

Social Studies, Social Justice: W(h)ither the Social Studies in High-Stakes Testing?

By Wayne Au

Introduction

High-stakes, standardized tests have become ubiquitous in public education in the United States. Teachers across the country are feeling the intensified pressures from high-stakes testing policies and are responding to these pressures by teaching to the tests in varying ways (Renter et al., 2006). Given the hegemony of high-stakes testing in schools today, this article seeks to explore the question: *W(h)ither the social studies in high-stakes testing?*

Drawing on the available body of empirical research, I will argue that social studies teachers are feeling the pressures of high-stakes testing, and that these pressures are causing social studies teachers to alter their classroom practices and curriculum. Further, I will posit that the social studies represent a special case in relation to other subjects because the changes to pedagogy and content are variable, and the amount or significance of these changes depends on specific factors such as test design or whether or not individual sanctions are tied to student performance on the tests. Finally, I argue that, because of the consistent variability connected to social studies teaching in relation to high-stakes tests, social studies

Wayne Au is an assistant professor of secondary education in the College of Education at California State University, Fullerton, Fullerton, California.

education, in many instances, is positioned to provide an education that challenges the hegemonic norms of high-stakes testing generally as part of a broader need to teach for social justice in today's schools.

A Brief Social History of High-Stakes Testing

A test is considered high-stakes when its results are used to make important decisions that immediately affect students, teachers, administrators, communities, schools, and districts (Madaus, 1988). These decisions may include whether or not a student graduates high school or is promoted from one grade to another, and they may also include the salary scales and tenure status of teachers and principals (Orfield & Wald, 2000). As part of the “accountability” movement, stakes are also deemed high because the results of tests, as well as the ranking and categorization of schools, teachers, and children that extend from those results, are reported to the public (McNeil, 2000), thus putting the reputation of states, districts, schools, principals, teachers, and students up to public scrutiny and judgment.

The publication of *A Nation At Risk* (National Commission on Excellence in Education, 1983) ushered in the contemporary standards and high-stakes testing movement (Nichols, Glass, & Berliner, 2005). This report was a product of the Reagan administration and, as such, it attacked public education for failing to keep up with other foreign powers within the context of Cold War geo-politics. Even though the data and analysis used in *A Nation At Risk* was later determined to be empirically false (Berliner & Biddle, 1995), nonetheless, this report “galvanized the fledgling accountability movement, transforming it into a national project with purported national security implications” (Sacks, 1999, p. 77). Within a year of *A Nation At Risk*'s publication, and following suit with many of the report's recommendations, fifty-four state level commissions on education were created, and twenty-six states raised graduation requirements. Within three years of its publication, thirty-five states had instituted comprehensive state education reforms that revolved around testing and increased course loads for students (Kornhaber & Orfield, 2001). Thus, the trajectory of education reforms into the 1990s was set, where forty-three states had statewide assessments for k-5 by 1994, and by the year 2000 every state but Iowa administered a state mandated test (Jones, Jones, & Hargrove, 2003).

The movement toward high-stakes testing continued through then Republican Vice President George H. Bush's campaign for the presidency where, as the self-proclaimed “education president,” he endorsed minimum competency testing for grade promotion and graduation. As President, G.H. Bush carried this forward into his America 2000 plan—focusing on testing and establishing “world class standards” in schools. Democrats Clinton and Gore later committed themselves to following through on the goals established by the America 2000 plan, maintaining the rhetoric of the necessity of “tough standards” in our schools and pursuing a national examination system to meet those standards. Al Gore, in his 2000 presi-

dential campaign, subsequently called for all states to establish high-stakes high school graduation tests, and within the first week of taking office, President George W. Bush pushed for federal Elementary and Secondary Education Act (ESEA) Title I funding to be tied to student test scores (Kornhaber & Orfield, 2001).

In 2002, the federal government reauthorized the ESEA, now renamed as the *No Child Left Behind Act*, or NCLB (U.S. Department of Education, 2002). NCLB is built around high-stakes testing, mandating that, by 2006, all students be tested in reading and math in grades 3-8 and once in high school. By 2008, NCLB also mandates that students will be tested at least once at the elementary, middle, and high school levels in science. If schools do not show consistent improvement on these tests, meeting what is called “Adequate Yearly Progress” (AYP) in subgroups related to race, economic class, special education, and English language proficiency, among others, they face sanctions such as a loss of federal funding or the diversion of federal funding to pay for private tutoring, transportation costs, and other “supplemental services.” Under NCLB, all students in all subgroups are also expected to score at 100% proficiency by the year 2014 or schools will face the above-mentioned sanctions.

Based on the historical record, Kornhaber and Orfield (2001) conclude that:

[T]here has been an unbroken line of rhetoric, extending across six presidential terms, fostering high-stakes testing. For almost two decades, all the national leaders of both parties have embraced the theory that our schools have deteriorated and that they can be saved by high-stakes tests. So have the state leaders in almost all of our states. (p. 4)

Counting President G.W. Bush’s re-election to office, the bipartisan rhetoric and policy focused on high-stakes testing now stretches across seven presidential terms, where, by 2004, in addition to the federally mandated tests, 24 states required high-school exit exams, with all but six withholding diplomas based on the test scores (Emery & Ohanian, 2004). Thus, testing has become firmly entrenched as the policy tool, bar none, for federal enforcement of educational reforms.

High-Stakes Testing and Classroom Control

A key issue for education researchers to examine is just what effects high-stakes testing has had on both curricular content and teacher pedagogy. Based on the research, we can conclude that high-stakes tests do generally leverage some control over the content of the curriculum. One of the most prevalent and consistent findings is that high-stakes testing narrows the instructional curriculum and aligns it to the tests. This happens because, to varying degrees, teachers feel pressured to shape content norms to match that of the tests. Several major surveys of teachers and principals finds decreases in the teaching of non-tested subjects and corollary increases in the teaching of tested subjects (see, e.g., Pedulla et al., 2003; Renter et al., 2006; von Zastrow, 2004). For instance, one nationwide survey found that

71% of the districts reported cutting at least one subject to increase time spent on reading and math as a direct response to the high-stakes testing mandated under NCLB (Renter et al., 2006). In addition to national surveys, quantitative and qualitative studies in many states (see, e.g., Gayler, 2005; Sloan, 2005; Wright & Choi, 2005) are finding the same thing: When punitive consequences are attached to test scores, teachers do indeed match their pedagogy and content to the test norms.¹

Even if teaching to the test means the alignment of content knowledge with the tests, it does not necessarily mean that pedagogy is affected. Again, the research does suggest that, in many cases, teachers' instructional practices have been altered by the pressures associated with high-stakes testing. In the classroom this translates into teachers preparing students for tests with pedagogies that focus on rote memorization and lower-order thinking as the tests themselves are usually structured to assess breadth of often shallow, fragmented bits of knowledge (see, e.g., Gayler, 2005; McNeil, 2000; Stecher & Barron, 2001; Taylor, Shepard, Kinner, & Rosenthal, 2001). In these cases, pedagogy is reduced to figuring out how to dispense "packaged fragments of information sent from an upper level of the bureaucracy" (McNeil, 2000, p. 5). This pedagogical trend towards "multiple choice teaching" (Smith, M. L., 1991, p. 10) also manifests in increased time doing test drills and practicing for the types of information, questions, and test-taking skills that the tests require (Hillocks Jr., 2002; Perreault, 2000; Teachers Network, 2007). This is becoming particularly true in regards to reading programs where many districts are becoming "more prescriptive about how and what teachers should teach" (Renter et al., 2006, p. 99).

Teachers, however, are highly cognizant that their pedagogy is being challenged by the pedagogic norms established by the high-stakes testing. Another related finding of the research is that, in response to this pedagogic control, teachers feel that they are teaching in ways that are contradictory to best practice (see, e.g., Agee, 2004; Brimijoin, 2005; Pedulla et al., 2003). For instance, in their nationwide survey, Abrams, Pedulla, and Madaus (2003) found that 76% of the teachers in states with "high" stakes tests and 63% of the teachers in their study from states with "low" stakes testing reported that their state testing programs were contributing to unsound educational practices. Given the above evidence, it does seem clear that, generally, high-stakes tests exert some level of control over teachers' instructional practice, and that this control often times contradicts what many teachers feel is good pedagogy.

Wither the Social Studies in High-Stakes Testing?

Given the state of teaching under high-stakes testing regimes, then, how do we situate social studies? The evidence supports the general claim that high-stakes testing influences both curricular content and teachers' instructional practice, and that this influence is largely negative: non-tested subjects and content are being reduced and teachers feel forced to used pedagogies that they feel are antithetical to

good teaching. As Grant (2006) notes, research on the specific impact of high-stakes testing on social studies teaching has been limited, with a small body of existing empirical research. But the literature base is growing, and based on the evidence, it is clear that the teaching of social studies is affected by high-stakes testing, although the nature of these effects varies depending on local contexts and conditions.

The Disappearing Subject

Because NCLB places so much emphasis on reading and math testing, many schools and districts are reducing the amount of social studies education they offer. For instance, in a survey of almost 1,000 principals in the states of Illinois, Maryland, New Mexico, and New York, 47% of the respondents who oversaw K-5 schools with “high minority” populations reported decreases in the teaching of social studies (von Zastrow, 2004). In a study of Colorado teachers, 82% of 161 teachers interviewed by phone reported reducing social studies teaching in response to high-stakes tests there (Taylor et al., 2001). Other studies have also reported cuts in social studies instruction in North Carolina (Groves, 2002) and Arizona (Wright & Choi, 2005).

Perhaps most disturbing for social studies education is a study done by the Center on Educational Policy (CEP) which received survey responses from all fifty states, including 299 school districts (Renter et al., 2006). The CEP found that 33% of the districts in their study reported reducing social studies in response to high-stakes testing. One teacher in the study states that, “[NCLB] has torn apart our social studies curriculum” (p. 10). The study goes on to highlight several schools, including North Tahoe Middle School in California, where students who are not meeting standard are taking up to three periods of reading and two periods of math a day. This has meant that some students are not taking social studies at all.

This finding should come as no surprise. The pressures associated with high-stakes testing results can be severe, and as schools scramble to improve their reading and math scores, it makes sense that administrators would make the pragmatic decision to increase teaching in those areas. Indeed, as Savage (2003) notes, rather than wage a concerted fight against high-stakes testing generally, many prominent social studies educators have called for social studies to be included on all high-stakes tests, thus guaranteeing their survival as part of the whole-school curriculum. While there is research to support this position (see, Stecher & Barron, 2001), this proposition is problematic for many reasons: It explicitly endorses a test-or-cut logic for all subjects; it blindly accepts the content norms of testing; it assumes that high-stakes, standardized test scores are valid, reliable measures of learning; and it glosses over the very serious and real relationship that high-stakes, standardized tests have with contributing to race and class inequality in education (Au, 2008; Darder & Torres, 2004; Popham, 2001).

Trivializing History and Pedagogy

In several studies, social studies teachers operating in states whose social

studies tests focused on the rote memorization of historical facts both added and cut curriculum content to align with the information on the tests (Fickel, 2006; Salinas, 2006; Segall, 2006; Smith, A. M., 2006; van Hover, 2006; van Hover & Heinecke, 2005; Vogler, 2005). A two year study of a high school social studies department in Kentucky found that teachers there adjusted their teaching to cover more content, and that their assessment practices changed to meet the tests too (Fickel, 2006). Vogler (2005) obtained similar findings in a survey of 107 U.S. History teachers across 55 districts in Mississippi, where teachers reportedly narrowed their content to match the tests and adopted more teacher centered instruction as a more efficient means of delivering the content breadth demanded by the tests. Similarly, in Virginia, social studies teachers not only adjusted content to meet the test requirements, but also resorted to fast-paced lectures to make sure everything was covered in time (Smith, A. M., 2006; van Hover & Heinecke, 2005). Teachers in Virginia have also been found to adjust their classroom assessments to match that of the high-stakes tests, where some do weekly fast-recall quizzes and others make sure to use multiple-choice questions on end-of-unit exams (van Hover, 2006).

Research has, in addition, found that high-stakes social studies tests that focus on the rote memorization of facts also promote a vision of social studies education as the collection of historical facts (Bigelow, 1999; Pahl, 2003; van Hover, 2006). Indeed, in order to maintain “reliability,” many history tests throw out high-level, critical thinking questions in favor of lower-order memorization of facts, thus resulting in the trivialization of historical knowledge (Pahl, 2003). This trivialization is further reinforced by budgetary and capacity issues faced by test designers, since it is cheaper, easier, and more efficient to construct, administer, analyze, and report results from a test that consists of close-ended, multiple choice questions (Pahl, 2003; Toch, 2006). As some social studies scholars have noted, tests that focus on collections of facts also promote lower order thinking for students (Grant, 2001; Grant et al., 2002).

The above findings should be disturbing to social studies teachers and teacher educators. It appears that, when high-stakes social studies tests consist mainly of multiple choice questions and dislocated fact memorization, social studies teachers feel compelled to align their content, instruction, and assessment to the test-defined norms. While this does not mean that these teachers changed their content, instruction, and assessment *en toto*, it does demonstrate that these tests do change the social studies curriculum and do cause teachers to at least reduce the amount of student-centered instruction and increase the amount of teacher-centered instruction in their classrooms. Indeed, as some of the research finds, teachers admit to using less class time for inquiry learning and critical analyses due to the content demands of the tests (Fickel, 2006; Smith, A. M., 2006), thus leading them to use more teacher-centered instruction, instruction which teachers recognize is incongruent with “wise” social studies teaching (van Hover & Heinecke, 2005). Thus, social studies teachers faced with fact-based social studies tests seem to respond to high-stakes tests in the same ways as other teachers: they teach to the test.

Test Construction Matters

The New York State global history and geography exam is one of the most studied examples (see, Bolgatz, 2006; Gerwin, 2004; Gradwell, 2006; Grant, 2001, 2003, 2005; Grant et al., 2002; Libresco, 2005). This high-stakes test is part of the Regent's exams, and it is one that mixes multiple-choice questions associated with rote memorization along with a "document-based question" (DBQ). The DBQ requires students to review source material and answer, in essay form, a question related to the historical material (Grant, 2003). Although imperfect, the DBQ gives social studies teachers much more leeway and seemingly allows for many of the curricular/pedagogic constraints associated with high-stakes and standardized tests to be lifted. Several case studies profiling singular instances of exemplary or "wise" social studies teaching in relation to New York's test highlight this occurrence. Bolgatz (2006) tells the story of one teacher that makes use of DBQ to explore primary documents while studying race and racism in a fourth grade U.S. History course. Libresco (2005) relates how another fourth grade teacher maintained "wise" social studies practice through her attention to DBQs in her classes. Gradwell (2006) tells a similar story of an eighth grade history teacher and her use of DBQ to encourage historical thinking amongst her students. Grant (2005) uses a case of a middle/high school social studies teacher to demonstrate that "ambitious teaching" exists even in a high-stakes testing environment like New York. Other studies of New York's tests unearth similar findings as well. Gerwin's (2004) study of pre-service teachers finds that teachers-in-training there did not feel the need to align their practice to the state exams. Research from other states supports these findings, where more complex tests which promote critical thinking and performance-based literacy practices are embraced much more strongly by teachers (Hillocks Jr., 2002; Yeh, 2005).²

Contraction and Expansion

In a study of how high-stakes social studies tests affected teachers in Kentucky, Fickel (2006) finds that, in response to the testing, the school in the study underwent a reorganization to block scheduling, which in turn created an increase in the number of credits needed for students to graduate. This test-induced reorganization made room for the social studies department there to create a new, required freshmen history course that was based on the department's "shared beliefs about good teaching and learning" (p. 94). Other research finds that social studies teachers, in response to test-related demands for more reading instruction in their schools, are developing interdisciplinary, integrated curriculum units (Vogler, 2003), or are integrating more traditional literacy practices into their own social studies instruction (Barton, 2005; Fickel, 2006; Hess, 2005). By in large this tact must mainly be viewed as defensive. However, the integration of literacy development in the social studies is also consistent with the teaching of historical thinking (VanSledright, 2004) and "wise" or "ambitious" practice as well (Yeager & Davis, 2005).

The Special Case of Social Studies and High-Stakes Testing

Some social studies scholars have argued that evidence of high-stakes testing affecting instruction has been ambiguous and inconclusive at best, and claims that high-stakes testing leads to pedagogic teaching-to-the-test are perhaps over stated (see, e.g., Gerwin, 2004; Gradwell, 2006; Grant, 2003, 2006; van Hover, 2006). Indeed, several of these scholars have staked their research as counter arguments to those critics of high-stakes testing that assert that the tests essentially perform a pedagogically oppressive function in schooling. For instance, Gradwell (2006), in a case study of a single teacher, cites Grant (2003) and Gerwin (2004) in support of her position and concludes that the pedagogy of the teacher in the case study is not affected by high-stakes testing. Such claims pose an interesting dilemma. Generally, as discussed above, empirical research finds teachers adjusting both their instructional content and practice to the norms of the high-stakes standardized tests, but some scholars in social studies research cannot verify this claim in their own studies. Assuming the empirical strength of both sets of research, one conclusion to be drawn is that social studies education, as opposed to other subject areas, represents a special case in relation to high-stakes, standardized testing.

There are several factors that make social studies education a special case. One is the inconsistency of social studies testing generally. Since 1999 roughly only half of the states have required history tests. Furthermore, the vast majority of these tests are not used for student level accountability—only 10 states use history tests to make high-stakes decisions regarding student progress (Grant & Horn, 2006). As discussed earlier, the research has found that, if stakes are attached to the tests, then teachers are more likely to teach to the tests, and that the higher the stakes, the more teachers feel compelled to adjust both their content and pedagogies. It seems reasonable to assert that if the stakes attached to social studies tests are wildly inconsistent, then figuring out just how much social studies instruction is being changed because of testing would be difficult at best.

A second factor that may make social studies a special case is the current state of status quo social studies instruction. Several of the studies note how the high-stakes social studies tests were causing some teachers to alter their instructional styles to less student-centered, more lecture-based, textbook style teaching (Fickel, 2006; Segall, 2006; Smith, A. M., 2006; van Hover, 2006; van Hover & Heinecke, 2005), even as some of these scholars take pains to assert that the tests are not restricting pedagogy (see, e.g., van Hover, 2006). Lurking in the shadows of the research, however, is the possibility that high-stakes testing is not dramatically altering the classroom instruction of social studies teachers because social studies teachers may already be using lecture-based, textbook style instruction normally. As Ross (2000) points out, “The dominant pattern of classroom social studies pedagogy is characterized by text-oriented, whole group, teacher-centered

instruction, with an emphasis on memorization of factual information” (p. 47), an observation supported by Gerwin’s (2004) study of New York social studies classrooms. Scholars in social studies must be open to the possibility that one of the reasons that social studies instruction may not be as dramatically effected as other subjects by high-stakes testing, such as reading, is that a substantial portion of the status quo norms for social studies instruction may fit the teacher-centered, history-as-a-collection-of facts model of instruction.

A third factor that makes social studies high-stakes testing a special case is the mixed nature of the social studies tests themselves. It should come as no surprise that much of the social studies research refuting claims to test-controlled teaching comes from New York. As mentioned above, the New York test is a mix of multiple-choice questions and a document-based essay question. It appears that the existence of the DBQ allows social studies teachers there an increased *possibility* for good, solid social studies instruction that promotes historical thinking (VanSledright, 2004). This is the case because answering an essay based DBQ is essentially a critical literacy skill—a reading and critical analysis of text coupled with the writing of an essay. As such, social studies teachers, in teaching to prepare students for DBQs, have the charge of teaching their students a specific skill set instead of being forced to focus on a rigidly imposed collection of historical facts. Such a test structure, for instance, allows Bolgatz (2006) to argue that “it is possible to open conversations with children about the racial history of the United States while at the same time preparing them for standardized and high-stakes tests...” (p. 133) because, “In addition to opening important conversations about race and racism, studying primary documents offers a viable route to learning skills required on various standardized tests” (p. 145). As long as they are using primary documents for their inquiry into historical issues, teachers in Bolgatz’s study and others (see, e.g., Gradwell, 2006; Grant, 2003; Libresco, 2005) find increased flexibility in the content and pedagogy they use to teach social studies in their respective high-stakes environments.

This factor points to a substantial critique of the above cited studies. In their research, it seems that these scholars failed to fully account for the special case that New York in itself might represent. It seems that it is only because the New York State tests have the DBQ section that the researchers for these studies find evidence that high-stakes tests do not affect social studies instruction: In seeking out individual cases of “exemplary,” “wise,” or “powerful” social studies teaching, such research may have focused on anomalous instruction in the face of high-stakes tests. Indeed, this observation points to another critique of the existing research on New York. Even using their research as a base, the argument can still be made that the high-stakes tests are indeed affecting practice, as every study noted how teachers clearly changed their content to meet the demands of the testing, and to varying degrees, every study also showed some teacher alignment of pedagogy as well—even if it meant an increase in the explicit use of primary documents for historical inquiry, as is the case with the DBQs. The main issue seemed to be that

the change to the use of DBQs was viewed more positively by the researchers and their study participants as being more congruent with good social studies education.

This last point also speaks to a fourth factor in the contradictory findings and conclusions of social studies scholars in relation to the effects of high-stakes testing: the political nature of research. I raise this because, in my review and analysis of the body of social studies research, there exists a distinct contradiction. Even though some of these scholars identify clear changes to the content and pedagogy in social studies classrooms due to the pressures created by high-stakes testing, some continue to assert that such tests do not effect classroom instruction. For instance, van Hover (2006), in a categorical sleight-of-hand, separates “instructional decisions” from “content decisions” and “assessment decisions,” allowing her to empirically observe that the teachers in her study were compelled to quickly cover a breadth of detailed historical facts and also used “daily fact-recall reading quizzes” and multiple-choice questions for end-of-unit assessments, while simultaneously concluding that the Virginia high-stakes social studies tests have “virtually no influence on instructional decisions” (p. 206).

Van Hover’s analysis is problematic for several reasons. First, and most glaring, is that such a neat separation of instruction from content and assessment is overly simplistic. While daily fact-recall quizzes and multiple-choice assessments are clearly “assessment decisions,” they are also simultaneously decisions about instruction and content. A daily memorization quiz immediately communicates a particular form of instruction. Likewise, test-driven content alignment is also both a “content decision” and an “instructional decision,” as a teacher’s choice (or lack thereof) of content implicates pedagogy as well. Van Hover’s lack of a critical lens is indeed curious because the evidence in her study finds what most other studies of high-stakes testing find: Social Studies teachers are shifting pedagogy, content, and assessment towards alignment with high-stakes social studies tests, particularly if these tests consist of multiple-choice, historical fact memorization. Van Hover’s analysis then serves both as a justification for the maintenance of such programs while simultaneously creating space for her to launch an explicit attack on the critics of high-stakes testing (see, e.g., van Hover, 2006).

I do not intend to pick on van Hover personally or individually, but her work does demonstrate how the unspoken political perspectives of the researcher do impinge on that researcher’s findings, conclusions, and recommendations. Van Hover’s research in social studies does represent somewhat of an extreme, however, as the vast majority of the scholars in high-stakes test-related social studies research do not take the same pains to defend testing. However, the social studies do constitute a special case. By highlighting how their research may contradict existing studies of the impact of high-stakes on teaching, while also neglecting to explicitly outline the specific contexts of their studies, some of these scholars, perhaps unwittingly, provide a defense of high-stakes testing generally and an attack on critics of high-stakes testing. In this manner, some of these researchers have taken up a highly

political position, but have hidden it under the guise of neutral, or value-free social studies research.

Conclusion: Social Studies, Social Justice

My goal here has been to situate the social studies within high-stakes testing via the existing literature, and as one might expect, the results have been decidedly mixed: Social studies teachers are feeling some content and pedagogical constraints due to high-stakes testing, but there are specific aspects of social studies tests and social studies teaching that influence how this process plays out. In this way we can see the effects of the hegemony of high-stakes testing, where broader, structural powers do maintain an overarching control over the process of education, but also where different actors in specific contexts also have the power to offer some resistance to those controls (Apple, 1995).

The relative control of social studies instruction exerted through high-stakes tests raises significant issues for social justice education because the tests systematically push multicultural subject matter out of school curriculum because the tests do not include multicultural knowledge as important, valuable knowledge (Berlak, 2000; Bigelow, 2001; Darder & Torres, 2004; McNeil, 2000; McNeil & Valenzuela, 2001; Themba-Nixon, 2000). For instance, in a study of the New York state world history and geography tests, Grant (2001) found that Western nations dominate the test content. In another example, an analysis of Oregon's history test questions found that the tests completely neglected the role of racism in Oregon's history, instead offering an approach to history that "vastly oversimplifies complex social processes—and entirely erases ethnicity and race as categories of analysis" (Bigelow, 1999, p. 39). These examples demonstrate one of the consequences of the trivialization of knowledge that results from a reliance on multiple-choice, historical fact-based assessments (Pahl, 2003), and represent a view of history education as "cultural literacy," a nationalistic view associated with neo-conservatives that emphasizes the rote memorization of specific sets of canonical historical "facts" in pursuit of the development of a shared vocabulary of U.S. history (Seixas, 1993).

Further, as the high-stakes tests function to force schools to adopt a generic, standardized, non-multicultural curriculum, the "voices, the cultures, and the experiences of children" (McNeil, 2000, p. 232) are silenced, particularly if those voices, cultures, and experiences fall outside the norms of the tests (McNeil, 2005). In this way, students' lives, in all their variation, are effectively locked out of the curriculum by high-stakes tests as schools press to structure learning to fit the curricular norms established by the tests. Thus, high-stakes testing systems *require* diversity to be subtracted because of their emphasis on standardization (Valenzuela, 1999). McNeil (2005) sums up the "subtractive" logic thusly:

From inside classrooms we know that the system has to de-personalize, has to exclude, has to structure out personal and cultural identities to claim objectivity.

Social Studies, Social Justice

It has to silence differences, whether cultural, developmental, or idiosyncratic, or it loses its potency. *The system has to be subtractive or it cannot function as a generic, standardized, system.* (pp. 93-94, emphasis in original)

As such, student identity, in all of its many facets, is restricted and bracketed as existing outside of the definition of a test-defined education.

All is not lost, however. As a special case, social studies teachers do have some levels of autonomy relative to high-stakes testing and its pressures to standardize both knowledge and students. Social studies teachers working within a context of high-stakes testing operate at a specific nexus, a convergence of factors that allows them the freedom to conceive of social studies for social justice—a freedom that teachers in other subjects may not have. Such a conception of social studies education is important and necessary, especially given the race and class inequalities associated with high-stakes testing programs (Au, 2008; Orfield & Wald, 2000). It is a conception of multicultural social studies that recognizes the diversity of the world and the complexities associated with issues of racism, sexism, class oppression, and other forms of inequality (Bigelow, 1999; Hursh, 1997). It is also a conception of the social studies that recognizes a commitment to developing culturally relevant practices so that the needs of students of color, needs that standardized tests exclude, be met (Ladson-Billings, 1997).

Social studies for social justice also means continuing the development of critical literacy in students (Hursh & Ross, 2000). In this conception social studies teachers and students are conceived as agents of transformation in classrooms, schools, and communities (Marker, 2000) where “teachers and students...raise questions of whose knowledge is in the curriculum and how power and equality are maintained” and where “[s]tudents begin to learn how to develop questions and gather information in ways that enable them not only to better understand society but also to change it” (Hursh & Ross, 2000, p. 10). Because of their unique position, social studies teacher have the space and ability to challenge the pedagogic and content norms being leveraged by high-stakes testing policies, and, given a commitment to social justice, social studies teachers also are in a position to challenge the hegemonic logics of the tests themselves. Indeed, I would like to echo Segall’s (1999) suggestion of critical history education that creates,

a pedagogical environment in which the very foundations of history, as a discipline are called into question; a space in which history...is shaken—it’s habitual meanings and ways of making meaning...exposed as custom and the prescribed is unsettled by a shift into the elsewhere of the possible. (p. 371)

This is social studies for social justice: a challenging of hegemonic, status quo norms of historical knowledge, with visionary and pedagogic “shift into the elsewhere of the possible.”

Notes

¹The pressures that teachers are feeling, however, are not universally the same across all contexts. Comparative studies of differing states find that the higher the stakes, the more teachers focus their teaching on the tests (Clarke et al., 2003; Hampton, 2005). This has meant that lower performing states are feeling the most intense pressure due to high-stakes testing and accountability systems (Nichols, Glass, & Berliner, 2005).

² However, as Gerwin (2004) also notes, the use of DBQs does not necessarily lead to quality social studies teaching and education. In his study, he also finds teachers are using the primary documents in a very simplistic and shallow manner, “as a source of closed questions that cover content the way the textbook did” (p. 74). This finding leaves open the possibility that, even though a high-stakes social studies test may allow for better social studies teaching through the use of DBQs, it does not guarantee that good, student-centered, historical inquiry will necessarily follow.

³ It is worth noting that with earlier research (see van Hover & Heinecke, 2005), van Hover adopts a model of teacher deficiency by suggesting that, after finding that ten Virginia social studies teachers “seemed to be forced to compromise their commitment to wise practice” because of the state’s testing, she and coauthor Heinecke suggest that “teachers need help teaching *through the standards* rather than teaching *to the test*” and that, “teachers need guidance about how to use ‘wise practice’ . . . within the context of high-stakes assessments and fact-oriented content standards” (p. 112).

References

- Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from the classroom. *Theory Into Practice*, 42(1), 18-29.
- Agee, J. (2004). Negotiating a teaching identity. *Teachers College Record*, 106(4), 747-774.
- Apple, M. W. (1995). *Education and power* (2nd ed.). New York: Routledge.
- Au, W. (2008). *Unequal by design*. New York: Routledge.
- Barton, K. C. (2005). “I’m not saying these are going to be easy”. In E. A. Yeager & J. Davis, O.L. (Eds.), *Wise social studies teaching in an age of high-stakes testing* (pp. 11-31). Greenwich, CT: Information Age Publishing.
- Berlak, H. (2000). Cultural politics, the science of assessment and democratic renewal of public education. In A. Filer (Ed.), *Assessment* (pp. 189-207). New York: RoutledgeFalmer.
- Berliner, D. C., & Biddle, B. J. (1995). *The manufactured crisis*. Reading, MA: Addison-Wesley.
- Bigelow, B. (1999). Why standardized tests threaten multiculturalism. *Educational Leadership*, 56(7), 37-40.
- Bigelow, B. (2001). Standards and multiculturalism. In J. L. Kincheloe & D. Weil (Eds.), *Standards and schooling in the United States* (Vol. 2, pp. 695-702). Denver, CO: ABC-CLIO.
- Bolgatz, J. (2006). Using primary documents with fourth-grade students. In S. G. Grant (Ed.), *Measuring history* (pp. 133-156). Greenwich, CT: Information Age Publishing.
- Brimijoin, K. (2005). Differentiation and high-stakes testing. *Theory Into Practice*, 44(3), 254-261.
- Clarke, M., Shore, A., Rhoades, K., Abrams, L. M., Miao, J., & Li, J. (2003, January). *Perceived effects of state-mandated testing programs on teaching and learning*. Boston: National Board on Educational Testing and Public Policy, Lynch School of Education,

Social Studies, Social Justice

- Boston College. Retrieved August 20, 2007 from <http://www.bc.edu/research/nbetpp/reports.html>.
- Darder, A., & Torres, R. D. (2004). *After race*. New York: New York University Press.
- Emery, K., & Ohanian, S. (2004). *Why is corporate America bashing our public schools?* Portsmouth, NH: Heinemann.
- Fickel, L. H. (2006). Paradox of practice. In S. G. Grant (Ed.), *Measuring history* (pp. 75-103). Greenwich, CT: Information Age Publishing.
- Gayler, K. (2005, June). *How have exit exams changed our schools?* Washington, DC: Center on Education Policy. Retrieved March 27, 2006 from http://www.cep-dc.org/highschool/exit/change/CEP_HS_EE_9June2005.pdf.
- Gerwin, D. (2004). Preservice teachers report the impact of high-stakes testing. *The Social Studies*, 95(2), 71-74.
- Gradwell, J. M. (2006). Teaching in spite of, rather than because of, the test. In S. G. Grant (Ed.), *Measuring history* (pp. 157-176). Greenwich, CT: Information Age Publishing.
- Grant, S. G. (2001, October 3). When an 'A' is not enough. *Education Policy Analysis Archives*, 9(39). Retrieved February 8, 2006 from <http://epaa.asu.edu/epaa/v9n39/>.
- Grant, S. G. (2003). *History lessons*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Grant, S. G. (2005). More journey than end. In E. A. Yeager & J. Davis, O.L. (Eds.), *Wise social studies teaching in an age of high-stakes testing* (pp. 117-130). Greenwich, CT: Information Age Publishing.
- Grant, S. G. (2006). Research on history tests. In S. G. Grant (Ed.), *Measuring history* (pp. 29-56). Greenwich, CT: Information Age Publishing.
- Grant, S. G., Gradwell, J. M., Lauricella, A. M., Derme-Insinna, A., Pullano, L., & Tzetzio, K. (2002). When increasing stakes need not mean increasing standards. *Theory and Research in Social Education*, 30(4), 488-515.
- Grant, S. G., & Horn, C. L. (2006). The state of state-level history tests. In S. G. Grant (Ed.), *Measuring history* (pp. 9-27). Greenwich, CT: Information Age Publishing.
- Groves, P. (2002). 'Doesn't it feel morbid here?' *Educational Foundations*, 16(2), 15-31.
- Hampton, E. (2005). Standardized or sterilized? In A. Valenzuela (Ed.), *Leaving children behind* (pp. 179-199). Albany, NY: State University of New York.
- Hess, D. (2005). Wise practice in an innovative public school. In E. A. Yeager & J. Davis, O.L. (Eds.), *Wise social studies teaching in an age of high-stakes testing* (pp. 131-151). Greenwich, CT: Information Age Publishing.
- Hillocks Jr., G. (2002). *The testing trap*. New York: Teachers College Press.
- Hursh, D. W. (1997). Multicultural social studies. In E. W. Ross (Ed.), *The social studies curriculum* (pp. 107-119). Albany, NY: State University of New York Press.
- Hursh, D. W., & Ross, E. W. (2000). Democratic social education. In D. W. Hursh & E. W. Ross (Eds.), *Democratic social education* (pp. 1-22). New York: Falmer Press.
- Jones, G. M., Jones, B. D., & Hargrove, T. Y. (2003). *The unintended consequences of high-stakes testing*. New York: Rowman & Littlefield.
- Kornhaber, M. L., & Orfield, G. (2001). High-stakes testing policies. In G. Orfield & M. L. Kornhaber (Eds.), *Raising standards or raising barriers?* (pp. 1-18). New York: Century Foundation Press.
- Ladson-Billings, G. (1997). Crafting a culturally relevant social studies approach. In E. W. Ross (Ed.), *The social studies curriculum* (pp. 121-136). Albany, NY: State University of New York Press.
- Libresco, A. S. (2005). How she stopped worrying and learned to love the test...sort of. In E.

- A. Yeager & J. Davis, O.L. (Eds.), *Wise social studies teaching in an age of high-stakes testing* (pp. 33-49). Greenwich, CT: Information Age Publishing.
- Madaus, G. F. (1988). The influence of testing on the curriculum. In L. N. Tanner (Ed.), *Critical issues in curriculum* (pp. 83-121). Chicago: University of Chicago Press.
- Marker, P. (2000). Not only by our words. In D. W. Hursh & E. W. Ross (Eds.), *Democratic social education* (pp. 135-148). New York: Falmer Press.
- McNeil, L. M. (2000). *Contradictions of school reform*. New York: Routledge.
- McNeil, L. M. (2005). Faking equity. In A. Valenzuela (Ed.), *Leaving children behind* (pp. 57-112). Albany, NY: State University of New York.
- McNeil, L. M., & Valenzuela, A. (2001). The harmful impact of the TAAS system of testing in Texas. In G. Orfield & M. L. Kornhaber (Eds.), *Raising standards or raising barriers?* (1st ed.), (pp. 127-150). New York: The Century Foundation Press.
- National Commission on Excellence in Education. (1983, April). *A nation at risk*. Washington DC: United States Department of Education. Retrieved.
- Nichols, S. L., Glass, G. V., & Berliner, D. C. (2005, September). *High-stakes testing and student achievement* (No. EPSL-0509-105-EPRU). Tempe, AZ: Education Policy Research Unit, Education Policy Studies Laboratory, Arizona State University. Retrieved September 27, 2005 from <http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0509-105-EPRU.pdf>.
- Orfield, G., & Wald, J. (2000). Testing, testing. *The Nation*, 270(22), 38-40.
- Pahl, R. H. (2003). Assessment traps in K-12 social studies. *The Social Studies*, 94(5), 212-215.
- Pedulla, J. J., Abrams, L. M., Madaus, G. F., Russell, M. K., Ramos, M. A., & Miao, J. (2003, March). *Perceived effects of state-mandated testing programs on teaching and learning*. Boston: National Board on Educational Testing and Public Policy, Lynch School of Education, Boston College. Retrieved August 20, 2007 from <http://www.bc.edu/research/nbetpp/reports.html>.
- Perreault, G. (2000). The classroom impact of high-stress testing. *Education*, 120(4), 705-710.
- Popham, W. J. (2001). *The truth about testing*. Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).
- Renter, D. S., Scott, C., Kober, N., Chudowsky, N., Jofus, S., & Zabala, D. (2006, March 28). *From the capital to the classroom*. Washington, DC: Center on Education Policy. Retrieved March 28, 2006 from <http://www.cep-dc.org>.
- Ross, E. W. (2000). Redrawing the lines. In D. W. Hursh & E. W. Ross (Eds.), *Democratic social education* (pp. 43-63). New York: Falmer Press.
- Sacks, P. (1999). *Standardized minds*. Cambridge, MA: Perseus Books.
- Salinas, C. (2006). Teaching in a high-stakes testing setting. In S. G. Grant (Ed.), *Measuring history* (pp. 177-193). Greenwich, CT: Information Age Publishing.
- Savage, T. V. (2003). Assessment and quality social studies. *The Social Studies*, 94(5), 201-206.
- Segall, A. (1999). Critical history. *Theory and Research in Social Education*, 27(3), 358-374.
- Segall, A. (2006). Teaching in the age of accountability. In S. G. Grant (Ed.), *Measuring history* (pp. 105-132). Greenwich, CT: Information Age Publishing.
- Seixas, P. (1993). Parallel crises. *Journal of Curriculum Studies*, 25(3), 235-250.

- Sloan, K. (2005). Playing to the logic of the Texas accountability system. In A. Valenzuela (Ed.), *Leaving children behind* (pp. 153-178). Albany, NY: State University of New York Press.
- Smith, A. M. (2006). Negotiating control and protecting the private. In S. G. Grant (Ed.), *Measuring history* (pp. 221-247). Greenwich, CT: Information Age Publishing.
- Smith, M. L. (1991). Put to the test. *Educational Researcher*, 20(5), 8-11.
- Stecher, B. M., & Barron, S. (2001). Unintended consequences of test-based accountability when testing in 'milepost' grades. *Educational Assessment*, 7(4), 259-281.
- Taylor, G., Shepard, L., Kinner, F., & Rosenthal, J. (2001, September). *A survey of teachers' perspectives on high-stakes testing in Colorado*. Boulder, CO: CRESST/CREDE/University of Colorado at Boulder. Retrieved February 24, 2006 from <http://education.colorado.edu/epic/COStdRprts/COStdrpretitle.asp>.
- Teachers Network. (2007). A survey of teachers on No Child Left Behind. Retrieved April 4, 2007, from <http://www.teachersnetwork.org>
- Themba-Nixon, M. (2000). Testing slights multiculturalism. In K. Swope & B. Miner (Eds.), *Failing our kids* (pp. 32). Milwaukee, WI: Rethinking Schools.
- Toch, T. (2006). *Margins of error*. Washington, DC: Education Sector. Retrieved June 13, 2006 from <http://www.educationsector.org>.
- U.S. Department of Education. (2002, September). *No child left behind: A desktop reference*. Washington, DC: U.S. Department of Education, Office of the Under Secretary.
- Valenzuela, A. (1999). *Subtractive schooling*. New York: State University of New York Press.
- van Hover, S. D. (2006). Teaching history in the old dominion. In S. G. Grant (Ed.), *Measuring history* (pp. 195-219). Greenwich, CT: Information Age Publishing.
- van Hover, S. D., & Heinecke, W. (2005). The impact of accountability reform on the 'wise practice' of secondary history teachers. In E. A. Yeager & J. Davis, O.L. (Eds.), *Wise social studies teaching in an age of high-stakes testing* (pp. 89-105). Greenwich, CT: Information Age Publishing.
- VanSledright, B. A. (2004). What does it mean to think historically...and how do you teach it? *Social Education*, 68(3), 230-233.
- Vogler, K. E. (2003). Where does social studies fit in a high-stakes testing environment? *The Social Studies*, 94(5), 207-211.
- Vogler, K. E. (2005). Impact of a high school graduation examination on social studies teachers' instructional practices. *Journal of Social Studies Research*, 29(2), 19-33.
- von Zastrow, C. (2004, March). *Academic atrophy*. Washington, DC: Council for Basic Education. Retrieved February 12, 2006 from <http://www.ecs.org/html/Document.asp?chouseid=5058>.
- Wright, W.E., & Choi, D. (2005, December). *Voices from the classroom*. Tempe, AR: Arizona State University, Educational Policy Studies Laboratory. Retrieved February 12, 2006 from http://www.asu.edu/educ/eps/EPRU/epru_2005_Research_Writing.htm.
- Yeager, E. A., & Davis, J., O.L. (Eds.). (2005). *Wise social studies teaching in an age of high-stakes testing*. Greenwich, CT: Information Age Publishing.
- Yeh, S. S. (2005, October 28). Limiting the unintended consequences of high-stakes testing. *Education Policy Analysis Archives*, 13(43). Retrieved February 8, 2006 from <http://epaa.asu.edu/v13n43/>.