

Present at the Creation: Social Studies Supervisors Develop Their Own Assessment

By Andrea S. Libresco, Ed.D.

Crisis equals opportunity

Who hasn't said to her or himself, "Well, if I were in charge of designing the social studies assessment, here's what I'd do?" This is a study of a group of New York State social studies supervisors who, in response to state budget cuts, took matters into their own hands and developed their own regional social studies test at the elementary level.

The catalyst for the formation of this group of supervisors came when the monies to create and distribute the fifth and eighth grade social studies tests were eliminated from the state budget. The local council for the social studies (LICSS, 2010) drafted a letter to the governor to restore the funding for the tests, arguing,

cutting social studies tests sends a message to teachers that the subject doesn't matter....Teachers and principals, hard pressed to raise test scores and earn bonuses, will now concentrate on the rote learning of math and reading skills. Content, context and citizenship will be ignored. In this time of increasing immigration and civil conflict, American history and civic education are crucial to responsible citizenship. Don't let history become history in New York State. Restore the funding that is necessary for state testing in Social Studies.

When their request went unheeded, the supervisors, operating on the notion that crisis equals opportunity, stepped into the void left by the state and met to devise their own test. They developed assessments that they offered to all interested districts in the region to ensure that social studies continues to be taught, even in the absence of state assessments. Examining this process has lessons for those in all states who value social studies and want to forestall its disappearance. In addition, this process of developing a regional social studies test yields information about the effectiveness of the current state curricula, as well as what supervisors believe is worth teaching in social studies. Finally, the process illuminates how social

studies supervisors view their own power vis-à-vis that of principals in the areas of curriculum, instruction and assessment.

Theoretical framework

A significant body of research has explored the effects of No Child Left Behind's emphasis on language arts and math at the expense of other subjects and has documented the reduced time devoted to, or actual disappearance of, social studies education from elementary classes (Leming, Ellington, & Shug, 2006; McMurrer, 2007; Van Fossen, 2005; O'Connor, et. al., 2007; Rock, et. al., 2006). This literature informed this study's interest in chronicling the dynamics of a group of supervisors who seek to retain a test in social studies to maintain the existence of the subject at elementary and middle school levels. As Grant (2007) has pointed out, the mere existence of a test can be a more important factor in teachers' instructional decisions than whether or not it is a high stakes test (which the elementary social studies assessment has never been).

As I prepared to observe the supervisors' meetings that would result in the creation of a regional elementary social studies test, I wondered the extent to which they would be guided by past state tests and by the existing state curricula. It is the case that teachers protest one or more features of state test construction or the ways in which scores are interpreted, but few protest against the very existence of a test (Grant, 2007). Would supervisors explore options beyond past state templates? Although teachers' power as curricular instructional gatekeepers is well-documented (Thornton, 1991), I also wondered to what extent the supervisors would seek to assume that role to ensure that their teachers kept social studies (and particular content, concepts, and skills) in the curriculum (given that the principal pedagogical effect of state social studies tests appears to be on teachers' content decisions (Grant, 2007)).

The literature on effective teaching (Good & Brophy, 2007), wise practice (Yeager & Davis, 2005), powerful social studies (Brophy & Alleman, 2006), thoughtfulness (Newmann, 1990), and depth over breadth (Wiggins & McTighe, 1998) all begin with an assumption that big ideas, themes and concepts must be the starting point for good instruction. Thus, the research questions for this case study were:

- What are supervisors' considerations when designing a regional elementary social studies assessment?
- To what extent do supervisors' discussions indicate that they view such an assessment as an opportunity to promote effective teaching of powerful social studies, including depth over breadth?
- What constraints, real and perceived, exist in the construction of such an assessment?

Method

Over the course of four months, I attended and took notes at the meetings of the group of social studies supervisors (including K-12 curriculum associates and 7-12 department chairs) who came together to create a regional elementary social studies test. The group of three males and seven females, all Caucasian, participated in four half-day meetings: the first resulted in a decision to substitute a regional assessment for the defunct state one. At the second, supervisors brought teachers with them who, in turn, brought possible test questions. At the third meeting, supervisors sought to finalize the test; and issues regarding the purpose and format of the test arose again. Finally, at the fourth meeting, the assessment was completed. The discussions at the meetings and the final assessment design provided data about the possibilities and constraints on elementary level social studies.

This case study discusses the four meetings and focuses on the final two: the third, where the richest discussion took place; and the fourth, where the test was finalized. I examined my notes of the third meeting with three codes: supervisors' considerations in assessment design, discussions of effective teaching, and discussions of constraints on the process. In addition to these pre-set categories, three others emerged as sub-topics of supervisors' considerations in assessment design: satisfaction/dissatisfaction with the state assessment, civic efficacy as a consideration in assessment creation, and assessments as a tool to improve writing. Walter Parker defines civic efficacy as "the willingness to assume citizenship responsibilities, which include informed decision-making and the belief that one can make a difference (Parker, 2011, 26).

I analyzed the final assessment designed by the supervisors by sorting test questions into three categories: content (turning on a piece of data), concept (turning on understanding one of the New York State social studies

concepts), and skill (turning on analysis of a map, cartoon, reading passage, chart, timeline, or other document). These three categories of analysis reflect the designations that New York State uses within its K-12 curricula (NYS Resource Guide). Each question was assigned to the content, concept, or skill category for one point. If a question was deemed to fit two different categories, it was counted in both, and scored with 0.5 in each category.

The Supervisors' Meetings

First meeting: Keeping the test and some type of DBQ

The first meeting in December 2010 resulted in a decision to substitute a regional assessment for the defunct state one, where the majority of participants echoed one supervisor's conclusion, "We must have a test to keep one on the schedule and keep social studies in the classroom." After they concurred about the necessity of an assessment, supervisors unanimously agreed that they wanted to retain the documents-based question [DBQ] on the fifth grade assessment. The state assessment had included multiple choice, constructed response questions, and a DBQ.

Dissatisfaction with state assessments

Three supervisors spoke highly of the DBQ as "the historian's tool," but there was also dissatisfaction with the state's version of the DBQ (as opposed to the College Board's version as seen on Advanced Placement Exams):

The state took it and ruined it.

The scaffolding questions are lower level, the documents are often not meaningful or thought-provoking, and the way they're chosen, there's no possibility for outside information.

Currently, all they do is "promote formulaic writing" (without outside information).

The rubric is problematic as well.

An opportunity to improve writing

Even as most present expressed their frustration with the state DBQ, the group decided to keep the present format for the first year, so that teachers who were already in the midst of teaching the curriculum would not be upset and refuse to administer the test. One supervisor looked to the future for change: "Hopefully, we can, though this process of designing our own assessment, choose themes and documents that are meaningful." Another sketched out those changes in the format of the DBQ: "We can lessen the amount of documents, re-do the scaffolding questions - they're spit-back. Maybe we won't even have scaffolding." A third discussed future opportunities to improve students' writing with better questions:

We value writing and hopefully, with better DBQs, we can change the way we teach writing. I've been experimenting at my school with one of the new teachers (not already wed to official DBQ format) with essays that employ three documents, not eight, and the students can bring in their own information and write less formulaically, as they're not trying to stuff in all of the documents into one essay....This is the beginning of a redefinition of what social studies will look like in the classroom. And how we *teach* writing, not just *assign* writing.

Even as supervisors imagined the future, those present imagined tinkering with the state test, not switching to performance-based assessments. They remained grounded in the present. At least five times, a member of the group reminded everyone that "It's December already. This stuff will have to wait 'till next year." With this admonition, those present concluded the meeting, deciding that everyone would bring at least one teacher for the next meeting in January 2011, and those teachers would use existing test banks (past New York State tests are all available on the Department of Education website) to choose multiple choice questions, but they would create their own DBQs. This was necessary because the new test would not be based on the same material. The state fifth grade exam, given in November, had been based on material from grades three and four. The new regional fifth grade exam would be given in June and be based predominantly on grade five.

Second meeting: Reality check - teachers share DBQs, supervisors share administrators' reactions

At the second meeting in January 2011, supervisors brought teachers with them who, in turn, brought possible test questions.

Constraints on the process

Before the test questions were examined, supervisors reported back from their districts as to whether their principals and assistant superintendents would support this new exam. Although each of the supervisors' districts was ultimately on board, the discussions had been difficult; principals were reluctant to add another assessment. Supervisors shared their administrators' responses:

Teachers and administrators were dancing in the street that there was one less assessment.

It was really hard to convince them to add this social studies assessment back in when they were already celebrating its disappearance.

The real test is whether we can keep this test in the face of all of the other programs within the district; all anyone is worried about these days is APPR [Annual Professional Performance Review - that ties teacher evaluation to student performance].

I could only get agreement if this year's test looks like past year's tests and it's given in June.

Thus, supervisors reiterated that the exam would look basically the same for this year. Several added that they would have more flexibility to re-think it in the future.

Dissatisfaction with state curriculum

Teachers shared their DBQs, and it became apparent that there was no uniformity of content in what was being taught in fifth grade. After the meeting was over, one supervisor stated, "The fifth grade was rough. We realized that people were doing all different things in 5th grade - teaching it as a straight US history course, doing no present day economics, if they were doing Latin America, it was more like a festival than an analytical study." On some level, this was not surprising, as the fifth grade curriculum is un-teachable in its totality. According to the state scope and sequence, it includes: The history, geography, economics, and government of the United States, Canada and Latin America (and it is worth noting that this is the first time students study any United States history). Here is but one bullet point under "History:"

Key turning points and events in the histories of Canada, Latin America, and the United States can be organized into different historical time periods. For example, key turning points might include: 15th- to 16th-century exploration and encounter; 19th-century westward migration and expansion; 20th-century population movement from rural to suburban areas (NYS Department of Education, grade 5).

Because so much of the meeting was taken up with discussing what could appear in a DBQ, given the very different content each teacher taught, teachers handed in their DBQs to one of the supervisors present who agreed to compile the multiple choice questions the teachers had brought and develop a DBQ, based on the teachers' input, for the next meeting of the supervisors in February. Mini-interviews at the end of the meeting revealed that supervisors were concerned both about common content and teacher-created DBQs.

One supervisor commented, "Our teachers need some serious work in exam construction" and looked ahead to professional development that could follow the regional tests, "We hope part of this process will be an opportunity to train teachers on exam construction and powerful writing (which the exams should allow for and reflect)."

Third Meeting:

The third meeting in February 2011 emerged as the one that prompted the most robust discussion, as supervisors grappled with the extent to which they could or wanted to deviate from the state assessment.

Civic efficacy as a consideration in assessment design

The most sustained discussion occurred around the DBQ, first generically, then specifically. Several supervisors reiterated (from the first meeting) their support for a DBQ but one more "meaningful" than those on past state assessments, as well as one that would encourage the teaching of "thoughtful writing."

The specific DBQ brought to the third session for consideration was a comparison of United States government to Canada's parliamentary system. The question asked for the similarities and differences between the two systems (though not the more upper level, "Are they more similar or different?"), and documents included diagrams of the three branches of federal governance in the United States and an intricate diagram of Canada's branches. One participant argued that the key difference, that the prime minister comes out of the legislative branch, is a pretty complex understanding for ten-year-olds and worried that, in an essay on similarities and differences, students would say, "The US has a president, Canada has a prime minister." This produced a lively discussion about what students can and should be able to handle, and whether or not the state had made good curricular choices. Almost everyone present weighed in:

"The US has a president, Canada has a prime minister." That's low level comment and could be done in a fill-in, as opposed to an essay.

But that's all they can handle.

Then why teach it if they can't truly understand the different between president and prime minister?

Then you're just asking students to make literal, as opposed to analytical, comparisons without meaning.

I'd love to get something beyond finding information in the documents.

You're not even asking them to analyze, just asking merely to copy rote information from the document on prime minister and president.

But it will be background knowledge, we're building vocabulary.

But they won't truly understand that the prime minister comes from the legislative branch, so they're building literal vocabulary without meaning. Besides, this won't come up again until 10th grade.

But it's in the curriculum.

But the curriculum is undo-able; it's not as bad as the sixth grade curriculum, but it's still un-teach-

able - US, Canada, Latin America - history, economics, government, geography... So we should be able to focus on what we want. We don't have to do Canadian Parliamentary government in a superficial way, just to say we've done it.

The old "Rose post office" [that had documents where citizens organized to retain their community post office] was a good DBQ model - it was on citizenship, the most important attribute of social studies.

But that's not what they do in fifth grade.

They're not doing rights and responsibilities of citizenship?

This discussion reveals a number of the supervisors' concerns: what students can and should be able to handle (literal or analytical meaning, going beyond the documents), whether the state made good curricular choices or if the fifth grade curriculum was "unteachable," and whether past state assessments had sufficient emphasis on citizenship ("the most important attribute of social studies").

The session ended with those present deciding that the comparison of American and Canadian governmental systems was not the best choice for the DBQ. Participants made a date for a fourth meeting and promised to send the coordinator documents for an essay that compared the two countries more broadly, including their climates, levels of diversity, economics, and governments.

Assessment and effective teaching

The final comments of the session reiterated that the format and content of this year's test were governed by the context of timing, of having to convince administrators that they should buy into this test, that familiarity was the only way to get the maximum number of districts on board. Some of the most interesting comments may have come during the informal discussion, after the official meeting had ended. One supervisor remarked, "In many ways, my teachers got off to a better start without the test....Maybe without the test, they'd do more depth over breadth." Another responded, "Not unless you eliminate the tests in *all* of the subjects." And a third said, "Well, as that's not happening anytime soon, this is kind of the beginning of redefining the social studies. I'm fine with what this is because it's a process."

Fourth meeting: The assessment is finalized

At the fourth meeting in April 2011, supervisors settled on both multiple choice questions and a DBQ. Discussion of ratios of past state assessments guided construction of the assessment. Past state test multiple choice questions broke down to 28% content, 14% concepts, and

58% skills (Libresco, 2007). Out of the thirty-three multiple choice questions collected by supervisors (from past state exams and from teachers in their schools) and shared at the meeting, the group classified fifteen as content, seven as concept, and eleven as skills.

Because the group wanted to retain the past emphasis on skills, they looked to cut some content questions as they worked to bring the total number down to twenty-five. Ultimately, the group agreed on ten (40%) content, six (24%) concept, and nine (36%) skills questions. Interestingly, despite stated intentions of emphasizing skills, more questions were based on content. On the regional (and past state) assessments, the DBQ would be entirely a skills section, as it is based entirely on documents students read during the test.

The group approved an essay that compared the United States and Canada more broadly than the government DBQ discussed at the third meeting: "The United States and Canada share many similarities and differences in their histories, economies, government and geography. Describe two ways the United States is similar to Canada. Describe two ways that the United States is different from Canada." Students had to include evidence from at least three of the five documents in their essays. The five documents were:

- 1) pictures of the three most popular sports in Canada and in the United States;
- 2) diagrams of the branches of government in Canada and in the United States;
- 3) bar graphs of the top ten ethnic groups in Canada and in the United States;
- 4) a chart of the languages spoken at home in Canada and in the United States; and
- 5) pie charts, as well as a brief paragraph, on the sources of electric power in Canada and in the United States.

Supervisors were satisfied with this DBQ as appropriate for the first year without a state assessment. They were cognizant of the powerful need for principals and teachers to feel comfortable with a familiar format. Their comments indicated that they saw this as "a first step," "the start of a process," "a floor not a ceiling" for elementary social studies assessment and instruction.

Significance

A regional group of supervisors who recognize and exercise their power to create social studies assessments and, by extension, curriculum, may be a route for supervisors in other areas of the state and in other states and for other subjects (e.g. science).

The process, itself, has implications for social studies supervisors and practitioners, with its rich discussion of: whether past NYS assessments are the best model for

future assessments. One may infer that those present in these discussions were tinkering with the former tests, not switching to performance-based assessments. The participants seem to avoid the deeper discussion of what social studies curriculum ought to be at the elementary and middle school levels. The ideal ratio of content to concepts to skills in curriculum and assessment at different grade levels was not achieved. They did not address policy issues sufficiently to recommend the extent to which the civic efficacy purpose of social studies is or should be reflected in assessments; the extent to which an unwieldy state scope and sequence can or should drive instruction; and the power struggle between subject supervisors and principals over the importance of curriculum and assessments of subjects that have not been identified as such by either NCLB or Race to the Top.

In addition, the process has implications with respect to the role of professors of social studies methods. This researcher's knowledge of the group of regional assessment designers was serendipitous. No doubt, there are methods professors who could have contributed to this particular group and to other such groups of social studies supervisors. However, social studies supervisors' listservs do not tend to include methods professors, and methods professors' listservs do not tend to include social studies supervisors. Perhaps college professors can be more attentive to connecting the two constituencies. In a time of de-emphasis on social studies, this separation of university professor and public school practitioner may not be effectively serving our profession, nor our most important constituency, our students.

With or without their university colleagues, if more teachers and supervisors of elementary social studies were given time and space to have such conversations about purposes of social studies, appropriate curriculum for students, and the possibility and value of thoughtful assessments, how would that affect how teachers and supervisors approached curriculum, instruction and assessment in elementary social studies? That's a conversation worth having.

References

- Brophy, J. and Alleman, J. (2006). A Reconceptualized Rationale for Elementary Social Studies, *Theory and Research in Social Education* 34 (4), 428-454.
- Good, T. & Brophy, J. (2007). *Looking in classrooms* (10th ed.). New York: Allyn & Bacon.
- Grant, S.G. (2007). High-stakes testing: how are social studies teachers responding? *Social Education* 71 (5), 250-254.
- Leming, J. S., Ellington, L., & Schug, M. (2007). The State of Social Studies: A National Random Survey of Elementary and Middle School Social Studies Teachers. *Social Education* 70 (5), 322-327.

Libresco, A. (2007). A Test of High-Order Thinking, *Social Studies and the Young Learner* 20 (1), 14-17.

Long Island Council for the Social Studies [LICSS]. (2010). Letter to Governor Cuomo. Retrieved from <http://licss.org/>.

McMurrer, J. (2007). Choices, changes, and challenges: Curriculum and instruction in the NCLB era. *Center on Education Policy*. Retrieved from <http://www.cepd.org/index.cfm?fuseaction=document.showDocumentByID&nodeID=1&DocumentID=212>.

New York State Department of Education. Grade Five: The United States, Canada, and Latin America, *Social Studies Core Curriculum*. Retrieved from <http://www.p12.nysed.gov/ciai/socst/pub/sscore1.pdf>.

New York State Resource Guide and Core Curriculum, Social Studies, New York State Education Department. Retrieved from <http://www.emsc.nysed.gov/ciai/pub/pubss.html>.

Newmann, F. (1990). Qualities of thoughtful social studies classes: an empirical profile. *Journal of Curriculum Studies*, 22, 443-461.

O'Connor, K. A., Heafner, T., & Groce, E. (2007). "Advocating for social studies: Documenting the decline and doing something about it." *Social Education* 71(5), 255-260.

Parker, W. C. *Social Studies in Elementary Education*. (14th ed.) Saddle River, NJ: Pearson Education, Inc., 2011.

Rock, T., Heafner, T., O'Connor, K., Passe, J., Oldendorf, S., Good, A., & Byrd, S. (2006). One state closer to a national crisis: A report on elementary social studies in North Carolina schools. *Theory and Research in Social Education*, 34(4), 455-483.

Thornton, S. J. (1991). Teacher as curricular-instructional gatekeeper in social studies. In J. P. Shaver (Ed.), *Handbook of research on social studies teaching and learning* (pp. 237-248).

Van Fossen, P. J. (2005). Reading and Math Take So Much of the Time...' An Overview of Social Studies Instruction in Elementary Classrooms in Indiana. *Theory and Research in Social Education* 33 (3), 376-403.

Wiggins, G. and McTighe, J. (1998). *Understanding by Design*. Alexandria, VA: Association for Supervision and Curriculum Development.

Yeager, E. and Davis, Jr., O. L. (Eds.). (2005). *Wise practice in teaching social studies in the age of high stakes testing*. Greenwich, CT: Information Age Publishing.

Andrea Libresco, Ed.D., is a Professor in the School of Education at Hofstra University, in Long Island, NY.



The BOCES/SCOPE **Outdoor/Environmental Education Program**

*Specializing in customized science programs that meet
New York State Standards*

**Offering Summer camps, field trips, and the award-winning
Outdoor Learning Laboratory**

Call (631) 360-3652 for details