

## Old Dominion University ODU Digital Commons

---

Educational Foundations & Leadership Faculty  
Publications

Educational Foundations & Leadership

---

2011

# An Investigation of Higher-Order Thinking Skills in Smaller Learning Community Social Studies Classrooms

Christopher Fischer

Linda Bol  
*Old Dominion University*

Shana Pribesh  
*Old Dominion University*

Follow this and additional works at: [https://digitalcommons.odu.edu/efl\\_fac\\_pubs](https://digitalcommons.odu.edu/efl_fac_pubs)

 Part of the [Curriculum and Social Inquiry Commons](#), [Educational Assessment, Evaluation, and Research Commons](#), [Educational Methods Commons](#), and the [Secondary Education Commons](#)

---

### Repository Citation

Fischer, Christopher; Bol, Linda; and Pribesh, Shana, "An Investigation of Higher-Order Thinking Skills in Smaller Learning Community Social Studies Classrooms" (2011). *Educational Foundations & Leadership Faculty Publications*. 38.  
[https://digitalcommons.odu.edu/efl\\_fac\\_pubs/38](https://digitalcommons.odu.edu/efl_fac_pubs/38)

### Original Publication Citation

Fischer, C., Bol, L., & Pribesh, S. (2011). An investigation of higher-order thinking skills in smaller learning community social studies classrooms. *American Secondary Education*, 39(2), 5-26.

This Article is brought to you for free and open access by the Educational Foundations & Leadership at ODU Digital Commons. It has been accepted for inclusion in Educational Foundations & Leadership Faculty Publications by an authorized administrator of ODU Digital Commons. For more information, please contact [digitalcommons@odu.edu](mailto:digitalcommons@odu.edu).

# AN INVESTIGATION OF HIGHER-ORDER THINKING SKILLS IN SMALLER LEARNING COMMUNITY SOCIAL STUDIES CLASSROOMS

## AUTHORS

---

**CHRISTOPHER FISCHER, PH.D.**, is an Administrative Coordinator for the Department of Curriculum and Instruction in Virginia Beach City Public Schools.

**LINDA BOL, PH.D.**, is a Professor at Old Dominion University in Norfolk, Virginia.

**SHANA PRIBESH, PH.D.**, is an Assistant Professor at Old Dominion University in Norfolk, Virginia.

## ABSTRACT

---

*This study investigated the extent to which higher-order thinking skills are promoted in social studies classes in high schools that are implementing smaller learning communities (SLCs). Data collection in this mixed-methods study included classroom observations and in-depth interviews. Findings indicated that higher-order thinking was rarely promoted in SLC classes. Interview data suggests several factors affecting teaching for higher-order thinking in SLC social studies classrooms. These include: high stakes testing, pacing pressures, teachers' dispositions and training, and teacher autonomy.*

Researchers and educators espouse the viewpoint that thinking skills are an essential component of a modern education. Gough (1991) illustrated this consensus:

[P]erhaps most importantly in today's information age, thinking skills

are viewed as crucial for educated persons to cope with a rapidly changing world. Many educators believe that specific knowledge will not be as important to tomorrow's workers and citizens as the ability to learn and make sense of new information. (p.1).

Unfortunately, there is not a long established legacy emphasizing higher-order thinking skills in American secondary schools (Au, 2007; Bol & Nunnery, 2004; Cuban, 1984; Goodlad, 1984; Raudenbush, 1993).

The current paradigm of high-stakes testing may exacerbate the issue. The enactment of the *No Child Left Behind Act* has led many states to establish teacher accountability programs with standardized tests in order to secure federal funds for their local schools (US Department of Education, 2001). Critics of high-stakes accountability programs and the standardized tests on which they are based contend that they measure low-level knowledge products, which drives instruction and curriculum toward preparing students only for the demands of the test (Au, 2007; Bol, 2004; Kohn, 2000). Barksdale-Ladd and Thomas (2000) noted that 75% of a sample of teachers surveyed changed their instructional practices as a result of pressures from state testing. The report suggested that teachers shifted from instruction involving higher-order thinking skills, collaboration and in-depth understandings of content to instruction that was specifically designed toward material on state tests.

For many teachers, the assessments that comprise their state's teacher accountability systems are their instructional focus. Bol, Ross, Nunnery and Alberg (2002) contended that assessments directly influence classroom instructional practices and student learning. Furthermore, Kohn (2000) suggested that teachers may be reluctant to implement assessments targeting higher-order thinking because they do not reinforce the skills and precise content knowledge that is aligned with the multiple choice style end-of-year tests that are often part of state accountability programs. Au (2007) found this problem to be pervasive. He conducted a metasynthesis of 49 qualitative studies examining the relationship between high stakes testing and curriculum and instruction. He found that high stakes tests narrows the focus of curricular content to tested subjects, addressing content topics in isolation exclusively in the context in which they appear on state mandated tests. Au (2007, p. 263) stated, "In tandem with both content contraction and fragmentation of knowledge, pedagogy is also implicated, as teachers increasingly turn to teacher-centered instruction to cover the breadth of test-required information and procedures".

Additionally, school reform programs that address achievement and

other outcome measures with at risk populations are often stymied by the high-stakes testing paradigm. Desimone's (2002) research implicated high-stakes accountability systems as a barrier to the successful implementation of comprehensive school reform models. She noted the concerns of teachers, principals, and district level leadership regarding alignment between the goals and demands of reform programs and state assessment systems. High stakes testing is also an issue in the sustainability of reform programs already in place. In an article using longitudinal data examining the sustainability of comprehensive school reform models, Datnow (2005) reported that high-stakes accountability systems hinder long term implementation of reform, particularly in schools with histories of low academic performance: "In schools where state accountability demands were high, reform strategies were abandoned in favor of test preparation (p. 147)."

Small learning communities (SLCs), one type of school reform model, have demonstrated much promise for promoting academic and school success outcomes for urban districts. Over the past twenty years, a solid body of research has described the impressive benefits of small learning communities, including lower drop-out rates, fewer incidences of school violence, increased graduation rates, more students taking advanced placement classes, and higher grade point averages. Such research has demonstrated that SLCs are superior to large schools on many measures and equal to them on others, thus helping to close achievement gaps (ABT associates, 2002; Cotton, 2001; Raywid, 1999).

Although SLCs have been linked to several positive student outcomes, there is a dearth of research examining the ability of SLCs to promote high-order thinking skills. Onosko (1991) examined barriers to the promotion of higher-order thinking in secondary schools. His findings suggest notable overlap between the structure of schools he identified as promoting higher-order thinking and the structure of SLCs. Onosko (1991) identified six barriers to higher-order thinking including:

- teaching as knowledge transmission;
- broad superficial content coverage;
- teachers' low expectations of students;
- large class sizes;
- lack of teacher planning time;
- a culture of teacher isolation.

These factors that impede higher-order thinking seem also seem likely to hinder small learning communities from promoting school reform, espe-

cially in the context of a high stakes testing paradigm. For these reasons, an examination of the extent to which smaller learning community settings foster higher-order thinking is warranted.

## PURPOSE OF THE STUDY

The purpose of this study was to examine how the smaller learning community structure affects higher-order thinking skills in social studies classes by addressing the following research questions:

1. To what extent do teachers in smaller learning community social studies classes promote higher-order thinking in their classes?
2. How does the smaller learning community structure affect teachers' planning for instruction regarding higher-order thinking?

## METHOD

### PARTICIPANTS

A convenience sample of nine ninth grade social studies teachers participating in the small school reform underway in four high schools in an urban public school district in Southeastern Virginia were the focus of this study. Ninety-minute class periods were observed in these teachers' classrooms a total of 17 times. Additionally, four of these teachers participated in in-depth interviews about their instructional practices. All of the teachers interviewed taught ninth grade *Honors World History Part One* in ninety-minute blocks. Two were female and two were male. They ranged in experience from two years to fifteen years, and they had been faculty members at their respective schools from two to twelve years. Only one held a post-graduate degree in education, and none held post-graduate degrees in history. *World History Part One* is among the high school social studies courses requiring administration of a Standards of Learning (SOL) end-of-year test. The test is administered in the spring of each school year, and students' performance has implications for individual students' graduation prospects, as well as the accreditation status of the high schools they attend.

All four of the high schools participating in the study offered small learning community programs for their first-time ninth graders, and all were in the second year of their respective programs. Each of the participating high schools varied in demographic make-up and in the extent of their implementation of their smaller learning community initiatives. The classes observed for this were assigned by building level SLC coordinators. Table 1 describes the demographic information most relevant to the populations SLCs are designed to serve.

**Table 1. Relevant demographic information for high schools**

	Teachers in SLCs	Total Students Enrolled	Percentage of 9th graders in SLCs	Percentage of African American students	Percentage of Economically Disadvantaged Students	Percentage of Disabled Students
High School #1	4 of 12	1,409	31%	86%	70%	16%
High School #2	4 of 17	2,275	70%	58%	48%	11%
High School #3	10 of 20	1,988	73%	61%	43%	12%
High School #4	4 of 12	1,914	17%	64%	52%	14%

## MEASURES

**Observations.** This research operationalizes higher-order thinking using the conceptual framework articulated in Newmann's (1988) analysis of the qualities of thoughtful social studies classrooms and Newmann's (1990a) follow-up research aimed at examining how teachers promote thoughtfulness in classroom instruction. I developed observation guidelines and an interview protocol based on Newmann's framework.

The observation guidelines employed Newmann's *Six Minimal Criteria of Thoughtful Classrooms* (1990a) to analyze the extent of higher-order thinking promoted in smaller learning community social studies classrooms. Lessons were examined for each criterion using a five point Likert scale ranging from 0 "Not Observed" to 5 "Strong Application." Prior to data collection, a team of four observers discussed each criterion after one round of pilot observations to enhance consistency regarding student and teacher behaviors along the continuum that Newmann described in his (1990a) conceptual framework. Each criterion, along with descriptors of each end of a thoughtfulness scale is described in Table 2.

<b>Table 2. Six Minimal Criteria of Thoughtful Classrooms</b>		
Criterion	Scale	
	Limited Application	Strong Application
1. There was sustained examination of a few topics rather than superficial coverage of many.	There was superficial coverage of many topics.	There was sustained examination of a few topics.
2. The lesson displayed substantive coherence and continuity.	Lesson teaches material as unrelated fragments of knowledge, without pulling them together.	Lesson builds on relevant and accurate content knowledge that works toward the logical development and integration of ideas.
3. Students were given an appropriate amount of time to think, that is, to prepare responses to questions.	Inadequate time for reflection is provided encouraging quick surface responses to questions.	Students are given appropriate time for reflection to prepare thoughtful responses to questions
4. The teacher asked challenging questions and/or structured challenging tasks (given the ability level and preparation of the students).	Questions/tasks require the mere retrieval of prior knowledge that is of little mental challenge.	Questions or tasks demand analysis, interpretation, or manipulation of information they require the use of prior knowledge to gain new information.
5. The teacher was a model of thoughtfulness.	Teacher demonstrates interest in final absolute answers with little attention to complexities of problems and little acknowledgment of students' ideas.	Teacher demonstrates interest in students' ideas, in alternative approaches to problems; demonstrates how he or she thinks through problems and acknowledges the difficulty of gaining a definitive understanding of problematic topics.
6. Students offered explanations and reasons for their conclusions.	Students are encouraged to give little support or explanation for their responses to questions/tasks	Students are encouraged to provide detailed explanations and reasons to support their responses to questions/tasks.

**Interviews.** The interview guide featured 27 open-ended and semi-structured questions (See Appendix A). Teachers were also asked follow-up questions asking for further explanations, clarification and examples. Questions were aimed at gauging teachers' dispositions toward teaching for higher-order thinking, the extent to which they fostered the *Six Minimal Criteria* in their planning for and implementation of instruction and how the SLC structure mediated thoughtful instruction in their classrooms.

## PROCEDURE

**Observations.** As part of a larger program evaluation associated with a Department of Education smaller learning community grant, four researchers observed classrooms in four participating high schools. They completed observations geared toward outcome goals of the smaller learning communities grant, and while in social studies classes they completed observations using the *Six Minimal Criteria of Thoughtful Classrooms* observation guidelines. All observations were pre-scheduled during varied times of the school day and teachers knew in advance their classes would be observed.

**Interviews.** Semi-structured open-ended interviews were conducted using an interview guide. I selected those social studies teachers observed most often in their respective schools. Each interviewed teacher had been observed at least twice. Interviewed teachers were contacted either personally or via email to participate in the interview and assured of their anonymity to encourage candor in their responses. Interviews were conducted in each teacher's respective classroom at the end of the school year or during the summer break.

## FINDINGS

### **OBSERVATION DATA REGARDING HIGHER ORDER THINKING**

Classroom observation data did not reflect higher-order thinking. Table 4 illustrates the mean rating and standard deviation for each *Criterion of Thoughtful Classrooms* across the 17 observed lessons. A rating of zero indicates the described target behavior for a given criterion was not observed, a rating of one indicates a limited application while a rating of five indicates a strong application. The mean ratings for all the criteria ranged from 1.47 to 2.47. The fact that none of the criteria rose above the mid-points on the scale indicates that observed classes were not reflective of higher-order thinking as Newmann (1991a) defines it.



		Mean	SD
Criterion One	There was sustained examination of a few topics rather than superficial coverage of many.	2.06	1.35
Criterion Two	The lesson displayed substantive coherence and continuity.	2.47	1.42
Criterion Three	Students were given an appropriate amount of time to think, that is, to prepare responses to questions.	2.12	1.45
Criterion Four	The teacher asked challenging questions and/or structured challenging tasks (given the ability level and preparation of the students).	1.94	1.25
Criterion Five	The teacher was a model of thoughtfulness.	2.06	1.25
Criterion Six	Students offered explanations and reasons for their conclusions.	1.47	1.33

n=17 observed classes

The percentages of responses in each rating category appear in Table 5. The majority of observations were concentrated in the *not observed to limited application* ratings for all criteria. The most observed rating was *somewhat limited application* for criteria 4 and 5 related to teacher questioning and the teacher as a model of thoughtfulness. Strong applications of the teacher as a model of thoughtfulness and opportunities of student elaboration were not observed.

	Not Observed	Limited Application	Somewhat Limited Application	Neutral Application	Somewhat Strong Application	Strong Application
Criterion 1	5.9	29.4	41.2	11.8	0	11.8
Criterion 2	11.8	11.8	23.5	29.4	11.8	5.9
Criterion 3	11.8	17.6	47.1	5.9	5.9	11.8
Criterion 4	11.8	17.6	52.9	5.9	5.9	5.9
Criterion 5	11.8	11.8	52.9	11.8	5.9	5.9
Criterion 6	29.4	23.5	29.4	5.9	11.8	0

The following findings examine both data from observational field notes and from interviews. They are organized via Newmann's (1992) conceptual framework, according to each Criterion of Thoughtful Classrooms.

**Criterion One: There was sustained examination of a few topics rather than superficial coverage of many.**

Observed lessons were characterized by more superficial coverage of many topics rather than in-depth examination of a few. *The Strong Application* and *Somewhat Strong Application* categories were observed in fewer than 12% of the observations. The average rating for this criterion was 2.06 on a five point scale with a standard deviation of 1.35. Field notes from the observation guidelines support this assessment. For example, a class targeting the Mongolian empire was described as “covering the entire Mongolian Empire in one class session,” while a class exploring African kingdoms “covered all aspects of African Kingdoms for SOL prep.” A trend in lessons designed as preparation for the SOL tests was also apparent from an examination of the observation field notes. For example, one teacher noted “many topics, definitions of terms were discussed. Very broad coverage preparing for SOLs.”

In all four interviews, the scope and pace of content coverage emerged as a major theme. All of the teachers agreed that focused examination of historical topics is an ideal instructional goal, but they also admitted to feeling pressure to complete their curriculum before the end-of-year test. Data from the interviews suggested the emphasis on broad coverage was a consequence of data-based decision making mandated by district leadership. These efforts were aimed at channeling instructional decisions through evaluations of student performance on district-created unit and quarterly tests designed to mimic the SOL tests.

Three of the four teachers interviewed mentioned data driven decision making as one of the major components of their planning for instruction. In one school, the administration mandated that collaborative SLC teacher teams analyze test data and post the results of their analysis in teacher workrooms. One teacher complained, “data analysis is a major push, but there is no time or any plan for remediation.” The value placed on data from these tests translated into pacing pressure in planning for instruction. All four teachers interviewed stressed the importance of completing the curriculum in anticipation of the SOL test when asked whether finishing the curriculum was more important than insuring in-depth understanding of topics covered. One teacher answered, “finishing the curriculum is most important, because we’re a data driven school.”

Coverage as an instructional goal led to piecemeal surface treatment of the content. As one teacher put it, “we have to use a shotgun blast approach instead of a sniper rifle, because of the randomness of the curriculum.” Other teachers mentioned time constraints and pacing pressures as a major

concern when asked about the tension between in-depth treatment of content versus covering a large breadth of information. Another lamented the elimination of enrichment activities like fieldtrips to allow time to complete the curriculum.

Three teachers mentioned using class time during an SLC freshmen transition/freshmen leadership course to cover material from their social studies classes that they did not have time to cover during the scheduled class period. In fact, co-opting the freshmen transition/leadership class was the only indication teachers gave of any impact the SLC structure had on the depth versus breadth issue. The interview data indicated that the SLC structure does little to alter teachers' instructional priorities except to allow for more seat time with students to cover an overly broad curriculum.

**Criterion Two: The lesson displayed substantive coherence and continuity.**

The extent to which lessons logically and systematically integrated prior material was observed by the observation team, albeit only moderately. The average rating for this criterion was 2.47 with a standard deviation of 1.42. The *Not Observed*, *Limited Application* and *Somewhat Limited Application* categories were observed in more than 45% of the classes. *The Neutral Application* was observed in 29% of the observations. However, field notes more often characterized observed lessons as covering material in unrelated fragments without linking them together in any coherent way. For example, one lesson was described as "not well organized in a logical progression of topics with the relation among topics not addressed." More than one observed lesson made reference to prior relevant course content, but there was "no effort to tie them together." Test preparation may have also been a mitigating factor in assessing this criterion. Field notes described lessons as "test review," "teaching to the test," with "no explicit attempt to explain the connection of review concepts other than they were for the review."

Interview data somewhat contradicted observation data. Teachers described different strategies for establishing relevance across course content, indicating thoughtfulness as an instructional goal. One teacher reported using generalizations that apply to concepts across units as a means of stressing lesson continuity and coherence. For example, she would use the maxim, "money makes the world go round," to emphasize the relationship between successful civilizations and flourishing trade in teaching about various ancient civilizations. Another teacher talked about establishing "bridges" from one topic to the next. "I try to leave them with a cliff hanger, asking what could happen next."

Coherence and continuity of instruction were stated instructional goals. Nevertheless, the extent to which they were established was directly related

to the priorities teachers placed on content coverage. All teachers indicated the state published SOL curriculum guide as the most important influence on the scope and sequence of lessons while planning for instruction. They mentioned issues like student relevance or personal interest as incentives to elaborate on course topics; however, the parameters of such extra emphasis were ultimately determined by pacing pressures to complete the curriculum guide. One teacher remarked, "I'm sure to cover everything in the curriculum guide, but I try to spend extra time on things according to my students' reactions." Others reported using feedback provided from prior versions of the end of year test to weight content found in the curriculum guide for emphasis, "I go into depth based on how many questions about a topic might show up on the SOL test."

Thus, the logical transition from one course topic to the next was determined largely without regard for elements making up the class dynamic and aligned with expectations for content that would be tested at the end of the year. As a result, historical content was taught as discrete events with no real relationship among them, except that covered events all belonged to the same time period.

**Criterion Three: Students were given an appropriate amount of time to think, that is, to prepare responses to questions.**

Time for reflection and deep processing in student responses was not observed to a great extent. A *strong application* and *somewhat strong application* were observed in only 17% of the classes visited. The average rating was 2.12 with a standard deviation of 1.45. Field notes indicate that, in many classes, instead of measured and organized discourse, "students called out answers to the teachers' questions." Questioning was observed in most lessons, but the method was machine gun style questioning where questions were close-ended in nature in a "finish the teacher's statement style," not encouraging much think time. When follow up questions were asked, they "did not guide students to correct responses nor to reflect on the material."

During interviews, teachers indicated that they allocated response time for questions according to the context and purposes of the discussion. They distinguished between questioning for review and questioning for "deeper discussion." When asking questions to review prior covered material, described as "knowledge questions," teachers reported allowing little wait time, but all the teachers qualified their responses by noting that questions that were part of deeper discussions were afforded more wait time. One teacher said he allowed up to one minute of wait time, another had a "30 second rule," and yet another said he gave students enough time for the

“wheels to start turning.”

All teachers reported adjusting their expectations of wait time for individual students, allowing more time if students are engaged in thinking. Teachers preferred a directed approach in describing how they promoted reflective behavior in their classrooms. When asked how they encouraged their students to think before they respond to questions, teachers reported they explicitly directed such behavior. Teachers mentioned not calling on the first student to raise their hand and discouraging impulsive answers.

All teachers said that “deeper discussions” were a normal element of class discussions, but they provided contradictory responses related to wait time and asking challenging questions. In describing the difficulties with facilitating Socratic seminars, one teacher explained that ninth grade students are “easily confused by all the rules” and that the “over speak and excited speak of freshmen is not really a good fit.” Thus, the wait time teachers observed was dependent on the goals of instruction, was usually explicitly directed from teachers’ prompts and was not behavior students or teachers felt comfortable engaging in.

**Criterion Four: The teacher asked challenging questions/structured challenging tasks.**

The prevalence of teachers asking challenging questions was also limited. Challenging questions were rated as *Somewhat Limited Application* in 53% of the classes observed and the *Somewhat Strong* and *Strong Applications* was only apparent in 12% of the observed classes. The average rating for the fourth criterion was 1.94 with a standard deviation of 1.25. Field notes descriptions varied from, “all questions were objective style aimed at memorization of content,” to “activity required students to make choices about what content to include, how to order the information and which graphics to accompany information.”

Those teachers who were interviewed provided several descriptions of ways they wove challenging activities into instruction. When asked about the types of challenging activities they implemented, teachers mentioned a wide spectrum of tasks including: using document-based questions from primary sources, requiring students to create their own assessments of content material, and asking students to reason from different perspectives during class discussions. However, much like the questions regarding wait time and reflection, there appeared to be a disparity between teachers’ perceptions of the level of challenge inherent in their lessons and the level of challenge indicated in the observation data.

During interviews, teachers also indicated the prevalence of challenging instruction was mitigated by administrative mandates aimed at addressing

increasing numbers of failing students and pressures to finish the curriculum before the end-of-year tests. One teacher said he was “a slave to PowerPoint lectures because it is impossible to keep pace without lecturing,” especially in the face of pressure from his administrators that his classes complete the curriculum with time to spare so they could review for the SOL test.

Teachers also described the difficulties of differentiating instruction for varied students’ abilities, students’ lack of confidence and motivation, and the trouble in making the transition from middle school to high school as obstacles to implementing challenging instruction. One teacher said, “my students are resistant to challenging questions. They want immediate answers and immediate results.” Teacher experience with the demands of implementing challenging instruction was another issue discussed. Teachers described their uncertainty for when and to what extent to provide support to students struggling with challenging tasks. “I have to walk a fine line between letting them do nothing and doing it all for them.”

**Criterion Five: The Teacher was a model of thoughtfulness.**

Modeled thoughtfulness was not often observed. Over 76% of the classes observed were rated from *Not Observed* through *Somewhat Limited Application* for modeled thoughtfulness. The average rating for this criterion was 2.06 with a standard deviation of 1.25. In classes where modeled thoughtfulness was observed it was through spontaneous responses to questions students asked about class activities and not systematic or planned as part of instruction. For example, one teacher described verbalizing her thinking when asked how the Five Pillars of Islam were similar to the Ten Commandments.

During interviews, teachers reported modeling thoughtfulness mainly in the context of practicing test taking strategies or analyzing primary sources for the purpose of answering questions on the SOL test. When asked how they were able to demonstrate thinking skills to their students, all of the teachers interviewed mentioned working through multiple choice questions with their classes to identify and apply test taking strategies. One teacher said, “I mimic how they should be thinking in tackling a question and stress the importance of staying on task.”

All of the teachers were excited about the promise of technologies like interactive white boards that have made it easier to demonstrate test taking strategies to large groups of students. “The Smartboard allows me to show the kids visually how to eliminate incorrect choices while I think aloud.” Other teachers also discussed “thinking aloud” as they analyzed primary source documents with their classes. “Sometimes I purposely make mistakes jumping to false conclusions until my kids catch on.”

**Criterion Six: Students offered explanations and reasons for their conclusions.**

Students were not often observed offering detailed explanations and reasons to support their answers to questions. None of the classes observed were rated as having a *Strong Application* for this criterion and only 12% were observed as *Somewhat Limited Application*. The mean rating for this criterion was 1.47 and the standard deviation was 1.33. The deficit in students' elaborated responses can be explained by teachers' questions and follow up questions requiring detailed responses. The low frequency with which challenging questions were posed is consistent with this explanation. Field notes also support this finding. For example one observer noted, "The teacher did not follow up or ask for much elaboration; he had fill-in the blank questions while students volunteered one word answers." It appears students were not given opportunities to substantiate their conclusions or articulate their reasoning regarding content topics.

In the interview analysis, class discussions emerged as the primary strategy teachers offered as a means to elicit elaboration from their students. When asked how they encouraged their students to provide more elaborate explanations to questions, all of the teachers mentioned asking follow up questions during class discussions. One teacher quipped, "I ask 'why' very often, since their first answers are infrequently satisfactory." Some were more systematic in their approaches than others. One teacher talked about how he used test data to identify which content objectives need to be emphasized, "to make sure we spend extra time discussing that information." He also admitted that there were not many opportunities for such discussions because of time constraints to cover the broad curriculum. Few of the teachers' comments specifically mentioned other types of class activities that would require students to explain their thinking. Additionally, short of following up initial questions with "why", none of the teachers offered any descriptions of questioning strategies that might elicit student explanations.

**DISPOSITIONS TOWARD THOUGHTFULNESS**

Newmann (1992) posited that teaching for thoughtfulness is a composite of knowledge, skills, and dispositions. Thus, identifying teachers' beliefs is an important part of examining the extent to which thinking skills are promoted in SLC social studies classrooms. Those beliefs included the value of instruction for thoughtfulness and the extent to which promoting thoughtfulness drives instruction, as well as teachers' beliefs about their efficacy to impact thoughtfulness. Consistent with the observation data, teachers indicated that the end-of-year Standards of Learning tests framed the goals that

drove instruction. One teacher explained that instructional objectives were, “driven by essential questions, which are part of a scope and sequence developed with student performance on the SOL test in mind.”

The other teachers also confirmed the primacy of the SOL test in developing lesson objectives. When describing how the content and skills objectives prescribed in the state issued SOL curriculum guide her lesson planning one teacher said, “the challenge in planning is matching up appropriate content and skill objectives. The content is too great. I start slow at the beginning emphasizing skills and then go rapid fire with the content as the school year goes on.”

Teachers’ responses in reference to fostering students’ higher-order thinking indicated little consideration of thoughtfulness as a valuable educational goal. Teachers believed they had an influence on their students’ thinking skills, but elaborated by emphasizing details regarding discipline and social-emotional connections, suggesting a misconception of what thoughtfulness as an instructional goal entails. One teacher commented, “you get back what you give.”

Clearly, the end-of-year standardized test loomed large in teachers’ minds when they were forging a philosophical approach to instruction and setting subsequent overarching class goals. These responses suggest a misconception regarding instructional goals that a thoughtful classroom implies. Teachers seemed to equate high performance on standardized tests as a byproduct of challenging instruction. The disposition toward structuring thoughtful classrooms was not evident in any of the teachers’ responses. Additionally, there was no evidence from the interview data that suggested the SLC structure mitigated pressures to “teach to the test” when teachers developed instructional goals for their lessons.

## DISCUSSION

With regard to the first research question, the results suggest that classroom instruction within these SLC schools rarely reflects higher-order instruction. Neither the observation nor interview findings seem to support the notion that thoughtful instruction was fostered. Mean ratings from the data on each observation item were below 2.5, beneath the standard considered adequate for promoting higher-order thinking established by Newmann (1991).

One theme that continually emerged from the qualitative and quantitative data was the contradiction between thoughtful instruction and teaching with the end of year test in mind. Teachers frequently said the right things when asked about thoughtful instruction in their classrooms, but the quantitative data suggests that thoughtful instruction and thoughtful behavior were



rarely observed. Bol, Ross, Nunnery and Alberg (2002) identified pressure from high stakes accountability programs as substantial influences on the allocation of resources and instructional time in school districts with teacher accountability plans in place. It is likely that the pacing pressures frequently mentioned in the qualitative data are at least partially responsible for the deficit of thoughtful instruction observed.

Geersten (2003) described the potential influence of accountability and assessment systems on instruction. "If standardized testing becomes a larger part of teaching assessment, teachers will teach to the test just as students direct their learning to the tests" (p. 15)." Bol (2004) elaborated this point, suggesting that teachers designed their assessments to mimic the close-ended assessments common among end of year high stakes tests. Darling-Hammond (1994), Kohn (2000) and Shepard (2000) all warned that an overemphasis on high stakes testing would impede instruction designed to promote critical thinking and problem solving. Confirming these warnings, Bol and Nunnery (2004) identified authentic assessments and student centered classroom orientations as among the first instructional elements that would be omitted in response to high stakes testing demands.

The qualitative data is replete with references to pacing pressure from administrators to implement instructional practices that are clearly at odds with Newmann's (1991) construct describing thoughtful classrooms, and many statements confirmed Desimone's (2002) and Datnow's (2005) admonitions of district pressures regarding high stakes testing and the success and sustainability of school reform models. The qualitative data also describes organizational structures in place in the schools studied (such as data analysis collaboration teams) that subvert thoughtful instruction in favor of assurance that students are prepared for the end-of-course test. Thus, in the SLC school in this study, it is highly likely that teachers designed instruction compelled by performance on end-of-year tests.

Another factor affecting the levels of thoughtfulness observed may be related to teachers' understanding of and training with instructional strategies and assessments that support higher-order thinking. There are several examples of contradictions between the answers teachers provided regarding their teaching and planning for higher-order thinking and the extent of higher-order thinking instruction actually observed. For example, in describing the context in which they provided appropriate wait time during class discussions, all teachers indicated that "deeper discussions" were a normal element of their class discussions, but substantial wait time was not frequently recorded in the observation data. Although teachers acknowledged the value of instruction for thoughtfulness, their responses indicated at least an

incomplete understanding of what thoughtful instruction entails. After completing one interview, a teacher remarked that the interview questions gave him pause to reflect on his teaching, "after thinking about these questions, I realize there is a lot more I could be doing to get my kids thinking."

One possible explanation for the disparity could be the amount of training and experience teachers have acquired. There is an extensive line of research connecting years of experience teaching, as well as training in assessment and measurement, to the use of alternative assessments that tap higher-order thinking skills (Hargreaves, Earl, & Schmidt, 2002; Mertler, 2000; Stiggins, 1992). Similarly, Bol (1998) showed that experienced teachers used alternative assessments more frequently than inexperienced teachers. This trend, when considered with the inconsistencies between the qualitative and observation data in this study, may indicate lack of training as the source of confusion regarding the purpose of assessment and the knowledge and reasoning (higher-order) skills targeted in particular assessments.

Bol and Strage (1996) discovered a misalignment between teachers' instructional goals and the demands of their classroom assessments. In their study, many teachers said that teaching students to problem-solve and think critically was an important instructional goal, yet their assessments were very low level in terms of cognitive demand. Therefore training may be key to promoting thoughtful instruction. Developers and researchers of effective thinking skills programs cite teacher training as a critical factor in program success. The majority of programs have a strong staff development component, and developers consider this training to be as important as the program content in bringing about learning gains (Cotton, 1991).

With regard to the second research question, the quantitative and qualitative data also suggest that the SLC structure has little impact on teachers' planning and implementation of thoughtful instruction. The quantitative data implicitly suggest this conclusion because instruction and behavior were not routinely observed. The qualitative data did not give any indication that the SLC structure influenced teacher planning for instruction except that it afforded teachers extra time to cover material they did not have time to get to during their regularly scheduled classes. Follow up questions revealed that instruction was not discussed during collaborative meetings of SLC teachers, but issues mandated by the administration such as student discipline, assessment data of quarterly and semester tests, and other logistical issues were the topics most often discussed in meetings.

Despite the potential of SLCs to address several of Onosko's (1991) barriers to higher-order thinking, the SLC reform as implemented in this study did not effectively address teaching as knowledge transmission, broad

superficial coverage, and teacher's low expectations for students. The presence of these barriers is described in the qualitative data and implied by the quantitative data. The absence of substantive change resulting from the implementation of the SLC structure in these high schools should not be surprising. Several issues described by Onosko (1991) as barriers to higher order thinking are evident throughout the data and can be traced back to pressures related to the high stakes testing paradigm. These results confirm Au's (2007) meta-synthesis findings regarding the detrimental effects of high stakes testing on curriculum and instruction and Desimone's (2002) assertion that high stakes accountability systems inhibit successful implementation of school reform models.

A limitation of these results is that they treat the SLC reform efforts in the district in this study as monolithic. Although part of the same grant, each of the four schools studied had its own SLC program in place, each at different levels of implementation. Even though all of the programs were required to meet specific criteria in order to be in compliance with the grant, schools initiated the SLC structures at different times and were at different levels of implementation. Results may have been more promising in schools with higher levels of SLC implementation for more extended periods of time.

Another limitation was that there was no control group or pre-implementation measures for comparison. The thoughtful instruction and behavior that was observed may have been more frequent and extensive than it was at non-SLC schools with similar populations. Given the paucity of the examples of thoughtful instruction observed, however, this would be an unlikely and disturbing trend.

This study was also limited by the time of the school year during which observations took place, which may have affected the extent to which the breadth and depth of content topics were emphasized. Therefore, it is critical to conduct further studies of SLC reforms with policies in place that specifically address issues of content coverage and teachers' expectations of their students. Perhaps studies involving SLC schools serving similar populations that are free from the pressures of high stakes accountability systems would provide a more accurate assessment of the promise of SLCs to promote higher-order thinking. Other areas of future study would include investigation into the structures that encourage collaboration among SLC teachers and to what extent instruction is emphasized in such collaborations. Additionally, more research must be done on the relationship between SLC teacher autonomy and challenging instruction.

## REFERENCES

- Au, W. (2007). High stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258-267.
- Barksdale-Ladd, M. A., & Thomas, K. (2000). What's at stake in high stakes testing: Teachers and parents speak out. *Journal of Teacher Education*, 51(5), 384-397.
- Bol, L., (2004). Teachers' assessment practices in a high-stakes testing environment. *Teacher Education and Practice*, 17(2), 162-181.
- Bol, L., Stephenson, P. L., Nunnery, J. A., & O'Connell, A. A. (1998). Influence of experience, grade level and subject area on teachers' assessment practices. *Journal of Educational Research*, 91, 323-330.
- Bol, L., & Nunnery, J. A. (2004). The impact of high-stakes testing on restructuring efforts in schools serving at-risk students. In G. Taylor (Ed.), *The impact of high stakes testing on the academic futures of non mainstream students* (pp. 101-117). Lewiston, NY: Edwin Mellon Press.
- Bol, L., Ross, S. M., Nunnery, J. A., & Alberg, M. (2002). A comparison of teachers' assessment practices in school restructuring models by year of implementation. *Journal of Education for Students Placed at Risk*, 7(4), 407-423.
- Bol, L., & Strage, A. (1996). The contradiction between teachers' in instructional goals and their assessment practices in high school biology courses. *Science Education*, 80(1), 145-163.
- Cuban, L. (1984). *How teachers taught: Constancy and change in American classrooms: 1890-1980*. New York, NY: Longman.
- Cotton, K. (2001). *New small learning communities: Findings from recent literature*. Portland, OR: Northwest Regional Educational Library.
- Darling-Hammond, L. (1994). Setting standards for students: The case for authentic assessment. *Educational Forum*, 59, 14-21.
- Data on Student Preparation, College. (2007). *Peer Review*, 9(1), 24-45.
- Datnow, A. (2005). The sustainability of comprehensive school reform models in changing district and state contexts. *Educational Administrative Quarterly*, 41(1), 121-153.
- Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research*, 72(3), 433-479.
- Geersten, H. R. (2003). Rethinking thinking about higher-level thinking. *Teaching Sociology*, 31, 1-19.
- Goodlad, J.I. (1984). *A place called school: Prospects for the future*. New York, NY: McGraw-Hill.
- Gough, D. (1991). *Thinking about thinking*. Alexandria, VA: National Association of Elementary School Principals.
- Hargreaves, A., Earl, L., & Schmidt, M. (2002). Perspectives on alternative reform. *American Educational Research Journal*, 39, 69-95.
- Kohn, A. (2000). Burnt at the high stakes. *Journal of Teacher Education* 51(4), 315-327.
- Mertler, C. A. (2000). Teacher centered fallacies of classroom assessment validity and reliability. *Midwestern Educational Researcher*, 13, 29-35.
- Newmann, F. M. (1988). *Higher order thinking in high school social studies: An analysis of classrooms, teachers, students and leadership*. (Report No. SO 019-569). Madison, WI: National Center on Effective Secondary Schools.

- Newmann, F. M. (1990). Higher order thinking in social studies: A rationale for the assessment of classroom thoughtfulness. *Journal of Curriculum Studies*, 22(1), 41-56.
- Newmann, F. M. (1991). Promoting higher order thinking in social studies: An overview of a study of 16 high school departments. *Theory and Research in Social Education*, 19(4), 324-340.
- Onosko, J. (1991). Barriers to the promotion of higher-order thinking in social studies. *Journal of Curriculum Studies*, 19(4), 341-366.
- Page, L., Layzer, C., Schimmenti, J., Bernstein, L., & Horst, L. (2002). *National evaluation of smaller learning communities: Literature review*. Cambridge, MS: ABT Associates.
- Quellmalz, E. D. (1985). Needed: Better methods for testing higher order thinking skills. *Educational Leadership*, 43(2), 29-35.
- Raudenbush, S., Rowan, B., & Cheong, Y. F. (1993). Higher order instructional goals in secondary schools: Class teacher and school influences. *American Educational Research Journal*, 30(3), 523-553.
- Raywid, M.A. (1999) Current Literature on Small Schools. *ERIC Digest*. Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Stiggins, R. J. (1992). *In teacher's hands: Investigating the practices of classroom assessment*. Albany, NY: State University of New York Press.
- U. S. Department of Education. (2001). *No child left behind*. Retrieved from <http://www.ed.gov/news/pressreleases/2003/02/02032003b.html>

## APPENDIX A

### ***Interview Guide Thoughtfulness in Smaller Learning Community Social Studies Classrooms***

#### Dispositions of Thoughtfulness

1. How do instructional goals guide your lesson planning?
2. What do you most want students to achieve as a result of taking your class?
3. How much influence do you feel you have over how your students thinking?

#### **Criterion 1. There was sustained examination of a few topics rather than superficial coverage of many.**

4. Which is most important in your lessons, finishing the curriculum or insuring students' in-depth understandings of the topics you cover?
5. How does the tension between going in-depth and covering a large breadth of topics affect your lesson planning? Instruction?
6. How does the small learning community structure affect this tension?

#### **Criterion 2. The lesson displayed substantive coherence and continuity.**

7. How do you determine the content, scope and sequence of your lessons while planning for instruction?

8. Besides the district's curriculum guide, what guides your lesson planning? What other frameworks guide your lesson planning?

**Transitions**

9. How do you go about maintaining continuity across lessons?
10. How do you tie together varied ideas and concepts across lessons?

**Criterion 3. Students were given an appropriate amount of time to think, that is, to prepare responses to questions.**

11. What are some ways you encourage your students to be reflective about the content you cover?
12. How much time do you typically allow students to answer questions during discussions of course content? Individuals? Groups? Socratic?
13. In what ways do you encourage your students to think before they respond to questions asked during class?

**Criterion 4. The teacher asked challenging questions and/or structured challenging tasks (given the ability level and preparation of the students).**

14. What sorts of activities do you plan that you would consider challenging for your students?
15. What sorts of thinking skills do you think these activities/questions require of them?
16. How do you offer them support in applying these skills?
17. How do your students respond to the more intellectually challenging activities or requirements?
18. In which contexts are your students most responsive to intellectually challenging tasks?

**Criterion 5. The teacher was a model of thoughtfulness.**

19. How do you go about demonstrating to your students the skills required to complete the activities you plan?
20. How often do you present historical content as problems to be solved?
21. How do you teach students to think about taking multiple perspectives or interpretations of historical content?
22. How do you handle the issue of truth in historical knowledge? Does it ever come up?
23. How do you teach students to think about different perspectives when thinking about historical content?
24. How do you ask students to think like the historical characters they study?

**Criterion 6. Students offered explanations and reasons for their conclusions.**

25. How do you encourage students to provide more elaborate explanations to questions they answer during discussions?
26. How do you insure that students understand historical concepts beyond the answers to objective questions they provide?
27. To what extent do the activities and discussions in your class provide opportunities for students to explain their thinking regarding the content delivered?

Copyright of American Secondary Education is the property of American Secondary Education and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.