections more than Western Hemisphere nations do.

The final data-based question features a web diagram of automobile production using straight lines and arrows illustrating the global connections between auto companies and the countries in which they originate. For example, Chrysler/USA is connected by a straight line to Renault/France and to Hyundai/South Korea and by an arrow to Mitsubishi/Japan. The distinction between straight lines and arrows is not explained. One question follows:

Which conclusion can be drawn about global economics in the 1990s?

1. Countries became more economically isolated.
2. Higher tariffs reduced trade between nations.
3. France dominated the world automobile industry.
4. Economies of the world were increasingly interdependent.

These questions meet the ostensible parameters of data-based situations: Students are presented with some data from which they must infer trends. That said, there are at

least two problems with these questions. One problem is that students need not manipulate any of the data in order to answer the attendant questions. Students must make an interpretation, but in all cases the "right" answers are fairly obvious.

The reason the answers are so obvious speaks to a second problem: Not one of the questions demands much in the way of prior social studies knowledge. In short, the questions are cast such that only one answer makes common sense. Consider just two examples. First, without knowing anything about China, the question about the projected change in its population can only be reasonably answered with response #3 since it is the only answer which explains a declining populace. The answer to the question about global economics is just as commonsensical. The only response consistent with the web-like diagram is #4 which features the language of "increasingly interdependent." (Note 9) True, students need to know the vocabulary used--population, interdependent, and the like--but these are hardly arcane words used only in social studies contexts. So while students are asked to make inferences from the information presented, not only are they low-level inferences at best, but the possible answers are phrased such that the answers are obvious.

Once again, then, the surface-level qualifications of the NYS global exam pass muster, but a peek below that surfaces undercuts any confidence in the A grade assigned by Education Week. As with each of the preceding criteria, the new test fails to push students' thinking in substantive directions. The appearance of asking students to manipulate and interpret data is not enough.

Higher Order Thinking

It is probably clear by now that this last criteria, the extent to which the new test asks students to engage in higher order thinking about significant social studies concepts and issues, lies at the heart of my critique. The test makers can legitimately claim *some* attention to each of the preceding criteria. On the level of that attention, however, reasonable objections can be lodged. I will not speculate as to why the exam was constructed in this manner, but that it came so close to being a rich experience for students only to fail, is discouraging.

Consider two examples of how the exam questions might have been enriched. I argue above that the DBQ is composed entirely on partisan views of capitalism and communism. A small, but significant improvement would be to substitute a graph offering descriptive data on the comparative economic productivity and/or the social service conditions of the two nations. Such an addition would not only expand the range of documents students consider, but it would also help them make a more reasoned response to the portion of the essay prompt that calls for them to "evaluate how successful each system has been at meeting the economic needs of the people."

The multiple choice questions might also have been improved. Consider this example from the 1994 NAEP Geography Assessment (National Assessment of Educational Progress, 1994):

Statistical Comparison of Two Countries

	Country A	Country B
Total Population	7,193,000	123,120,000
Urban-Rural Urban	49.0%	76.7%

Rural	51.0%	23.3%
Religions	Rom. Cath.:92.5% Baha'i: 2.6% Other: 4.9%	Shinto*: 89.5% Buddhist*: 76.4% Christian: 1.2% Other: 9.3%
Life Expectancy at Birth (years)		
Male	50.9	75.9
Female	55.4	82.1
Age Distribution		
Under 15	43.4%	19.0%
15-29	26.4%	21.6%
30-44	15.7%	22.4%
45-59	9.3%	20.1%
60-74	4.4%	9.2%
over 74	0.8%	7.7%
% of the Population over 25 with No Formal Schooling	48.6%	0.3%
Leading Exports (as % of total exports)	Natural Gas: 21.0% Tin: 12.0% Zinc: 5.7% Silver: 5.6% Antimony: 4.0% Coffee: 2.0% Sugar: 1.5% Hides: 1.4%	Motor Vehicles: 18.4% Machinery: 10.9% Iron and Steel: 5.8% Chemicals: 5.3% Textiles: 2.6% Vessels: 1.5% Radios: 0.8% Televisions: 0.7%
		(*Some persons practice both religions.)

Which of the following statements most accurately describes Country A?

- A. It is dependent on raw material exports.
- B. It probably has a high literacy rate.
- C. It has a predominantly urban population
- D. It will experience slow population growth.

Which of the following statements most accurately describes Country B?

- A. It has few medical facilities.
- B. It is industrialized.
- C. Its primary imports are manufactured goods.
- D. Its population is primarily employed in agriculture.

Country B is most likely

- A. Botswana
- B. India
- C. Ireland
- D. Japan

Like the questions on the NYS global exam, these examples require understanding of significant social studies terminology. Unlike those questions, however, these examples push students to do more than define those terms. The first two questions demand that students evaluate the data in each cell and to draw conclusions across those cells. For example, in the first two questions, each of the possible answers directs students toward at a different cell of data. To select the best answer, then, a student must evaluate the information across the chart. The third question also asks students to look across multiple cells, but it adds a twist: Students must compare their assessment of Country B's attributes with their previous knowledge of world countries in order to select the best response.

These two brief examples point to the possibilities missed on the current exam. It seems safe to say that the bulk of the new global exam aims at low-level knowledge and understanding. The majority of the objective questions call for defining terminology, identifying significant people, places, and events, and in the case of the short answer section, describing the main point of a document. Surprisingly, the essays push no harder. The thematic essay asks students to complete several tasks, but by giving the students numerous examples of human rights cases, it is difficult to imagine many students struggling. The DBQ seems similarly poised. Students must synthesize the views from eight different documents, but there is no nuance in any of them and the clean divisions among them play directly into the tasks to which students are assigned. Taken together, the array of questions on this exam promise much. They do not deliver.

There is one more dimension that is worth note under the general criteria of higher order thinking. The NYS exam presumably scores high on the Education Week criteria in part because of the "extended response" items or essays. Moreover, the DBQ seems designed to signal a change in the structure of the social studies exams: One might argue that such a question represents a major shift away from traditional testing and toward more authentic assessment of students' historical understanding and reasoning. The scoring guide for the test, however, mitigates that claim: In short, students can easily pass the test without a single DBQ point. In fact, students can pass the exam without any essay points at all. A conversion chart on the last page of the teacher guidelines indicates that if students total a minimum of 54 points from the total of 61 possible multiple choice and short answer questions, they pass with a converted score of 65 regardless of whether or not they even attempt the essays. (Note 10) In this light, one can argue that the written portion of the new test has been substantially discounted compared to the previous exams. Where the essays once counted for 45% of a student's score, they now account for only 29%. Thus students can leave the essays blank, answer correctly

approximately 72% of the multiple choice and short answer questions, and still pass the exam. Adding the DBQ, then, can be read as a minor revision at best.

Implications

Since the mid-1990s, state policymakers have introduced a number of curriculum reforms such as new state standards for social studies. Preliminary indications (Grant, Derme, Gradwell, Lauricella, Pullano, & Tzetzso, 2000) suggest, however, that NYS global teachers view the curriculum and assessment changes as a mixed bag. Some applaud the state's move to a chronological approach as a more coherently historical mode. Others condemn this move (and some individual teachers and whole departments have rejected it) arguing that it undercuts the power of a cultural studies approach.

More important than the curricular changes, however, are teachers' concern about the new state tests (Grant, 1997a, 2000). This makes sense for two reasons. First, the curriculum documents produced thus far offer teachers little assistance in making concrete instructional decisions (Grant, 1997b). Second, the messages teachers receive often promote the view that tests are intended to drive change (Grant, 1996). For example, during sessions introducing the new state social studies standards, one representative from the New York State Education Department said that new tests will "help grow change in the system." During another session, a different SED representative said, "New assessments will represent a change in instruction....Kids won't perform well until (teachers') instruction reflects this." And at yet a third meeting, NYS Commissioner Richard Mills added, "Instruction won't change until the tests change." The message that tests matter also surfaced during local school and district meetings. A suburban district social studies supervisor, for example, told teachers that "change in content will come if we change the tests." An urban district supervisor observed, "If we change the assessments, we'll change instruction" (Grant, 1996, p. 271). One might question the focus of test influence--instruction, curriculum, or the "system" in general--but it is hard to miss the larger point: tests matter.

But how the new tests will matter deserves continued investigation. Our initial work in this area (Grant, Derme, Gradwell, Lauricella, Pullano, & Tzetzso, 2000) suggests that teachers' views of the new tests reflect some ambivalence. Most teachers support the use of documents and the DBQ. Yet from what teachers have seen in the test sampler disseminated by the state education department, few see this move as necessitating a fundamental shift either in their own pedagogies or as indicating a fundamental shift in the state's emphasis on social studies knowledge as represented in multiple-choice questions.

The analysis above, which focuses on the first test administered to NYS tenth graders last spring, suggests that teachers have it about right: The new test represents little in the way of fundamental change, and so can be read as demanding little change in classroom practices. True, some teachers report a ratcheting up of anxieties by students, parents, and administrators as test scores become media fodder. But responding to test score concerns and responding to the tests at hand may be two very different things.

Notes

1. The impetus for this action was a symposium entitled, "State Standards-Based Assessments and the Social Studies" held during the annual conference of the National Council for the Social Studies, San Antonio, Texas in November, 2000. Pat Avery and I were joined by Robin Chandler (Kentucky), Jean Craven (New

Mexico), and Ceola Ross Baber (North Carolina).

2. Thanks to Sandra Cimbricz for bringing this quote to my attention.
3. Mock test items, called test samplers, are available for the grades 5, 8, 10, and 11 tests (see <http://www.emsc.nysed.gov/ciai/assess.html>). The first administration of the grade 5 test is scheduled for November 2001; the new grade 8 and 11 tests are scheduled for June 2001.
4. Seven additional questions lump together people, places, and events such that it is difficult to ascribe them to a category.
5. Three additional questions span these time ranges and thus are difficult to categorize.
6. The other two segments are: The Resurgence of Renaissance Europe and Global Trade and Interactions.
7. Alert readers will note that Statement C is the key to the correct answer. Students might consider Statements A and B, but these are general features of most economic systems.
8. The quote from "Katia" might be double-counted as both in support of communism and in opposition to capitalism in that, before the bulk of the quote which does the former, she offers this presumed critique of capitalism: "Capitalists are rich people who own factories and have lots of money and workers."
9. I recognize that the adverb "increasingly" is problematic since no comparative data is presented. Nevertheless, in testmakers' parlance, it is clearly the "best answer."
10. Even more startling is the fact that in those many districts that opted to lower the passing score to 55, students need only get 44 of the possible 61 points to pass.

References

- Barton, K., & Levstik, L. (1998). "It wasn't a good part of history": National identity and students' explanations of historical significance. *Teachers College Record*, 99(3), 478-513.
- Barton, K. C., & Levstik, L. S. (1997, March 1997). *Middle Graders' Explanations of Historical Significance*. Paper presented at the American Educational Research Association, Chicago.
- Carr, E. (1961). *What is history?* New York: Vintage.
- Cimbricz, S. (in press). State testing and teachers' thinking and practice: A synthesis of research. *Educational Policy Analysis Archives*.
- Cohen, D., & Barnes, C. (1993). Pedagogy and policy. In D. Cohen, M. McLaughlin, & J. Talbert (Eds.), *Teaching for understanding: Challenges for policy and practice* (pp. 207-239). San Francisco: Jossey-Bass.

- Cornbleth, C., & Waugh, D. (1995). *The great speckled bird: Multicultural politics and education policymaking*. New York: St. Martin's.
- Firestone, W., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, 20(2), 95-113.
- Grant, S. G.. (1996). Locating authority over content and pedagogy: Cross-current influences on teachers' thinking and practice. *Theory and Research in Social Education*, 24(3), 237-272.
- Grant, S. G.. (1997a). Opportunities lost: Teachers learning about the New York state social studies framework. *Theory and Research in Social Education*, 25(3), 259-287.
- Grant, S. G.. (1997b). A policy at odds with itself: The tension between constructivist and traditional views in the New York state social studies framework. *Journal of Curriculum and Supervision*, 13(1), 92-113.
- Grant, S. G.. (2000). Teachers and tests: Exploring teachers' perceptions of changes in the New York state testing program. *Educational Policy Analysis Archives*, 8(14), 1-16.
- Grant, S. G.. (2001). It's just the facts, or is it? An exploration of the relationship between teachers' practices and students' understandings of history. *Theory and Research in Social Education*, 29(1), 65-108.
- Grant, S. G.. (in press). An uncertain lever: The influence of state-level testing in New York State on teaching social studies. *Teachers College Record*. Grant, S. G., Derme, A., Gradwell, J., Lauricella, A., Pullano, L., & Tzetzko, K. (in press). Teachers, tests, and tensions: Teachers respond to the New York state global history exam. *International Forum for Social Studies*.
- Hartocollis, A. (1999,). The man behind the exams: New York's education chief pushes agenda of change. *New York Times*, pp. B1, B10.
- MacDonald, H. (1992). The Sobol report: Multiculturalism triumphant. *The New Criterion*, 10(5), 9-18.
- National Assessment of Educational Progress. (1994). *NAEP Geography Assessment*. Washington, DC: National Center for Educational Statistics.
- New York State Education Department. (1998). *Social studies resource guide*. Albany, NY: Author.
- Seixas, P. (1994). Students' understanding of historical significance. *Theory and Research in Social Education*, 22(3), 281-304.
- Seixas, P. (1997). Mapping the terrain of historical significance. *Social Education*, 61(1), 22-27.
- Smith, M. L. (1991a). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20(5), 8-11.

Smith, M. L. (1991b). Meanings of test preparation. *American Educational Research Journal*, 28(3), 521-542

Stake, R., & Rugg, D. (1991). Impact on the classroom. In R. E. Stake (Ed.), *Advances in program evaluation* (Vol. 1, pp. xix-xxii). Greenwich, CT: JAI Press.

About the Author

S. G. Grant

Associate Professor
517 Christopher Baldy Hall
University at Buffalo
Buffalo, New York 14260
716-645-2455 x1135
<http://www.gse.buffalo.edu/fas/grant/home.html>

Email: sggrant@acsu.buffalo.edu

S. G. Grant is an Associate Professor of Social Studies Education in the Department of Learning and Instruction at the University at Buffalo. His research interests lie at the intersection of state curriculum and assessment policies and teachers' classroom practices, with a particular emphasis in social studies. In addition to publishing papers in both social studies and general education journals, Dr. Grant has published *Reforming Reading, Writing, and Mathematics: Teachers' Responses and the Prospects for Systemic Reform* (1998; Lawrence Erlbaum Associates) and (with Bruce VanSledright) *Constructing a Powerful Approach to Teaching and Learning in Elementary Social Studies* (2001, Houghton Mifflin).

Copyright 2001 by the *Education Policy Analysis Archives*

The World Wide Web address for the *Education Policy Analysis Archives* is epaa.asu.edu

General questions about appropriateness of topics or particular articles may be addressed to the Editor, Gene V Glass, glass@asu.edu or reach him at College of Education, Arizona State University, Tempe, AZ 85287-0211. (602-965-9644). The Commentary Editor is Casey D. Cobb: casey.cobb@unh.edu .

EPAA Editorial Board

[Michael W. Apple](#)
University of Wisconsin

[John Covalleskie](#)
Northern Michigan University

[Sherman Dorn](#)
University of South Florida

[Greg Camilli](#)
Rutgers University

[Alan Davis](#)
University of Colorado, Denver

[Mark E. Fetler](#)
California Commission on Teacher Credentialing

Richard Garlikov
hmwkhelp@scott.net

Alison I. Griffith
York University

Ernest R. House
University of Colorado

Craig B. Howley
Appalachia Educational Laboratory

Daniel Kallós
Umeå University

Thomas Mauhs-Pugh
Green Mountain College

William McInerney
Purdue University

Les McLean
University of Toronto

Anne L. Pemberton
apembert@pen.k12.va.us

Richard C. Richardson
New York University

Dennis Sayers
California State University—Stanislaus

Michael Scriven
scriven@aol.com

Robert Stonehill
U.S. Department of Education

Thomas F. Green
Syracuse University

Arlen Gullickson
Western Michigan University

Aimee Howley
Ohio University

William Hunter
University of Calgary

Benjamin Levin
University of Manitoba

Dewayne Matthews
Education Commission of the States

Mary McKeown-Moak
MGT of America (Austin, TX)

Susan Bobbitt Nolen
University of Washington

Hugh G. Petrie
SUNY Buffalo

Anthony G. Rud Jr.
Purdue University

Jay D. Scribner
University of Texas at Austin

Robert E. Stake
University of Illinois—UC

David D. Williams
Brigham Young University

EPAA Spanish Language Editorial Board

Associate Editor for Spanish Language
Roberto Rodríguez Gómez
Universidad Nacional Autónoma de México

roberto@servidor.unam.mx

Adrián Acosta (México)
Universidad de Guadalajara
adrianacosta@compuserve.com

Teresa Bracho (México)
Centro de Investigación y Docencia
Económica-CIDE
bracho dis1.cide.mx

Ursula Casanova (U.S.A.)
Arizona State University
casanova@asu.edu

J. Félix Angulo Rasco (Spain)
Universidad de Cádiz
felix.angulo@uca.es

Alejandro Canales (México)
Universidad Nacional Autónoma de
México
canalesa@servidor.unam.mx

José Contreras Domingo
Universitat de Barcelona
Jose.Contreras@doe.d5.ub.es

Erwin Epstein (U.S.A.)
Loyola University of Chicago
Eepstein@luc.edu

Rollin Kent (México)
Departamento de Investigación
Educativa-DIE/CINVESTAV
rkent@gemtel.com.mx
kentr@data.net.mx

Javier Mendoza Rojas (México)
Universidad Nacional Autónoma de
México
javiermr@servidor.unam.mx

Humberto Muñoz García (México)
Universidad Nacional Autónoma de
México
humberto@servidor.unam.mx

Daniel Schugurensky
(Argentina-Canadá)
OISE/UT, Canada
dschugurensky@oise.utoronto.ca

Jurjo Torres Santomé (Spain)
Universidad de A Coruña
jurjo@udc.es

Josué González (U.S.A.)
Arizona State University
josue@asu.edu

María Beatriz Luce (Brazil)
Universidade Federal de Rio Grande do
Sul-UFRGS
luceb@orion.ufrgs.br

Marcela Mollis (Argentina)
Universidad de Buenos Aires
mmollis@filo.uba.ar

Angel Ignacio Pérez Gómez (Spain)
Universidad de Málaga
aiperez@uma.es

Simon Schwartzman (Brazil)
Fundação Instituto Brasileiro e Geografia
e Estatística
simon@openlink.com.br

Carlos Alberto Torres (U.S.A.)
University of California, Los Angeles
torres@gseis UCLA.edu