

Think Globally, Assess Locally

by Janet Price

This school year, 2002-03, is the first year that my high school is administering the New York State Regents Global History and Geography Exam. This test consists of multiple-choice questions, short answers, and two essays. The content potentially tested by this instrument is daunting.

Up until now, the school's philosophy, and my own classroom practice, has emphasized depth over breadth of coverage and performance assessment over on-demand testing. The challenge, clearly, was to maintain our best practices while at the same time helping students pass the test, a requirement for the high school diploma.

Our review of recent Regents exams left us with some questions and pushed me to develop an action research study to determine how and to what extent my students would demonstrate mastery of global history concepts and content in these two very different arenas of the classroom and the high stakes exam setting. Previously, a study entitled "The Empire Strikes Back" compared students' classroom work in literature to a similar prompt on the English Language Arts Regents (in Meyers & Rust, 2003). I decided to do a similar study, comparing students' test essays on the global exam to their work on a few different classroom assignments.

While the two test essays seemed more in keeping with our emphasis on language development and thinking skills, we had grave concerns about how to prepare our

students for 50 multiple-choice questions that could range from Hammurabi's code through global warming. I reasoned that my focus should be on the essays because they are so closely aligned with Regents standards for learning and specifically for World History. Thus, I began this study feeling that my students would probably fare well on the essay but experience difficulty with the short-answer questions. What follows is the story of what I learned about teaching to the test.

Context

International High School at LaGuardia Community College is a small school that admits only students who have been in this country for less than four years and who have failed the Department of Education's English language test. In the early 1990s when the State Education Department first introduced tests in history and science as a prerequisite for high school graduation, the school felt teaching to such tests would interfere with the school's emphasis on in-depth projects, small group instruction, constructivist practice and portfolio assessment which the school community felt were essential to accelerating the language growth of English Language Learners.

On the strength of its outstanding graduation rate, low dropout rate, and consistently good performance on state tests, IHS applied for and received a five-year waiver from the state testing program. Beginning in the 1995-96 school year, graduates were required to present a portfolio of major projects in lieu of passing the Regents Competency Tests. The development of this portfolio system and its positive effect on student achievement are documented in "The Empire State Strikes Back" (in Meyers and Rust, 2003).

In 1997, the New York State Board of Regents decided to phase out the Regents Competency Tests and, instead, require passing scores on five new Regents exams as a condition for receiving a high school diploma. A group of public high schools in New York City, including IHS, formed the New York Performance Standards Consortium and sought state approval for a common graduation portfolio aligned to state standards as an alternative to the new tests. In April 2001, the state education commissioner formally rejected the consortium's proposal but gave member schools extra time to phase in the American history, global studies, science and math exams. (The schools were already required to administer the English Language Arts exam.) Thus, our current tenth graders, the class of 2005, will be the first in our school to have to pass five Regents exams to graduate.

It is customary to administer the global studies, science and math exams to tenth graders and the ELA and American history exams to eleventh graders. Some of my tenth grade students volunteered to take the global exam early, in January 2003, providing data for this study.

Method

How do students respond to similar tasks in dissimilar contexts? How do different kinds of history writing provide different opportunities to demonstrate mastery?

In our early analysis of the global test structure, social studies teachers at our school observed that there are two types of essays on the exam and that they are not that dissimilar to some of the work we assigned students and were clearly designed to assess skills and knowledge that we deemed important. One essay, called the "Thematic Essay Question" asks students to write about a particular topic, such as industrialization or

cultural diffusion but permits students to choose their own examples depending on what they have studied. The test also gives suggestions to help students choose topics. The second essay requires students to read mainly primary source documents, answer short questions about the documents, and write an essay.

On various versions of the test, students have been asked to compare such topics as revolutions or human rights in various countries. In the January 2003 version of this test prompt, students were asked to "compare and contrast the effect of geographic factors such as location and availability of resources on the political and economic development of Great Britain and Japan" using information from at least five of the eight documents provided --four maps and four short texts. The texts were undated, although judging from their antiquated writing style, some may have been written during the industrial revolution.

As I had done in my earlier study (in Meyers and Rust, 2003), I decided to create a document-based essay embedded in the class curriculum and then compare student performance on the test prompt and the class assignment. The in-class essay was part of an extended unit on 20th century wars. Prior to writing the essay, the students had made a list of the causes of World War I, learned about trench warfare, seen *All Quiet on the Western Front*, read war poetry and designed group diagrams connecting the Versailles Treaty to the rise of Adolf Hitler. They had also kept current event logs on the developing situation in Iraq and, in a class discussion, they had compared the unfolding Iraq story to the events leading to World War I. For instance, students discussed whether the American interest in Iraqi oil could be compared to the European imperialism that was a cause of World War I.

For the in-class essay, students read 10 to 12 documents concerning the causes of World War I, World War II, and the recent war in Iraq. Their task was to use at least three of the documents to compare and contrast the causes of two of the wars. They were given appropriate background on the authors and dates of the documents to help them evaluate them. So for instance, since they were reading an excerpt from *Mein Kampf*, I explained when and where Hitler had written this book. Moreover, students were encouraged to discuss the documents with each other in small groups before writing an individual essay. While the in-class essay was one that students were required to revise, for purposes of this study, I used their first draft, unaffected by my corrections and comments, to compare with their essay on the exam. Nonetheless, there was not only classroom groundwork and context but also ample opportunity for student reflection and revision in a low-pressure environment.

To compare performance on the classroom DBQ essay and the Regents essay I used the state's own grading criteria for the Regents essay, obtained through the same training booklet graders had been used to score the January exam. (I did not participate in grading that exam.) I included the 16 students who had both taken the January global Regents and completed a first draft of the classroom essay.

Other major projects completed by the students this year were simply too unlike a test prompt to use to make a direct comparison. However, I looked at two projects, to see how they aligned with state standards in World History. The French Revolution diaries, point of view writing based on both classroom study and independent research, could not be directly compared to either exam essay. I analyzed excerpts from some of my subjects' diaries to identify what kinds of knowledge or skills are demonstrated by their

work. This allowed me to compare the opportunities to demonstrate mastery afforded to students through a more creative project as opposed to the on-demand exam essays.

Just before the January Regents, my students turned in research reports on a revolution of their choice, the culminating activity in the revolution unit. Since a research report is a different kind of task than a closed-book essay, the essay rubric was not a useful tool. However, as luck would have it, the January test was on the same topic as the reports my students had turned in the week before. The test directions were to "Choose one political revolution" and explain the causes, describe the effects on society and "evaluate whether the changes that resulted from the political revolution resolved the problems that caused it." These directions were remarkably similar to the framing questions I had given the students to organize their reports. I was, therefore, able to compare their effectiveness in answering these questions in the two different forums.

Findings and Analysis

1. Test results

As students were actually taking the test, I took a long, hard look at it. I did not expect them to do well on the multiple choice. Many of the questions were on topics we had not studied. Moreover, students had been given a mere hour of practice with answering multiple-choice questions. On the other hand, I had every reason to expect that students, at least those who had advanced enough English skills, would fare well on the two essays. Not only was the thematic essay on the subject that they had studied for a full semester, but they also had some classroom grounding on the DBQ topic. The directions on the document-based essay were to "compare and contrast the effect of geographic factors...on the political and economic development of Great Britain and

Japan." Two months before, my students had studied British geography as a factor in the industrial revolution. Just before the January Regents, they had reviewed in a special test prep class how Japan's physical proximity to China had affected its development. Thus, they had some classroom background to guide them in interpreting the documents.

I was floored by the test results. Eight of my students, about one-third of the test takers, passed the exam but they did so on the strength of their multiple choice and short answers scores, not their essays. A few students who actually did better on the essays failed the exam. It turns out the scoring is heavily weighted towards the multiple choice and the short answers as demonstrated by the January 2003 grading grid shown in Table 1 below (see Appendix). Indeed, if a student does very well on these parts of the test, they can pass without even writing the essays.

Why did some students do so well on the short answers? Many of the DBQ questions can be answered by copying appropriate phrases from the documents themselves. And many of the multiple choice questions include maps, cartoons, graphs and quotations that can be correctly interpreted without any prior knowledge of the history content. On the other hand, both the multiple choice questions and the documents were dense, wordy and full of terms and idioms that would be unfamiliar to most English language learners. Verbatim translations were available but only in a few of the languages our students speak. And as on any multiple choice tests, test-taking skills were important.

What surprised and dismayed me was how poorly many students did on the essays, including students who turn in excellent essays in class. For example, one student had just turned in a workmanlike report on Simon Bolivar in which she had answered

most of the questions posed in the essay prompt. Yet on the test itself she only scored a 1 out of a possible 5 points for her essay on the Spanish American revolution. I asked her why she hadn't used more of the information from her report and she told me that by the time she got to the essays she was "just too tired."

2. Comparing test results on the Regents document-based essay to a documents-based essay completed as a classroom assignment:

The table in Appendix A compares the scores students received on their Regents essay with the scores I gave them for their classroom essay, using the same criteria as the Regents. Possible scores range from 0 ("fails to address the theme, is illegible, or is blank paper") to 5. Three students scored three points higher on the class essay. Four students scored two points higher. Five students scored one point higher. Three students scored the same. No students scored higher on the Regents exam.

Why did the students do so much better on the classroom assignment? Certainly, the lack of test pressure and the ample time to read, reflect and revise created better conditions for demonstrating mastery. Obviously, students should do better when they are using documents to further explore topics they have been recently studying in class. However, excerpts from student work on the two different tasks illustrate another possible reason for the huge discrepancy in performance, particularly considering that the Regents task was conceptually simpler and most of the documents were shorter and less difficult than those used in conjunction with the class assignment. Comparing student work on the two tasks, it is clear that students found it hard to work up enthusiasm for the rather dry Regents topic whereas most students seemed to enjoy the opportunity to compare the causes of the Iraq war to World War One or World War Two.

Consider, for instance, student A's conclusion to his Regents essay which received a score of 1. He writes: "Japan and Great Britain are very wonderful place because there is water all over. Fishing Industries make a lot of money and trading is good too." The essay ends with a drawing of a guy sticking his tongue out.

In contrast, here is how student A ends his essay comparing the causes of World War II and the war in Iraq:

"Many people sacrifice their life for their country. Some are forced and some do it willingly. I salute all the people there are fighting to keep America safe. I don't like war but when there are bad, cruel, ruthless leaders in the world, war is the only way to stop them from going wild. I hope the communist government of China learns a big lesson after the U.S. and British army defeats and Liberates the Iraqi people from Shadam Hussein. May be one day hope I can see my country [Tibet] get freedom like the Iraqi people got from Shadam Hussein and Afghans from the Taliban, It would be the happiest day in my life. This lesson has taught me a lot about the wars that happened and the wars that are happening."

Most students, of course, took the Regents essay more seriously than A. Some tried to make connections between the geography of Great Britain and Japan and the effects on the peoples these ocean nations colonized. For instance, Student I writes "Both of these countries wanted to be powerful nations. Japan wanted to be like England, of the East it wanted to acquire some countries. Korea and Manchuria are some examples of this."

Student F goes even further in using the documents as a starting point to write about imperialism: "Many lands had been taken over by other countries. They have established there but later people start to gain their land again and make it like it was before but they learn how to build their countries stronger and richer. Just like the United States this country started by 13 colonies being govern by an imperialist country. They gained their independence and enlarge their country by conquering other lands."

Unfortunately, in their quest to do something interesting with the essay, both students move away from the essay task and directions. Both received a score of 2.

Students had some trouble working through the archaic language in some of the test documents under test-taking conditions but were able to accurately interpret equally difficult documents on the classroom essay. Here is a portion of a Regents document that gave students a hard time:

"a country where broad and gently-flowing rivers act as natural canals will have advantages in internal communications over a country broken up by mountain ranges...When we recognize that England is rich in these advantages, that she has coal and iron lying close together, that her sheep give the best wool, that her harbours are plentiful, that she is not ill-off for rivers and that no part of the country is farther than some seventy miles from the see, we have not said all..."

Here is how student F interpreted this document: "...They also have rivers that act as natural canals that help them with internal communications over a country broken up by mountain ranges." Student M wrote "They don't have rivers, but that okay because because they have a lot of other things which they can make money with."

In contrast, on the class essay, both students did a stellar job interpreting challenging documents. Student M deftly interpreted a speech by Winston Churchill decrying Chamberlain's appeasement policy, writing: "In document#6 Churchill says that Britain should have stopped Hitler but they didn't it was their mistake. If they would have stopped them then this war wouldn't have happened."

Student F tackles a very daunting document, a speech by the leader of the German delegation to the Versailles Peace Conference that read, in part, "In the last fifty

years, the imperialism of all European states has chronically poisoned international relations. Policies of retaliation, policies of expansion and disregard for the rights of peoples to determine their own destiny, have contributed to the European malady which came to a crisis in the World War."

Student F writes: The German representatives at the Versailles Peace Conference in 1919 explained why imperialism led to war they said, "In the last fifty years the imperialism of all European states has chronically poisoned international relations" (document 5). In this speech he meant that fighting over colonies led to many different conflicts between European countries, poisoning their feelings towards each other.

Student C did a workmanlike job and scored a 3 on her Regents essay but there was only so much she could do with the dry documents. For instance she wrote: "Thanks to its [Japan's] cost they could get different kind of fish, they were rich in fish, so the start a fish Industry which then help them to get recruiting ground. Now know as one of the most strongest and most formidable navies of the modern times. (Documents 6 and 7).

Here, in contrast is how she used an excerpt from *Mein Kampf*, "He said (document 1) 'One blood demands one Reich,' with this he meant that all Germans demand one power, so all Germans must be governed by only a ruler. Then he also said 'Oppressed territories are led back to the bosom of a common Reich, not by flaming protests but by a mighty sword.' (p. 265) With this he meant that the territories and the colonies that Germany lost were going to get back to their owner not by protesting but with strong weapons, in other words they would have to fight for them."

Recently, a panel of history professors convened by the New York Performance Standards Consortium reviewed a Regents DBQ essay. According to the resulting report, several panelists approaching the question from divergent ideological standpoints concurred that the exam failed utterly to pose any intellectual or analytical task, and that it similarly failed to spark students' innate interest or curiosity and denied them any voice in the response they were expected to produce. Students were neither required nor expected to assume a point of view about history, nor were issues in history presented as fodder for debate. The panelists wondered how the material thus presented would ever engage students. (Performance Consortium 2002)

Judging from their responses to the Regents task, my students would concur. I question, however, whether it would be possible to construct an engaging exercise that encouraged students to assume a point of view about history within the time constraints of an on-demand half day exam. That sort of task lends itself better to classroom projects, sustained effort over time and performance-based assessment.

The panel report also notes a disconnection between the Regents exam and the state standards. Both the exam questions, and the sample essays produced in response to those questions, reflected an assumption that there was a only one correct interpretation of individuals and events. (Performance Consortium 2002) This completely contradicted the Learning Standards' expectation that students would become able to investigate differing and competing theories of history, and hypothesize about why interpretations change over time (Performance Consortium 2002). As one panelist expressed it, the exam was deficient in terms of the critical skills, which are part of the Standards but not at all reflected in the test. (Performance Consortium, 2002)

As a practitioner, I am concerned that interesting and engaging projects that in fact help students achieve mastery of important state standards do not seem to be useful in preparing my students for the high stakes exam. On the day of the test, faced with an essay on the subject they had studied for the better part of a semester--revolutions--my students were not able to use what they had learned by writing a research report on a revolution or writing a diary in the voice of an actor in the French revolution. Yet these projects both addressed state standards in a way a time-constrained essay could not.

In the commencement standard 2--World History, subtopic 1, the state says that students should be able to "analyze historic events from around the world by examining accounts written from different perspectives" and that "This is evident, for example, when students analyze important events and developments in world history through the eyes and experiences of those who were there as reported in their literature, diaries, letters, debates, art and music." In sub-topic 3 the state says students should "analyze the roles and contributions of individuals and groups to social, political, economic, cultural and religious practices and activities" and that "This is evident, for example, when students prepare narratives that describe important historic events and developments (e.g.... revolution...) from the perspectives of the individuals and groups who witnessed them." My students evidenced these masteries by choosing a character from the French Revolution (e.g. Robespierre, Napoleon, Marie Antoinette), finding information about and primary sources by that character and then writing diary entries discussing historical events from that person's perspective. The project also gave them practice finding and utilizing internet information and helped them develop the habits of revision and sustained effort over time. Above all, it engaged students. In their end of semester class

evaluations, most students pointed to this project as their favorite activity of the semester.

Here is a sample diary entry by a student who wrote as Napoleon:

October, 1805

Horatio Nelson had destroyed the combined Franco-Spanish fleet off Cape Trafalgar.

This loss has made it virtually impossible for me to invade England. Now what I do? Think, Think! I know...I'm going to introduce the Continental system, designed to exclude all British goods from Europe. In this manner I hope to ruin the British economy and to force the "nation of shopkeepers" to make peace on French terms.

Like I always say, "the most dangerous moment comes with victory."

Policy Implications

Should I give up a project that addresses state standards that are not effectively assessed by the high stakes, on-demand test given that the project is time consuming and does not seem to prepare students for that test? This is the dilemma faced by global history teachers across New York State today. How do we provide the depth necessary to address state standards and to make history meaningful and engaging to students while providing the breadth necessary to prepare students for a test that covers everything from the Neolithic Age to NAFTA?

Ultimately, this is not a choice that we should have to make. An on-demand test is not the best assessment of many of the state standards nor, as the comparison between

performance on the Regents test and the class essay demonstrates, is it the best way for many students to demonstrate mastery of skills and knowledge.

Multiple forms of assessment offer a fuller, fairer picture of achievement. As Orfield and Wald note in an article on high stakes testing, the National Research Council, who are national testing experts, the American Educators Research Association, the American Psychological Association, the National Council on Measurement in Education and the Department of Education all agree that no decision of serious consequence in a child's life should be made on the basis of a single test score. (Orfield and Wald, 2001)

Not only is over-reliance on high stakes testing potentially unfair. It may also have the unintended consequence of lowering rather than raising student achievement. A recent Arizona State University study by Amrein and Berliner of state testing programs in 28 states concluded that the implementation of high school graduation exams results in a decrease in academic achievement. After high-stakes graduation exams were imposed, ACT, SAT, and AP scores declined (Amrein and Berliner 2002). A companion study suggested that the adverse effects of high-stakes testing may outweigh any benefits. Among the many adverse effects identified in the study were higher numbers of urban school teachers, in particular, "teaching to the test," limiting instruction to only those things that are sure to be tested, requiring students to spend hours memorizing facts, drilling students on test taking strategies, and higher numbers of teachers who leave their public school positions to teach in private schools free of state testing because state rules make teachers feel compromised as professionals (Amrein and Berliner 2002).

On-demand tests pose special problems for English Language Learners such as my students at International. The lack of native proficiency in English makes them particularly vulnerable to the inherent problems of on-demand testing including arbitrary and unrealistic time limits, lack of context or meaning in the assessment tasks, inappropriate assumptions about prior

knowledge, and unfamiliarity with vocabulary and idiom. These problems are greater in history than in language arts. Thomas and Collier's 1997 study of English Language Learners found that students of all ages reached grade-level achievement in mathematics and language arts in a shorter period of time, but required many years to reach grade level in reading, science and social studies in English. (Thomas and Collier, 1997) Students who arrive after age 12 even with good formal schooling in their native language, had run out of time by the end of high school for their grades to catch up to the native English speakers, who were continually pulling ahead. (Thomas and Collier 1997) In his review of the literature on English language acquisition Garcia, 2000, notes, "Students with limited literacy in native language will take even longer to learn the English skills associated with science, social studies and higher order mathematics" (Garcia 2000, p. 9).

That is not to say that English Language Learners cannot meet history standards. However, high stakes, on-demand tests are particularly inappropriate for this group. Offering the test in the students' native languages is not an adequate solution. For one, the Regents tests are available currently in very few of the languages our students speak. Second, much of their study is done in English and they may in fact know more specialized vocabulary in English than in their first language.

Policy Recommendations

- End the practice of making any one test a determinant of who receives a high school diploma. If the global Regents is required, permit schools to use it as one piece of evidence in making high-stakes decisions.
- Grant variances to schools that wish to use performance-assessments aligned to state standards in lieu of the Regents global exam.
- Adjust the passing score for English Language Learners from 65 to 55 to partially compensate for the special difficulties the test poses for this population.

Bibliography

- Amrein, Audrey L. and Berliner, David C. (2002). *An Analysis of Some Unintended and Negative Consequences of High-Stakes Testing*. Education Policy Research Unit, Arizona State University. Available on line <http://edpolicylab.org>
- Garcia, G. (2000). *Lessons from Research: What Is the Length of Time It Takes Limited English Proficient Students to Acquire English and Succeed in an All-English Classroom?* Issue Brief. Washington, DC: National Clearinghouse for Bilingual Education.
- New York Performance Assessment Consortium (2002). *Regents Exam Critiques Report of Review Panel, New York State Regents Examinations in Global and American History*. Available online www.performanceassessment.org
- Orfield, G., and J. Wald. (2001). "High Stakes Testing." *In Motion Magazine* (29 April). Available online at www.inmotionmagazine.com/er/gojw.html.
- (Self Reference)
- Thomas, W. and V. Collier, (1997). *School Effectiveness for Language Minority Students*. NCBE Resource Collection, no. 9. Washington, DC: National Clearinghouse for Bilingual Education.
- The University of the State of New York, The State Education Department, (1996). *Learning Standards for Social Studies, Revised Edition*.

Appendix A:**Table 1: Comparison of scores on Regents document-based essay and class essay**

Student	Regents score (65=pass)	Score on Regents DBQ essay	Score on class essay
A	69	1	4
B	57	2	3
C	68	3	4
D	73	3	4
E	45	2	3
F	61	2	4
G	75	3	3
H	67	3	3
I	65	2	4
J	86	3	4
K	60	2	4
L	59	2	2
M	55	1	3
N	44	0	3
O	53	1	4
P	48	2	3