

12-2011

Assessing the Authenticity of the Document-Based Question Featured within the New York State Global History and Geography Regents

Clinton DePew
St. John Fisher College

[How has open access to Fisher Digital Publications benefited you?](#)

Follow this and additional works at: http://fisherpub.sjfc.edu/education_ETD_masters

 Part of the [Education Commons](#)

Recommended Citation

DePew, Clinton, "Assessing the Authenticity of the Document-Based Question Featured within the New York State Global History and Geography Regents" (2011). *Education Masters*. Paper 196.

Please note that the Recommended Citation provides general citation information and may not be appropriate for your discipline. To receive help in creating a citation based on your discipline, please visit <http://libguides.sjfc.edu/citations>.

This document is posted at http://fisherpub.sjfc.edu/education_ETD_masters/196 and is brought to you for free and open access by Fisher Digital Publications at St. John Fisher College. For more information, please contact fisherpub@sjfc.edu.

Assessing the Authenticity of the Document-Based Question Featured within the New York State Global History and Geography Regents

Abstract

This study questioned the validity of the Global History and Geography Regents exam DBQ as an authentic learning event and assessment. The previous five years of the Global History and Geography Regents DBQs administered in June were evaluated to assess the authenticity of the documents in regards to their context, the level of questioning featured in the task and constructed response questions, and the readability levels of the documents. Ultimately the analysis of the data disproved the validity of the DBQ as an authentic task as presented in the Global History and Geography Regents. In the end, teachers must infuse their curriculum with both challenging texts and authentic tasks to prepare students in accordance to new Common Core Standards.

Document Type

Thesis

Degree Name

MS in Literacy Education

Department

Education

Subject Categories

Education

Assessing the Authenticity of the Document-Based Question Featured within the
New York State Global History and Geography Regents

By
Clinton DePew

Submitted in partial fulfillment of the requirements for the degree
M.S. Literacy Education

Supervised by

Dr. Joellen Maples

School of Arts and Sciences

St. John Fisher College

December 2011

Abstract

This study questioned the validity of the Global History and Geography Regents exam DBQ as an authentic learning event and assessment. The previous five years of the Global History and Geography Regents DBQs administered in June were evaluated to assess the authenticity of the documents in regards to their context, the level of questioning featured in the task and constructed response questions, and the readability levels of the documents. Ultimately the analysis of the data disproved the validity of the DBQ as an authentic task as presented in the Global History and Geography Regents. In the end, teachers must infuse their curriculum with both challenging texts and authentic tasks to prepare students in accordance to new Common Core Standards.

Assessing the Authenticity of the Document-Based Question Featured within the New York State Global History and Geography Regents

Introduction

In 1965, President Lyndon Johnson enacted the first federal legislation regarding public education in an initiative to afford students from high poverty districts the same educational opportunities available through more affluent schools (Forte, 2010). Nearly three decades later President Clinton authorized the Improving America's Schools Act (IASA) that required states to establish common standards for all students in English and Mathematics grades 3-8 as well as assessments aligned to these standards. The assessments under IASA would allow the state to evaluate the effectiveness of the district, therefore establishing a hierarchy of accountability (Forte). This framework was further entrenched into public education with the passing of the No Child Left Behind Act (NCLB) in 2001 which established high stakes standards for districts as an effort to uphold accountability.

Though the federal government has been providing aid to public school districts for nearly half a century, the decade since the passage of NCLB has been wrought with controversy. As explained by Ellen Forte, NCLB "is supposed to be about improving achievement among low-achieving students in high poverty schools" (Forte, 2010, p.76). In a nut shell, assessments are utilized to identify schools that are in need of improvement, these schools then develop an improvement plan which enables greater success for its students in accordance to measurable objectives (Forte). Very few would dispute the need to improve struggling schools districts, the controversy lies in the methodology in which students and schools are assessed and whether or not interventionist practices on behalf of the government are beneficial.

Although federal legislation mandates the implementation of standards and assessments, states maintain various degrees of autonomy in creating standards for each grade level and subject as well as assessments that mirror the goals of these standards. As a result of NCLB however, each state's standardized assessments have become the sole basis for measuring student, teacher, administrative and district growth in the eyes of the federal government.

Challenges facing school districts during an economic recession may often feel overwhelming. While districts are forced to cut spending to the bare bones they must maintain or improve student scoring on high stakes assessment mandated by the Federal and state government. These tests not only intimidate the students that are required to take them they also create unparalleled levels of anxiety amongst the district's administrators, teachers and community members. According to NCLB legislature districts that fail to meet set standards or improve over time are victim to funding cuts for faculty and programs that appear crucial to student success (Forte). In many cases as funds are reallocated teacher aides, reading specialists and special education teachers are put on the chopping block to pay for outside tutors (Forte). In the most extreme circumstances, all of the educators in the building are required to reapply for their jobs as the school is scrutinized by the government or state (Forte). These potentially dire consequences force the community and its members to question: what if our students had done better on the test? A test, probably the educator's oldest tool, a way to measure retention or success in an easily calculated formula, a one size fits all solution to the question: what do my students know? What do they remember? Am I successful teacher? Is this a successful school? The answers to these complex questions are often interpreted through results on standardized assessments. With such black and white standards for success one would assume that the creators of the exams have done all that is possible to determine that the test is without bias. To

affirm that students from various economic, social, and cultural backgrounds are on equal footing when undertaking these exams. While many other papers and years of research have sought to identify these potential biases including Dodge (2009), Forte (2010), Rubin (2008) the tests may be flawed in the seemingly most basic of structures, their reading level. Assumptions would guide the reader to think that as the student progresses through the grade levels the difficulty of the text within the test increases at a correlating rate. Simply, the text should read at the grade level in which the test is given. Yet, teachers have repeatedly expressed at the conclusion of exams that documents utilized in the test were written at a level that was incomprehensible for their students. Indeed, by simply examining this year's 11th grade New York State U.S. History Regents Exam Document-based question (DBQ), Flesch Kincaid indicated that the readability of the documents ranged from a 7th grade level to that of a 5th year college student at the graduate level. Such disparities beg to question who believed that the chosen text would be a valid assessment of the student's ability to critically examine historical text? And does the task of the assessment reflect the objectivity of authentic learning?

Throughout their schooling students assimilate various literacies by actively engaging in authentic learning events. Essentially, they see value in gaining acceptance within a particular discourse and more importantly they have the tools and guidance to acquire it. Currently, the format of the DBQ tests only a subset of skills students must exemplify in real world learning environments. For these reasons, the construct validity of the NYS Global History and Geography Regents DBQ can be questioned.

It is understood that students taking the exam also range in their abilities to comprehend text, however the documents contained within the DBQ are frequently above grade level reading standards. As illustrated by Johns (2008), students forced to read texts that are too difficult or at

the ‘frustrational level,’ are rarely successful. Congruently, Johns (2008) asserts that students should be given materials at an instructional level in accordance to their reading abilities to promote growth and encourage academic success. At the instructional level students are taught to utilize a variety of resources to decipher difficult text, particularly in classrooms that emphasize collaborative learning and incorporate multimedia tools and methodologies for expression. To suddenly rob students of these tools causes the validity of the exam to fall into jeopardy. If the Regents is designed to assess the comprehension skills of students, then a test formatted in a structure that resembles an authentic literacy event with resources such as reference materials, peer review and more time would be necessary (Williams, 2003).

In an effort to assess the authenticity of the Global History and Geography Regents exam DBQ, this study analyzed what constitutes an authentic learning event in comparison to the tasks and materials available to students within the DBQ. According to Wiggins (1993), an authentic task reflects “the extent to which students experience questions and tasks under constraints as they typically and ‘naturally’ occur, with access to the tools that are usually available for solving such problems” (p. 214). More specifically, in regards to historical studies students must learn to “construct their story on the basis of evidence by selecting and arranging the facts” to “develop a persuasive argument” (Williams, 2003, p.11). In congruence with these demands, the previous 5 years of the June issued Global History and Geography Regents DBQs were evaluated based on the authenticity of the documents included and their presentation, the level of questioning featured in the task and constructed response questions, and the readability levels of the documents. To validate the study the data was collected and synthesized through previously employed formulas including Grant et al.’s (2004) basis for authenticating documents, Blooms Taxonomy levels of questioning, and three readability formulas readily available to the public.

The results of the data analysis revealed that the majority of documents were presented in an inauthentic manner as their source of original publication could be validated and the consensus of documents failed to offer differing perspectives of the topic of study. Congruently, the questions associated with the documents and task consisted of low level questions that almost never required students to synthesize the material into an argument. Finally, the readability levels of the documents were found to be better suited for students reading at the collegiate level which does not coincide with the 10th grade populace taking the exam. To conclude, the triangulation of these results disproved the validity of the DBQ as an authentic task as presented in the Global History and Geography Regents.

Theoretical Framework

The definition of literacy or the act of being literate is a social construction that when achieved signifies an individual's ability to effectively interact and engage with several discourses. Both Larson and Marsh (2005) and Gee (2001) agree that language can be summated as a social construction as a means of interacting with one's environment, or more simply, "learning language is learning how to mean" (Goodman, 2001, p. 317). A child does not become affluent in a particular literacy in isolation without guides or mentors. As explained by Gee (2001) the role of a mediator and immersion into the literacy are essential for growth. The school setting is no exception and the most exceptional teachers introduce new literacies through a collaborative process with learners at various levels of mastery of the skill or discourse. New literacies such as Wikis, blogs, chat rooms, Youtube, and even Twitter perhaps best exemplify the use of children's more contemporary literacy practices. These new literacies, in the eyes of Lankshear and Knobel (2003) signify the end of the Typographic Era in which texts centered on print, and the rise of the Post-typographic Era that encompasses a plethora of new multimedia

texts. The term ‘text’ may now assume an innumerable amount of roles including video, social networking, and online gaming. At the same time, never before has the technology to reproduce and manipulate such texts been so accessible, resulting in what Jenkins (2006) dubs the ‘insider generation.’ In accordance to Jenkins (2006), this generation is composed of literacy learners that can effectively and efficiently traverse new techno literacies without explicit instruction. Literacies that once took years to master are suddenly becoming implicit to modern students. Furthering his argument of new literacies, Jenkins (2006) portrays the manifestation of a participatory culture that focuses on collaborative experiences mediated by one’s peers that emphasizes core media literacy skills such as play, multitasking, collective intelligence, and transmedia navigation. While these new literacies appear to signify a revolution in the perceptions of literacy, government and educational policy have reverted to reductionist policies that characterized earlier eras based on skill and drill practices. These practices are best exemplified in the tasks of high stakes assessment which do not allow the resources students have become dependent on and continually engage with as a means of synthesizing information.

The counter argument, of course, is that contemporary literacies are essentially pseudo new literacies that fundamentally are reproductions of older, more proven methodologies of interacting with text. As an example, an email can be viewed simply as the electronic version of the letter, and chat rooms were once referred to as ‘sitting rooms’ or ‘social clubs.’ While these arguments are not without merit, the revolutionary aspects of new literacies are their accessibility, spontaneity, and ability to produce instantaneous results. Though ironic, it seems fitting that Gee’s (2001) definition of discourse as “a socially accepted association among ways of using language, of thinking, and of acting that can be used to identify oneself as a member of

a socially meaningful group or ‘social network’” (p. 537) utilized language that would be borrowed by the largest social networking site in the world, Facebook.

In accordance to psychological theory and cognitive development, children are believed to come “prewired” with structures or tools that support the process of language learning. Yet as argued by most contemporary theorists including Kucer (2005), Gee (2001), and Halliday (1969) a child actively constructs his or her perceptions of language and its functions. The new generation of students belongs to an insider generation whose definitions of text are drastically altering the ways in which people communicate in every aspect of their lives. Today’s students are continuously plugged in to interconnected digital literacy communities in which “Validity of knowledge...is established through peer review in an engaged community, and expertise entails understanding disputes and offering syntheses widely accepted by the community” (Greenhow, Robelia, & Hughes, 2009, p. 247). These practices have long been advocated by sociocultural and sociohistorical theorists that assume, “learning derives from participation in joint activities” and “is inextricably tied to social practices, and is mediated by artifacts over time” (p. 248). As students increasingly participate within digital literacies many researchers “argue that literacy today is necessarily social, “situationally specific” and a “multimodal, multimedial, dynamically changeable enterprise” (p. 250). In accordance to this trend, researchers such as Jenkins (2006) have debated the influence of Web 2.0 literacies on students’ primary literacy practices and the necessity of various stimuli to be engaged, one of which is instant and permanent access to others around them.

As expressed by Larson and Marsh (2005), “literacy is intimately tied to...what people *do* with literacy” (p. 20), or the literacy events that occur within sociocultural practices in which text plays an integral role. However, high stakes tests reduce the competency of the individual

by alienating them from their peers in an attempt to test their literacy skills in a medium that is foreign and intimidating in comparison to their primary discourse and most easily traversed literacies. Those that design the assessments such as the Regents exams would point out that students must also be independent and capable of being productive critical thinkers (NYSED, 1996). No one debates these points; instead it is the manner in which the student is assessed that proves problematic. Thankfully, Jenkins' (2006) work primarily entails new literacies in the ontological sense where his core media literacy skills incorporate properties of literacy acquisition that may be easily applied to all New Literacy Studies. By emphasizing the positive attributes of play as engagement in authentic problem solving, performance and simulation to promote improvisation and discovery, Jenkins (2006) appeals to student interests which increases motivation, and directly correlates with high success rates (Kucer, 2005). Accordingly, students assume responsibility for literacy learning. In addition, collaborative networking, and negotiation practices result in a larger collective intelligence that students learn to navigate for information (Jenkins, 2006). In essence, teamwork that mirrors the demands of the workplace (Jenkins), allows students to explore literacy in meaningful contexts, similar to the manner in which they develop their initial uses of language. In spite of Jenkins (2006) work NCLB mandates continue to emphasize high stakes testing that assess only a subset of skills (Horn, 2003). Compounding the issue is the disingenuous context in which the tasks of these tests are presented. By constricting the data available to students, denying students the ability to work collaboratively, limiting the time they are permitted to accomplish a task, and demanding students interpret texts that are at a frustrational level, the Global History and Geography Regents DBQ fails to adhere to the objectives set forth by NYSED and the objectivity of an authentic assessment.

Research Question

Unfortunately, standardized high stakes assessment has become a mainstay in educational dogma that influences everyday classroom practices particularly in schools that repeatedly fail to achieve desired marks. As an educator the current atmosphere may appear grim, and for the students even more disparaging. However, new literacies and consequently New Literacy Studies offer a basis for instruction that engages all the members of the classroom on an authentic stage. As language, like learning, is a socially mediated process which commences with an effort to manipulate, and correspond with, one's surroundings this paper will question: Does the NYS Global History and Geography Regents exam DBQ mirror an authentic learning task as suggested within NYSED's standards?

Literature Review

The literature review that follows examines the implementation of standardized assessments and their increased emphasis in light of a globalized economy. As indicated through the literature, high stakes assessment was instituted to promote accountability within the field of education from administrators to students. Though early evidence is mixed, the majority of research indicates that high stakes assessment have negatively influenced graduation rates, further alienated disenfranchised students and limited curriculum's in a manner that disallows for authentic learning. In response to these findings, when developing the document based question as a task on the Global History and Geography Regents NYSED sought to challenge students with an authentic task that mirrored the demands of historians. Although there is little research pertaining specifically to the Regents DBQ, Grant, Gradwell, and Cimbricz (2004) disprove the validity of the exam as an authentic task. The value of authentic learning tasks is defined and exemplified through several case studies which indicate increased student motivation and

production when engaging in real world problem solving. As a component of authenticity Johns (2008) and Fink (2006) advocate for differentiated instruction in accordance to students' reading levels. To effectively differentiate instruction educators need to establish the reading levels of the student and the readability of text with which the student is paired. Because the Regents exam does not take into account the reading levels of each student taking the exam this research was cited sparingly. However, as illustrated by Johns (2008) the readability level of a text greatly influences reader success. As a result of the aforementioned influence the works of O'Toole and King (2011) and Burke and Greenburg (2010) were cited for their various methodologies for determining the readability level of texts. Unfortunately, no previous research was discovered that discussed the reading levels of text featured on the Global History and Geography Regents DBQ.

Standardized Assessment

Test Based Accountability

As discussed within the introduction section, standardized assessment has steadily risen to prominence with the increased presence of federal legislation and funds for public education. In accordance to Loveless (2005) this movement began as a means of enforcing accountability systems at multiple levels within the field of education. Loveless (2005) aptly names this movement, 'test-based accountability.' Through his compilation of research on accountability systems, Loveless found these programs to be successful in boosting student performance during the latter half of the 1990's. Citing the research of Carnoy and Loeb (2003) as well as John Bishop (2001), Loveless (2005) concludes that the publication of test results and the implementation of other accountability driven incentives resulted in higher levels of achievement for specific districts and states including New York. In light of his findings Loveless

hypothesizes that the resistance to test-based accountability systems, led by students, teachers, and parents, is not the result of the compiled data. Instead, the backlash stems from the lack of desire to be held accountable and the inability for one to separate their philosophies on education from their political stance (Loveless). Loveless argues that, “Traditionalists tend to support measureable learning standards, describing in clear language the knowledge and skills that the students will learn” and “They are not offended by standardized tests with multiple choice items” (p.21). Whereas progressives “view educational curriculum more holistically, valuing the acquisition of inquiry and problem-solving skills as much as factual knowledge...Many progressives favor “real world” learning – that is, experiential as opposed to book learning – and “authentic assessment” as opposed to standardized tests with multiple choice items” (p.21). Most teachers and students, in an effort to maintain autonomy, believe in the overarching principles of progressive philosophies (Loveless). In the end, Loveless (2005) debates whether or not accountability systems will succeed in the face of such heavy opposition and if the field of education permits itself to current practices of accountability. Though Loveless achieves a relatively objective stance in his presentation of his findings the majority of his cited research concluded prior to the installation of NCLB and therefore failed to evaluate many of its hotly debated components.

Education in a Global Economy

While Loveless (2005) examined the effects of test-based accountability, Hursh (2007) sought to identify “the changing historical context of education and in particular how education is positioned differently within a globalized economy (p. 495).” According to Hursh the passage of the No Child Left Behind signified the rise of “neoliberalism” which as described by Tabb (2002) stresses, “the privatization of the public provision of goods and services (p. 29)”

including education (Hursh, 2007). Though prescribed as a predominantly economist policy the implications on education that entail from pro neoliberalism policies are startling. As neoliberals push for the “elevation of the free market above the public interest,” the government uniformly takes a back seat to its customarily interventionist strategies” (p. 496). Rather than the federal government regulating trade, welfare, Medicare and education, these industries are turned over to the free market to be run by private for profit industries (Hursh). Likewise, as schools, and thus education, are converted into profitable commodities competition among students, parents, districts and states will theoretically fuel progress and growth (Hursh). As a basic premise, parents and their students then become customers of the education system choosing where to invest their intellectual currency. Those in favor of school choice argue, “efficiency and equity in education can only be addressed through ‘choice’ and where family or individuals are constructed as the customers of educational services” (Robertson, 2000, p. 174). As a counter argument, one could question how a free market would create equal education opportunities when societies that employ capitalism are anything but equal in socioeconomic terms. Or as Hursh (2007) more succinctly stated, “such educational triage exacerbates educational inequality as the students who either pass or are close to passing the test become valued commodities and those students who need the most help are left to fend for themselves” (p. 507).

Hursh (2007) leaves no doubt as to his sentiments towards this ideal, warning the reader that the signs are already on the wall in the form of Charter schools, reconstruction of schools and school choice, all of which is currently funded by the government but determined mainly by high stakes assessments created by for profit companies. As an analogy imagine a child’s education in terms of a share of stock. No one buys a stock unless they believe it is a profitable

investment. Predicting a stock relies on measured assessment. Assessment of a company under this guise could mean a school, an AP program or a school struggling to pass proficiency requirements. The free market system was never intended to trade education like a commodity but if it is trusted in this manner, Hursh (2007) suggests that problematic disparities between achievement groups will only worsen and the authentic learning environment will be replaced by a rigorously structured curriculum that robs all members of the educational community of their autonomy.

Arnold Dodge attributes this rationale to availability heuristics. As defined by Dodge (2009), availability heuristic is “an oversimplified rule of thumb which occurs when people estimate the probability of an outcome based on how easy the outcome is to imagine” (p. 2). Thus, emotionally charged predictions that people can more easily relate to, are more easily imagined than vague, abstract thoughts or ideas (Dodge). More simply, a heuristic is a short cut; people will choose the heuristic that is most easily defined and recognizable whether or not they contain errors in marginalizing information. How this applies to schools is more aptly summarized by Dodge when he explains, “that the accountability of schools is fundamentally based upon the extent to which they satisfy the public’s perception of legitimacy” (p. 3). So, “If we can find criteria that the public perceives as legitimate, then we can use the criteria to measure the success of our schools” (p. 3) despite the fact that the measurement may not express improvement in learning. In summation, by dumbing down student test data to a few easily understood statistics the public will view the statistics and the assessments as a valid measurement, an availability heuristic (Dodge).

Dodge’s theory of availability heuristic demonstrates the power of public perception and its influence on the presentation of data. Congruently, Rubin’s (2008), “Theorem of intellectual

measure” further validates Dodge’s (2009) work as he analyzes the manner in which society and science attempt to measure human intelligence. Though there are numerous theories of intelligence Rubin (2008) summates three critical themes: 1. “the capacity to learn; 2. the total knowledge acquired; and 3. the ability to adapt successfully to a changing environment” (p. 5). While these three components can be assessed through a variety of means, Rubin (2008) emphasizes the necessity in maintaining ‘construct validity.’ According to Rubin (2008), construct validity “refers to the extent to which a measure correctly operationalizes the concepts being studied” (p.5). In other words, how accurately does the assessment test what is being measured? Rubin readily admits that “the ability to directly measure skills related to intelligence remains an elusive goal” and that “the score on a standardized test shows the degree to which an individual responded to the educational environment” (p. 5). Because standardized tests are intricately tied to the context in which the material is presented Rubin (2008) discovers that “the score on the standardized test may actually reveal differences in educational opportunities better than useful comparisons of intellectual capabilities” (p.7). Therein lays the connection to Dodge’s (2009) availability heuristic. As society seeks ways in which to measure an abstract capability such as intelligence, they marginalize their results by assessing a few concrete skills while ignoring contributing factors that cannot be easily accounted for such as socioeconomic status and educational opportunities. As an example, Rubin (2008) asks the reader to “consider a student who had every possible resource and opportunity for educational support with a score just above the median range on a national standardized test, compared to a student with sparse educational resources who scored just below the median level” (p. 7). Though the reader may at first assume that the first student was more intelligent, an argument could be made that the limited educational opportunities more greatly affect student two’s score and therefore

jeopardize the construct validity of the assessment. Despite the constraints of standardized tests exposed by Rubin (2008), the author defends these assessments “as the best alternative to date for determining a comparative measure of accumulated knowledge” (p. 11).

The Effects of High Stakes Assessment

In accordance to Rubin’s (2008) research on intellectual theorem, one would assume then that standardized assessments are used sparingly and as a way to judge the accrued knowledge of a student. However, as part of NCLB states such as New York developed rigorous standards that involve authentic and abstract skills such as the ability to find problems, solve problems identified by themselves or the teacher, work with others to arrive at solutions, and to present the results of their toils and findings. Indeed even the U.S. Department of Commerce, Education and Labor teamed with the National Institute of Literacy and the Small Businesses Administration to outline 21st century job skills that every student should possess upon graduation from high school. These skills included “the academic basics of reading, writing and computation” the ability to use an “array of advanced information, telecommunication and manufacturing technologies” and organizational skills such as “communication, analytical, problem solving, and interpersonal skills; creative thinking; and the ability to negotiate and influence and to self-manage” (Horn, 2003, p.37). The concern therein is in the ability of teachers to instill these skills in students and the standardized assessments’ ability to calculate student competency in such a large array of practices.

In spite of such lofty goals, well documented standards, interventionist practices in failing schools, and test-based accountability measures several researchers proclaim NCLB to be further widening the gap between high and low achieving students as well as white students and minority students. Horn’s (2003) research, though conducted at the onset of NCLB, documents

an increase in retention rates and dropout rates as a result of mandated high stakes standardized assessment as a the sole measure for graduation.. Forte's (2010) more contemporary research found that NCLB's objectives are not being met across the board as entire districts continue to miss Annual Yearly Progress goals (AYP) even while demonstrating growth amongst their student body. As further evidence, Forte explains the failure of remedial efforts instituted by the federal government for struggling districts which include the hiring of outside tutoring for low achieving students, federal funding cuts that undermine improvement plan goals and restructuring of schools. Whereas Forte (2010) puts the blame on flawed remediation strategies, Dodge (2009) attacks the philosophy of high stakes assessment and their use in measuring student performance and knowledge. As evidence to their flawed nature, Dodge quotes Nichols and Berliner (2008) "a system of rewards, punishment and pressures on self-esteem sounds like a logical way to motivate teachers and students, and some psychologists support this approach. But it doesn't work very well." (p.149). Or as Dodge (2009) puts it, "the pressure to perform may suit those who voluntarily choose such venues but to foist this arrangement onto a captive audience of youngsters is beyond the pale" (p.6). What results is a relatively new phenomenon of stress amongst school age children called 'test anxiety.' Though some stress is required to boost motivation as it intensifies "performance and learning collapse" (Goleman, 2007, p.271). To conclude, there are obviously many concerns surrounding test-based accountability and the foretold rise of neoliberalism by Hursh under NCLB legislation.

Teacher Anxiety and Effectiveness

Many of these issues are compounded by teachers who either lack the necessary skills to be effective or are so intimidated by accountability measures associated with high stakes testing that they are forced to teach to the test. These trends were blatantly evident to Gerwin (2004)

when he interviewed pre-service teachers in Queens, New York about their job prospectus, their willingness to teach in grades that were state or nationally assessed, and their ability to integrate historical documents into their curriculum. Immediately Gerwin (2004) noticed that nearly all of the teachers he observed and interviewed expressed a desire to teach in a grade that did not have a New York state Regents exam at its conclusion. Gerwin (2004) dubbed this preference the ‘steering effect’ and noted that most pre-service teachers believed that the rigors of the Regents exam and the necessity to continually review would constrain their teaching practices.

Interestingly however, Gerwin (2004) observed almost zero discrepancy in the manner in which social studies material was presented prior to the integration of the Regents exam and the methodologies that were being used by the pre-service teachers many years later. For this reason, Gerwin (2004) concluded that the Regents exam weighs heavily on the mind of the pre-service teachers interviewed and observed, but has minimal influence in daily lesson planning .

Gerwin (2004) attributed this to the stagnant practices that social studies teachers have used for nearly 30 years that are dependent on rote memorization and devoid of critical thinking skills.

Although Gerwin’s (2004) work examined only secondary social studies teachers, Dodge (2009) also uncovered the disturbing trend of teachers feeling forced to teach to the test. Much of the pressure that teachers feel is placed on them by administrators who seek higher test scores to improve their district’s NCLB profile. Contrary to Gerwin’s (2004) study Dodge (2009) was able to ascertain specific examples from across the nation of teachers modifying their daily teaching practices. This phenomenon of teaching to the test, was best exemplified by a parent letter that read, “My son attends arguably the best public middle school program in Baltimore, and the language arts teachers there have been told not to teach novels until the spring, after the state testing is over” (Myers, 2007, p. A35). As Dodge illustrated, even highly successful

teachers in superb programs have felt the pinch of high stakes assessment and have modified their curriculum to prep for the specific demands of these exams.

Unfortunately, not all teachers are as adept to change or skillful enough to prep their students for these exams. More often than not, those that are the least qualified are employed by low-achieving districts with a high population of low income minority students. Or as Haycock, Lankford, and Olson (2004) succinctly put it, “typically, and this is the case across the country, students who are the most dependent upon their teachers for academic learning are systematically assigned to teachers with the weakest knowledge and skills” (p. 230). As further evidence of this injustice, Haycock et al. (2004) points out that “poor and minority children are more likely than other children to be taught by uncertified teachers” (p.231) teachers “with no previous teaching experience” (p. 232) “teachers who do not have a major or minor in the subject they are teaching” (p. 233), or teachers that “have failed either the general knowledge or liberal arts and science certification exams” (p. 234) and teachers that attended lower quality undergraduate institutions. These discrepancies are due to a multitude of reasons including a lack of desire for highly qualified teachers to work with low achieving, high poverty minority students and the fact that the majority of teachers, 85 percent, teach within forty miles of where they grew up, meaning local talent often stays local (Haycock et al.). In an effort to validate the importance of effective teachers, Sanders founded the Value-Added Research and Assessment Center at the University of Tennessee, Knoxville where he examines individual teachers and the growth of their students (Haycock et al.). On average, “he finds that low-achieving students gain about 14 points each year on the Tennessee test when taught by the least effective teachers, but they gain more than 53 points when taught by the most effective teachers” (p. 237). These gains were also mirrored in middle and high achieving students (Haycock et al.). In summation, schools that

face the most scrutiny under NCLB legislation are at distinct disadvantages due to a lack of educational opportunities and perhaps more importantly, high percentages of ineffective teachers.

Faults in the Construction and Execution of NYS Regents Exams

Throughout this literature review it has become clear that high stakes assessment including the NYS Regents exams contain faults in their conceptual framework as well as their presentation. In Hursh's (2007) analysis of Regents exams administered and constructed at the start of the new millennium he discovered glaring injustices in the composition and scoring of numerous Regents exams. As examples Hursh (2007) points to the "Living Environments" exam in which only 39% of students who took the exam passed with above a 55% correct response and the June 2003 Math A Regents exam that recorded a 37% passing rate but was deemed to be so poorly constructed that the results were thrown out. These are extreme examples of the inadequacy of the Regents but Hursh (2007) also emphasizes the overriding power of the SED to change evaluative scoring of particular exams to obtain certain scores based on the needs of the state. So, if New York needs more federal grants but is deemed ineligible because their students test at proficient levels SED doctors the scoring of the exams to reflect lower achievement levels (Hursh). All of these indiscretions perpetuate a climate of distrust between the SED and the educational backbone of teachers, students and parents. How can those being assessed believe that the assessment and the scoring are valid?

One of the more highly anticipated changes to NYSED learning standards was the emphasis placed on multiculturalism and multiple perspectives (Maestri, 2006). In a state that epitomizes the 'melting pot' culture as a premier destination for immigrants over the centuries, one would assume that New York would continually be at the forefront in ensuring a

multicultural curriculum. Unfortunately, as discovered by Maestri (2006) in her analysis of the New York State U.S. History Regents, New York has failed to make a concerted effort to implement a multicultural curriculum. While two of the eight Learning Dimensions developed by NYSED (1996) are titled, “Unity and Diversity, and Multiculturalism and Multiple Perspectives,” minority groups including women continue to be represented in less than 20% of U.S. History Regents exam multiple choice questions (Maestri, 2006). Some may be quick to point out that the scope and sequence section of New York State standards includes many minority figures and events, thus they should be included in classroom instruction. However, as Commissioner of Education Richard Mills stated, “Instruction won’t change until the tests change” (p. 383). Or as Diane Ravitch put it, “Tests drive the curriculum...teachers teach what they think is likely to be on the standardized tests that their students will take” (p. 383). “Likely to be” may even be an understatement based on the work of David Bally (2010), a school teacher in NYC, in which he uncovered obvious trends of questions and skills that repeatedly appear on Regents exams. By simply focusing on redundant skills and material within the Regents exams Bally was able to significantly raise his students’ tests scores (2010). Similar to Bally, Maestri’s research implicates a perverse ignorance of the test coordinators to include disenfranchised people regularly on the U.S. History exam (2006). Case in point, by Maestri’s (2006) calculations from 1998 to 2005 not one question had featured Hispanics, on average there is less than one question per year on Native Americans and Asian Americans, an average of three on African Americans and one pertaining to women. Considering America’s history as a refuge for immigrants, proprietor of chattel slavery, champions of Manifest Destiny and home to a constitution created by the people and for the people, how can this discrimination be so grossly perpetuated by not including these histories on the state assessment? The answer to such a

question can be found seventeen years ago when in 1994 Thomas Sobol, then Commissioner of Education, proposed adopting the newly minted National History Standards to New York's Social Studies curriculum (Maestri). The proposal was met with such ferocity from critics, "who claimed that the standards included too much information on race and gender but not enough data on the traditional "facts" of American History" (p. 382) that Sobol balked on the initiative and reformed the standards in a second movement to more closely adhere to "traditional" history (Maestri). Though the blatant disparities in equity on the multiple choice questions in regards to gender and race seem obtuse, researchers including Maestri (2006), Fine (2005), Horn (2003) and Hursh (2007) have unearthed much more alarming research in regards to race, ethnicity, and gender in relation to success on the Regents.

The underrepresentation of minorities' histories on the NYS Regents U.S. History Regents exam as illustrated by Maestri's (2006) research may be just one link in a long chain that reveals the vast disparity of success between white students and minority students on high stakes assessment. Although a direct correlation cannot be explicitly established given the data gathered by Maestri (2006), her citation of NYSED's work on minority issues claims that while 81.8 percent of white students passed the U.S. History exam 77.6 percent of Asians, 63.7 of American Indians, 52 percent of African Americans and only 48.6 percent of Hispanics accomplished the same feat. Again, given the lack of culturally relevant material minority students may not perceive the information as relevant or worth knowing. An analogous bias was also discerned when comparing the success of males versus females (Maestri). Women given their hierarchy in the material presented may perceive the curriculum in the same light as minorities; their histories' do not matter. Congruently, while the education field is typically dominated by the female gender a much larger portion of males receive their bachelor's in Social

Studies Education than females (Maestri). Maestri's (2006) compiled data certainly raises many questions as well as a call for more research within the years since her study concluded.

In comparison to some of her peers, Maestri's (2006) conclusions about race and gender disparity within the NYS U.S. History Regents exam appear to just scratch the surface of a much deeper seeded issue within education reform. For years researchers such as Fine (2005), Horn (2003) and Hursh (2007) have been investigating the widening gap between the success of white students in comparison to students belonging to minorities in spite of reforms instituted by NCLB meant to counter the growing inequality. In a presentation to the Board of Regents of New York State in 2005, Fine plainly outlined the miscues associated with high stakes assessment, namely the Regents, as a graduation requirement for all students. As a result, Fine (2005) explains graduation rates have dropped significantly, "with rates less than 40% for black and Latino students" (p.25). Even more discouraging is the rise of what Fine coins, "disappeared" students, students that inexplicably fall of the radar but have not been officially recorded as drop outs, the majority of which are students of, "color attending under-resourced schools in low income neighborhoods" (p. 25). Horn (2003) reflects the same concern in her research on high stakes assessment in Texas, North Carolina and Massachusetts specifically. Because of the increased scrutiny under which districts and states are held, many have enacted extraordinary measures to ensure that their school is successful...on paper. Both Horn(2003) and Hursh(2007) revealed schools in Texas that retained students, labeled them as learning disabled or simply shipped them out of district to achieve higher marks on the TAAS, an exam similar to the Regents. As exemplified by the aforementioned researchers NCLB legislation appears to be causing the reverse of their intended changes by further alienating disenfranchised youth.

Inauthenticity of the New York State Global History and Geography Regents

In 2005, the New York State Education Department declared that, “social studies skills are not learned in isolation but rather in context as students gather, organize, use and present information” (p. 12). Yet, when comparing the goals and objectives of NYSED and the Board of Regents to their own examinations there lies a distinct lack of correlation. Though assessing different content the construction of the two Regents history exams, United States History and Global History and Geography, both contain the same simple format: fifty multiple choice questions, a thematic essay on a predetermined topic, and a Document Based Question essay (DBQ). All students within New York State regardless of ability and intellect must take and pass these two examinations if they wish to receive a Regents diploma upon graduating high school. The tests are administered in June for the majority of students, and January and August for students in advanced standing or for those that did not pass on the previous attempt. The tests are taken in isolation within a three hour time limit unless indicated otherwise by a student’s individualized education plan or 504 plan. Unlike the standards set forth by NYSED (1996) the examination inhibits students from collaborating with their peers, expressing their findings through a variety of mediums, validating sources, and developing uniquely created questions and hypotheses, all of which are objectives created by NYSED (p. 13-14).

Grant et al. (2004) arrived at similar conclusions in their analysis of the document based question citing research by Wiggins (1989) that, “schools assessments are “typically *inauthentic* , designed as they are to shake out a grade rather than allowing students to exhibit mastery of the knowledge” (p. 310). As a counter argument to the assessment status quo Wiggins (1993) argued that contemporary assessments value, “reliability over validity” and that by subjecting

students knowledge to the constraints of “forced choice” questions test makers were sacrificing the validity of their exam as an accurate measurement of “intellectual value” (p. 310). When developing the modern Regents exam, as demonstrated earlier, it was clear that NYSED was well aware of the power of authenticity and the value of its use within the classroom. Thus, they implemented what they perceived as an authentic challenge with the implementation of the DBQ.

The DBQ as it appears on the Regents typically consists of 7-10 primary and secondary source documents that are of a variety of modalities including political cartoons, quotes, maps, legislation, court cases, inauguration addresses, photographs and diary entries. Following each document are one or two constructed response questions addressing the main idea of the document which correlate with an essay prompt and an accompanying historical context. As evidence to the tasks authenticity, in 2002 Larson, then a NYSED representative responded to Grant and his fellow researchers with a curt email that read, “When SED (state education department) moved to a standards-based curriculum, instructional and assessment program, the DBQ was introduced. DBQ’s are examples of authentic assessment as it (sic) mirrors what historians do” (Grant et al, 2004, p. 314). Though not explicitly stated above, the majority of Larson’s argument is based on the premise that the DBQ’s of the Regents exam are similar to those featured within AP exams that have withstood critique as a valid assessment (Grant et al.). However, the AP exam DBQ is constructed so that the student must undertake a perspective and support their argument using the documents. To contrast, the essay prompt written for the majority of the Regents exams requires students to perform lower level thinking skills such as discuss, describe and explain (Grant et al.) in other words avoid synthesis, analysis, questioning, and critique. By subscribing to lower level questioning, the task of the DBQ jeopardizes its namesake as an authentic assessment.

True, part of what historians do involves analyzing documents to discern their meaning but as reiterated by Grant et al. (2004) historians search for these documents and select them based on the contextual questions of their research. Unlike the documents that appear on the Regents, historians' sources are not collected for them nor do they appear in prearranged groups that in turn make their orientation a representation of the biases of the original collector. Compounding the issue in selecting the documents is the blatant editing from their original form to what appears on the Regents exam (Grant et al.). If the DBQ is an authentic assessment because, "it mirrors what historians do" than children across the country would be traveling the world to examine documents in their original state as historians frequently enjoy (Grant et al.). Obviously students are not afforded the same opportunities as professional historians for innumerable reasons, but the point remains that as a method to improve authenticity historians search out their own resources and rely on their own interpretations, not those of a governing power.

As No Child Left Behind continues to push for privatization of schools, competition among students and to take away state and district autonomy, "civil society is weakened and is held accountable by the government rather than the other way around" (Hursh, 2007, p. 514). The result is a counter revolution that stresses authentic learning environments and tasks in which student centered learning and teaching are the primary avenues of knowledge acquisition. Though these ideals originated nearly one hundred years ago Dewey's dream for education to become central to all our activities is now perpetuated by researchers such as Olssen et al. (2004) that call, "for an education state, claiming that a deep and robust democracy at a national level requires a strong civil society based on norms of trust and active responsible citizenship with

education central to such a goal” (p. 1-2). In an area of uncertainty and change educators must ensure that today’s youth will become critical, active citizens in the future.

The Value of Authentic Assessment

Wiggins (1993) declared that, “validity of assessments should be considered in terms of authenticity” which Wiggins generally defines as “the extent to which students experience questions and tasks under constraints as they typically and ‘naturally’ occur, with access to the tools that are usually available for solving such problems” (p. 214). As explained by Grant et al. (2004) “ a discipline that features argument, interpretation, and multiple perspectives , history is especially resistant to simple forms of assessment” (2004). For that reason, high stakes assessment as highlighted throughout this essay have proven to be unsuccessful. Rather than assessing students after a cumulative year or two of study theorists and researchers such as Wiggins (1993) and Grant et al. (2004) suggest using authentic assessment, “as a regular feature of classroom practice” (p. 314). Specifically, “students should regularly solve engaging and worthy problems, produce a quality product and/or performance, undertake projects that allow for frequent interactions between teacher and student, and have the opportunity to demonstrate habitual patterns of thinking and performing” (Ward, 1995, p. 206-7). Research has repeatedly demonstrated the success of teachers who effectively implement authentic learning events and projects within their classrooms. The following passages exemplify some of these teachers whose ideas have been documented within the last decade.

Some of the most powerful and authentic literacy events are created or developed from critical literacy activities and knowledge acquiesced during the student’s exploration of the topic and themselves. Pestacore (2008) came to this realization when she decided to institute current events into her Regents level high school English class. The premise came to Pestacore as she

was taking classes for her Ph. D in education when she began to see critical literacy as an avenue to create, “citizens who are empowered and emboldened to act as a result of their enlightenment” (p.330). Through the use of a New York Times article on Global Warming, Pestacore (2008) was able to strike a chord with her students as she revealed hidden biases of writers that can be found within their publications. Pestacore (2008) carefully scaffolded the student’s questioning and readings, but she was inspired by the students’ enthusiasm and their own questions for her, the authors, other students and of themselves. The consistent scaffolding of questions and research allowed Pestacore’s students adequate time to assimilate to the material and better engage with their peers in a critical discussion. Though the New York Times articles and subsequent research varied in pertains to their reading level the students continuously utilized their peers and alternate resources to grasp the contents of the material (Pestacore). With a growing understanding of the material, the research began to mean something different for each student in the class. Through further interaction with their peers and reflective writing processes they were able to clarify their thoughts and construct valued opinions (Pestacore). These practices directly coincide with Jenkins (2006) belief in the rise of a participatory culture and Greenhow et al.’s (2009) argument that literacy “is necessarily social” (p. 250). In the end, many of the students decided to write letters to the editor of the Times sharing their perspectives (Pestacore, 2008). By allowing students to question, research, validate, reflect, share, and construct new ideas with a real world issue they developed “cultural capital – the ability, knowledge, and skill to manipulate, strategize, and position themselves in the culture to maximize their gain” (p. 335). Ultimately, critical literacy and the usage of authentic learning environments equip students with tools to become, “catalyst(s) for action when one sees injustice or oppression” (p. 335). Although Pestacore (2008) is a strong advocate for more local,

autonomous, meaningful assessment she cites the benefits of teaching students to be analytical thinkers are their improved results on state assessments.

Scheidet's 2003 study of a classroom in Mount Sinai New York reflected many of the same results exhibited by Pestacore (2008) as he witnessed a 10th grade Global History and Geography teacher implement a web based curriculum with one of his classes. The study was carried out over the course of a year through several observations and interviews with the teacher and students of the class (Scheidet, 2003). As a control, one class was infused with the new web based curriculum while the teacher's other Global class was taught with traditional classroom practices that depended on the textbook (Scheidet). At the conclusion of the study, Scheidet and the teacher noted marked increases in student motivation and interest, parent involvement and higher test scores on the Regents exam for those that participated in the web based classroom. Both Scheidet and the teacher attribute these gains to the framework of the web based classroom, which allowed students "to build on previous knowledge, develop personal connections to conceptual material, and to improve their ability to apply information to solve problems" (p.90). Additionally, the teacher perceived the web based curriculum to be more beneficial because "Project based learning provided more options to help meet individual needs... There were more opportunities for students to operate at their own pace... The teacher became the facilitator of information rather than a director of the class" (p.91). Scheidet's (2003) study was not without its limitations. Most notably the teacher did not attempt to implement a project based learning curriculum with traditional materials in the classroom. Instead the teacher relied solely on the capabilities of the internet and computer based technologies to integrate project based learning. Overall, however the study verified the power of authentic learning to better engage students and improve achievement.

Though Scheidet (2003) neglected to include examples of student tasks and to define project based learning (PBL), Toolin (2004) readily outlined the goals of project based learning: “to investigate real world, standards-based problems that are of interest, relevance, value and worth to students and teachers over a sustained period of time” (179). “Projects” are defined by questions or problems that are collaboratively investigated by students and teachers utilizing technology and resulting in a series of artifacts or products that address the question of problem over time” (p. 179-180). In Toolin’s (2004) examination of two New York City districts from the respective east and west sides, she determined that science teachers that infused PBL with their standards based curriculum better motivated students, developed critical thinking skills and achieved higher scores on the Regents examination at the end of the year. Further validating Toolin’s (2004) study was the refusal of two observed teachers to implement PBL and the struggles they had in motivating students and congruent low test scores. As noted by Toolin, a critical component to the creation of a PBL curriculum is the opportunity for continuous professional development for teachers.

Reich and Bally (2010) echoed the necessity of professional development as they, presented the benefits of “community of practice” in which “groups of teachers...meet regularly to discuss their practice” (p.179). These communities “are able to build a sense of shared goals, values, and ideas about what is effective” and “are able to successfully improve their teaching” (p.179). In Bally’s experience, the community of practice analyzed the New York State Global History and Geography Regents to identify “patterns in the knowledge and skills that reliably appeared” (p.180) which they then used to develop a document “that outlined challenges, skills, big content areas, and themes that the exams consistently focused on” (p.180). With the support and tools to tackle the Regents, Bally was able to better prepare his students for the demands of

the Regents exam (Reich & Bally, 2010). To conclude, the influence of PBL and professional development on improving teacher and student success cannot be ignored and deserve further inquiry.

Reading Difficulties in Regards to Documents

When searching for publications on the readability levels of texts included in Regents exam I found nothing within many the online article databases. However, a particular area of concern in Bally's (2010) analysis of the Global History and Geography exam was his students' ability to read and respond to documents within the exam. It quickly became apparent to Bally (2010), "that when we read in class, my students gave up after a paragraph out of frustration or fatigue" (p.180). Although the readability of the documents was not assessed by Bally the frustration of his students may in large part be due to that fact that the reading levels required to comprehend the text were above his students' levels or at the frustrational level. In accordance to Johns' (2008) publication "the frustration level is that level at which a student should never be given materials to read" (p. 12). Determining the frustrational level of text requires first an analysis of the text to determine its readability level and more importantly an informed assessment of what levels of text the student succeeds and struggles. Students should be evaluated at three reading levels, independent, instructional and frustrational, which reflect the student's ability to fluently read the text (Johns). As discussed by Johns (2008), "If students are placed in instructional materials...they tend to be successful readers who are on task. Unfortunately, many students are placed in materials that are too difficult for them. These students fail to benefit much from lessons using grade level texts." (p. 4). Considering the incredible array of students and their coinciding range of abilities that take the exam, one could postulate that a number of students' struggles may in part be the result of frustrational level texts

within the exam. The lack of research in regards to this concern further warrants the investigation of this study into the readability levels of documents included in the DBQ portion of the Global History and Geography Regents.

Readability

In accordance to previously stated standards, in order for a document to be presented in an authentic context, the reader must be able to ascertain the validity of the material through analysis of its contents and assumed bias. Thus, the validity of a text is contingent on the researcher's ability to read and comprehend the diction within the document. As documents may refer to a number of materials that span various areas of space and time, one should never assume that all documents can be deciphered by all readers. For this reason, educators and researchers alike must tackle the task of evaluating the readability of a document. As defined by Zakaluk and Samuels (1988), readability "is a concept that attempts to capture the ease with which learners access that material" (O'Toole & King, 2011, p.181). Though experienced teachers and educators may believe themselves capable of an 'eye test' as an accurate assessment of the readability of a document, Burke and Greenburg (2010) strongly advise against this informal assessment. Instead Burke and Greenburg in their 2010 study "Determining readability: How to select and apply easy-to-use readability formulas to assess the difficulty of adult literacy materials" have outlined several formulaic methodologies for determining the readability of a text. Two of the most widely used methods exemplified by Burke and Greenburg (2010) are the Flesh Kincaid and Dale Chall formulas. As explained within their research, both methods can be easily accessed, Flesh Kincaid is contained with Microsoft Word and the Dale Chall can be found at Okapi! Website (<http://www.interventioncentral.org/htmldocs/tools/okapi/okapi.php>). However, the formulas of

both are dependent on differing variables. Flesh Kincaid's formula relies on sentence length, the number of words within a sentence, and word difficulty, which takes into account the number of syllables in each word. Presumably, the greater the sentence length and word difficulty the more difficult the text passage is to comprehend. According to Burke and Greenburg's (2010) research there are limitations to the Flesh Kincaid formula as they discovered that it works best for a running narrative and the grade equivalency determined by the software "tends to underestimate the difficulty of the passages by approximately two grade levels" (p.35). While the Flesh Kincaid emphasizes the importance of word difficulty in relation to syllables, the Dale Chall formula identifies words that are not commonly found within a list of 3,000 predetermined words to judge difficulty. The Dale Chall combines the words not found on its list and couples this measurement with sentence length to gauge the difficulty of the text. The resulting grade equivalency proves the "most reliable and validated of the readability formulas" in the words of Burke and Greenburg (p. 36). Because each readability formula takes into account different variables one must learn to use numerous formulas to find an average readability and understand that the resulting estimates are not absolute. Despite the various tools available to the consumer, O'Toole and King (2011) make a strong case for determining the readability of a text via cloze test that can be developed and scored by hand. According to O'Toole and King (2011), the cloze test proves most valid when determining the readability of a document for an individual student. However, there are apparent difficulties in preparing a cloze test and the scoring of it can be quite time consuming, both of which jeopardize the practicality of its use. As with any assessment discretion on the behalf of the researcher must be used in choosing the appropriate readability formula in regards to the type of text as well as the reader.

The ability to accurately determine the readability of a text better enables educators to pair the text with a reader's instructional level. Fink (2006) dubs this approach 'responsive instruction' and claims that it makes "instruction easier and more effective" (p. 131). As claimed by Johns (2008) readers at the instructional level "can make maximum progress in reading with teacher guidance" (p. 7) and the independent level is achieved "when students read "fluently with excellent comprehension" (p.7). These levels are established so that teachers can differentiate instruction based on the strengths and needs of their individual students. Based on the descriptors of the various reading levels students are most successful when working alone when given materials that are at their independent level. Yet, as illustrated by Bally (2010) students frequently encounter text on exams that does not take into account their independent reading levels and consequently forces students to engage with materials that are at the frustrational level. As stated previously, students that are given materials at the frustrational level fail to benefit from its contents as it is beyond their comprehension abilities.

Conclusion

Despite the positive intentions of NCLB and test-based accountability, numerous studies have noted its lack of success and negative consequences for both students and teachers. New York state and the Regents exam are no exception as these assessments have greatly influenced classroom practices in a negative fashion and often reduced the curriculum to only a small subset of skills. This is not to say that all educators "teach to the test." On the contrary, Toolin (2004), Scheidet (2003) and Pestacore (2007) all demonstrated the power of authentic assessment and project based learning to boost student engagement and to facilitate critical thinking skills. While authentic tasks require real world problems that require inquisitive, reflective, and collaborative learning, assessments such as the New York state Global History and Geography

Regents continue to test students in inauthentic contexts. Further compounding the difficulties of the Global History and Geography Regents is the readability levels of the documents included in the DBQ section of the exam. The combination of these issues requires an investigation as to whether or not the DBQ within the Global History and Geography Regents exam presents an authentic learning task.

Methods

Context

In accordance to New York State Education Department law, all students are required to take the Global History and Geography Regents after completing the course over a two year period (NYSED, 2011). The exam is administered in both June and January, however in recent years a lack of funding has caused some inconsistencies (NYSED, 2011). Traditionally the exam is taken by students nearing the end of their sophomore year of high school, though students who do not achieve a 65 on the test must retake the exam either in January or June the following year. Students, except those with disabilities, are given three hours to complete the exam which consists of 50 multiple choice questions, a document based essay and a thematic essay. According to NYSED (2011) regulations the exams are scored by qualified teachers, in this case those that are certified Social Studies 7-12 and have had experience grading classroom exams or state required tests. The scorers of the exams may score the multiple choice section by hand or by machine. When evaluating the DBQ and thematic essays two raters utilize rubrics provided by the state to score the essays on a scale of 1-5 with one being the lowest and 5 the highest grade possible (NYSED). Once the grading of the exam has been completed by two raters the student's scores from all components of the test are converted into a numerical score via chart

developed by NYSED. Afterwards the score is finalized unless an exceptional circumstance occurs.

The participants within the study both teach at the high school within the Grand Valley district (pseudonym) located in the Southern Tier of New York. Geographically the district remains unique in comparison to all other districts across the U.S. as it is the only public school district located entirely on an Indian Reservation. The reservation is home to the Seneca Nation of Indians who comprise 34% of the student population. However, this number only reflects those students who are enrolled as members of the Seneca Nation through their mother's lineage. 59% of the remaining student body is white, while 2% are African American and another 2% claim Hispanic or Latino heritage. According to the school report card (2011) 40% of the students are eligible for free lunch and another 11% of students are eligible for reduced lunch. In relation to AYP goals the district failed to graduate 66% of students set to graduate in the 2009-2010 school year, graduating only 63% of eligible students. The graduation rate was even worse for those who were economically disadvantaged which graduated only 53% of the 40 students within this sub group (NYSED). Overall, the district's student body proves ethnically, economically and academically diverse.

Participants

Aside from the DBQ's that I analyzed, this study included email exchanges with the two Global History teachers at Grand Valley High School. The first to respond to the email exchanges was Erie (pseudonym), a five year teacher who currently teaches 9th grade students, though she will loop with them to 10th grade. Erie is a white female who has been with the Grand Valley District for all of her five years in teaching. The other participant Tom

(pseudonym) has been teaching Global History for ten years at the Grand Valley District.

Currently, Tom teaches the 10th grade half of the Global History course although he will teach 9th grade students next year as Erie loops with her class

Researcher Stance

For this study, I have assumed the role of a passive observer simply collecting data from the documents and two of my colleagues. For the past three years, I have been a long term substitute at Grand Valley Middle School fulfilling the role of the 7th grade and 8th grade history teacher. Since graduating from SUNY Fredonia in 2008 with a Bachelor's degree in Adolescence Education with a major in history I have sought my Master's degree in Literacy at St. John Fisher. As I do not directly have a stake in the outcome of this research, my objectivity to the data is shaped only by my findings. Upon completion of this study, I will share my findings with other history teachers within the Grand Valley district.

Methods

Gathering DBQ's that were administered over the last five year was achieved by accessing the archives from the NYSED website. The analysis of the documents in congruence to their authenticity was be based on a number of factors. Williams' work "The Historians Toolbox" (2003) deftly outlines the tasks historians tackle around the world when constructing a historical research paper or essay. Language borrowed from Williams as well as Bloom's taxonomy was compared to that used in the constructed response questions accompanying the documents as well as the DBQ's central task. For example, Williams (2004) states several questions and concepts that are key to historical research including "creating a narrative or an argument based on verifiable evidence," and "what is the meaning of the events studied in terms

of both the past and present?” (p. 12). At the same time, Bloom’s taxonomy and its congruent vocabulary indicates the cognitive attention necessary to complete a given task.

Equally important in regards to authenticity are the documents that were chosen. Do they display various perspectives, from what sources they were derived, in what manner or form are they presented (are they truncated) and what is the reading level necessary to adequately comprehend the document’s language? Documents that have been compromised include those that are truncated, whose sources are not verifiable, contain obvious bias, and are presented solely through one vantage point as these documents do not present a viable record of the event. To evaluate the source of the document, its presentation and the perspectives or viewpoints utilized I conducted Google searches for each document. If the document was not readily found, I reworded the search until relevant results appeared. For each DBQ a separate chart was used to record the findings for all of the documents.

As a means of testing the documents’ readability three methodologies were utilized including the Lexile framework, Flesch Kincaid and Dale Chall. These methodologies have been chosen because their formulas do not measure text readability using the same calculations. Thus, by triangulating the readability ratings of these three programs the results became more credible. In addition, all of these methods have been endorsed by the new Common Core Standards that are replacing much of New York State’s learning standards at all levels and subjects (Common Core, 2011). The Flesch Kincaid test, contained within Microsoft Word 2003-2007, depends on word length and sentence length to illustrate semantic and syntactic complexity. Simply, the longer the words and sentences the more difficult the text appears to be. Dale Chall results differ by replacing the stress on word length with word frequency. Words that appear commonly are assumedly not as difficult whereas words that appear less are unfamiliar

and therefore more challenging. The Lexile Framework measures a text in regards to word frequency and sentence length and an analysis utilizing MetaMetrics. As a result of the analysis the researcher will better match a student to a text at their reading level.

Quality and Credibility

According to Mills (2011) credibility “refers to the researcher’s ability to take into account the complexities that present themselves in a study and to deal with patterns that are not easily explained” (p. 104). Congruently, Mills advocates adhering to Guba’s (1981) framework which includes “prolonged participation at the study site,” “persistent observation,” “peer debriefing,” “triangulation,” “collection of documents,” “member checks,” and “structural corroboration” (p. 104). The credibility of this study was reassured through a number of these practices. By utilizing a variety of data collection methods the resulting triangulation should help to eliminate discrepancy. Congruently, by analyzing five years of the DBQ I hope to “overcome distortions” within the test they may occur between exams (Mills, 2010, p. 104). Additionally, the use of a critical colleague and the consultation of other social studies teachers will provide insight, reflection, and help in constructing a valid study.

In addition to Guba’s (1981) framework for credibility, the researcher recognizes that when performing qualitative research all data is context bond and must be presented as such. Guba (1981) refers to the contextualization of data as ‘transferability.’ Because of the contextual nature of qualitative studies I that this study has some limitations. The interviews that were conducted were meant for a qualitative study and taken from two teachers who teach the same subject within the same district. Although the exam is the same across the state their interpretation of its effectiveness as well as their development of the curriculum contains much

contextual bias. Accordingly, the exams studied are not indicative of all DBQ's on other Regents exams or others that are presented across the nation. The focus is on a relatively narrow field that includes only DBQ's and their constructed response questions on the NYS Global History and Geography Regents. To preserve the stability or 'dependability' of the data the researcher utilized overlapping methods of data collection. As stated previously the triangulation of data collection or overlapping of methods diminishes the weaknesses of some evaluative processes by relying on the strengths of many. Also, an audit trail of all DBQ's used and the subsequent evaluation was recorded and kept in their numerous forms of progress. To ascertain the neutrality or confirmability the researcher and St. John Fisher ensure that the study will be made accessible to other researchers so that they may access the same materials to perform similar or further evaluations of the content. Additionally, all questions used in interviews or by the researcher to guide their study will be presented within context so that underlying bias or assumptions are revealed.

Informed Consent and Protecting the Rights of Participants

Before exchanging emails with the participants, I asked for their informed consent with an attached document that summarized my study, how it would be used, and how the participant would be protected through the use of pseudonyms. Both of the participants acknowledged the consent form with their signature.

Data Collection

In an effort to uphold the credibility of my study I collected four forms of data that were triangulated and synthesized in the discussion and findings section of this paper. The derivatives of the majority of the data came from the June issued DBQ's within the Global History and

Geography Regents. These DBQ's were obtained through the NYSED website and were located in the test archives section. After downloading and printing the DBQ's, I asked the participants, that were described earlier, to answer a set of questions I typed into a Microsoft Word document and forwarded to them through our district email service. The participants typed their responses into the document, saved their additions, attached the modified document to an email and sent them back to me. Neither of the participants asked any clarifying questions, and they both answered all ten questions. Their answers to the questions are also found in the appendix of this paper.

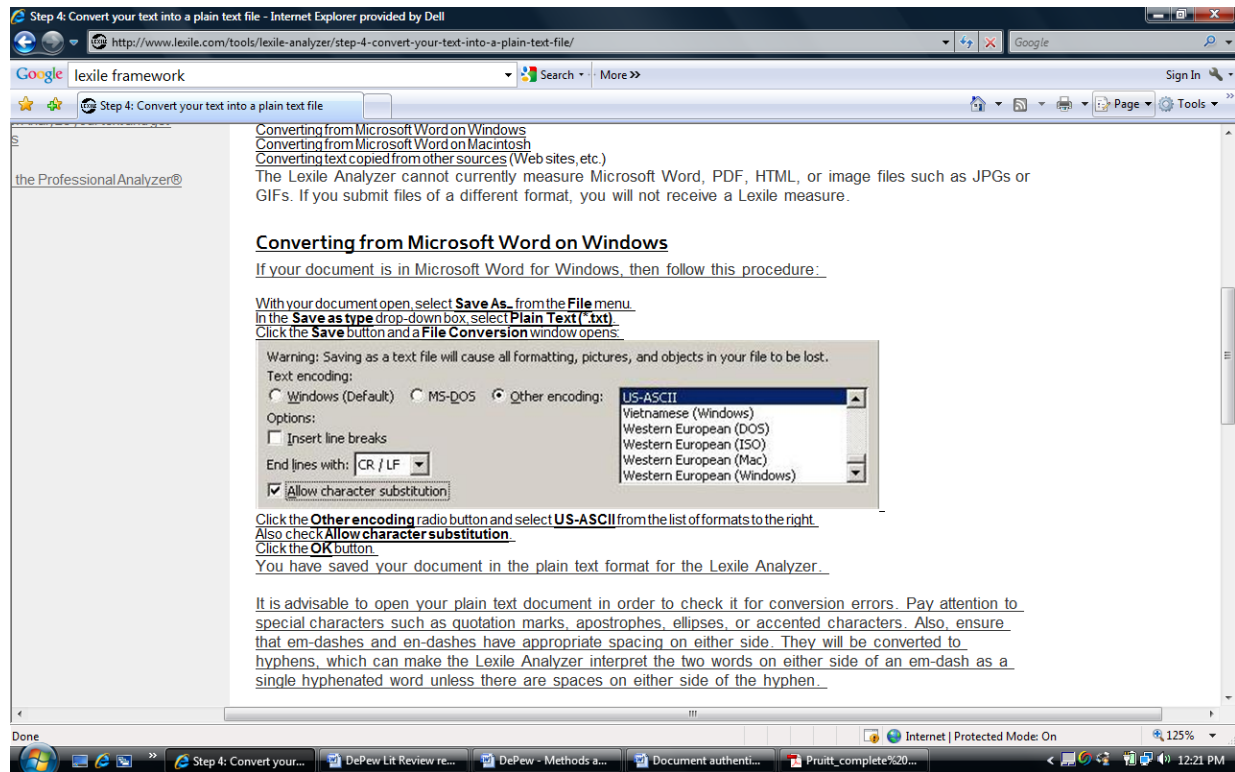
Data Analysis

From here, I separated the DBQ's by years and focused first on the task and constructed response questions of each DBQ. Utilizing a chart based on Blooms Taxonomy levels of questioning I sorted the questions within the DBQ in accordance to the level of questioning and thinking that was necessary to answer the question. Patterns across the five years of data emerged quickly that demonstrated the lower level questioning featured on the DBQ. After sorting all of the questions into Blooms Taxonomy levels I created a second chart based off of Grant et al.'s (2004) work that analyzed the authenticity of the tasks and documents within the DBQ. The chart was composed of four headings including: What type of document was it (this includes whether it was a primary or secondary source)? Was the source given and has the document been adapted? Could the document readily be found for further investigation or inquiry? And, do the documents as a whole offer differing perspectives of the issue? Every document from the last five years of the June issue DBQ's were researched using Google. This type of search was meant to mirror the resources and research that students may be granted if given the opportunity to further investigate the documents and topics presented in the exam. If,

where, and in what form the document was found was indicated within the chart and can be viewed in the appendix of this paper.

To measure the readability levels of the documents I employed three different readability formulas that are endorsed by the Common Core Standards and are readily available online for free. Flesch Kincaid was the easiest readability formula to locate and use as it is installed within Microsoft Word and coincides with the spelling and grammar check function of the software. The software presents the results of its analysis in the form of a grade level score, so it is easily compared to other readability formulas. After running each narrative type document through Flesch Kincaid I abbreviated the document to less than 200 words so that it could be analyzed using the Dale Chall formula found at:

<http://www.lefthandlogic.com/htmldocs/tools/okapi/okapi.php>. When using the online software I had to change the formula settings to Dale Chall and after pasting the document into the processor I had to edit the document to space out words that had been lumped together and eliminate punctuations such as parentheses, colons and semi-colons. Once the software completed its analysis the results were shown as a raw score and grade level equivalent. Within the chart I created for comparing the readability formula results I included the raw score as well as the grade level equivalent, so that the results could be verified through a separate analysis. For a third readability formula I chose the Lexile Framework for Reading. As stated previously the Lexile Framework employs MetaMetrics to analyze word frequency and sentence length to determine a Lexile score that can then be converted into a grade level equivalent. There are however, a few steps that must be followed for the Lexile Analyzer to work properly. The screen below illustrates the steps required to save the text file as a plain text, which once completed several times seems quite simple.



After the document has been converted to a plain text you can upload the document to the Lexile Analyzer through the browse function. Conclusions of the analysis are presented in terms of a Lexile measure, a word count, and mean sentence length. The Lexile measure can then be converted into a grade level equivalent using the following Table:

Table 1

Typical Text Measures, by Grade

Grade	Text Demand Study 2009 25th percentile to 75th percentile (IQR)	"Stretch" Text Measures 25th percentile to 75th percentile (IQR)
1	230L to 420L	220L to 500L
2	450L to 570L	450L to 620L
3	600L to 730L	550L to 790L
4	640L to 780L	770L to 910L
5	730L to 850L	860L to 980L
6	860L to 920L	950L to 1040L
7	880L to 960L	1000L to 1090L
8	900L to 1010L	1040L to 1160L
9	960L to 1110L	1080L to 1230L
10	920L to 1120L	1110L to 1310L
11 and 12	1070L to 1220L	1210L to 1360L

The results of all three readability formulas for each document are illustrated in the appendix of this paper and are broken down by the year in which they were featured on the DBQ.

Surprisingly, the results from each readability formula varied greatly, for this reason I calculated the lowest and highest possible grade level average and included the results within the chart.

Despite the variance amongst the results of the three readability formulas there were some consistencies that mirrored the results of Burke and Greenburg (2010). True to Burke and Greenburg's (2010) findings the Flesch Kincaid continuously rated the texts much lower than both the Dale Chall and Lexile Framework. Interestingly, the Flesch Kincaid measurement was almost always at grade level or below the 10th grade students taking the exam which leads one to

postulate that this maybe the formula that NYSED utilizes to evaluate the level of texts included on the DBQ.

From all of the evidence discerned from the data there arose three overarching themes that addressed the authenticity of the DBQ. First, the presentation of the documents and lack of available resources greatly limits the ability of students to synthesize an objective argument. Second, according to the analysis of the questions featured on the DBQ students are not required to perform high level thinking skills or synthesis nearly as often as they are forced to answer lower level questions based on comprehension. Finally, despite the emphasis on comprehension the majority of narrative style documents were found to be at a readability grade level equivalent that was higher than the students taking the exam. Accordingly, the findings of the data are organized as follows: Authenticity of the documents, Authenticity of the task and constructed response questions, and the readability levels of the documents.

Discussion and Findings

In the analysis of the data three features of authenticity were consistently misrepresented or neglected on the Global History and Geography Regents DBQ. By framing the analysis of the documents after Grant et al. (2004) and Williams (2003) I concluded that the presentation of the documents in regards to the balance of primary source and secondary source texts, the ability to validate sources, and the inclusion of various perspectives amongst the texts were all found to be disingenuous as part of an authentic task. The questions that correlated with the documents in accordance to Blooms Taxonomy Levels also neglected to require synthesis and evaluative level thinking, which are essential to real world problem solving. In addition, the readability levels of the documents were most commonly above grade level making it extremely difficult for students to comprehend the text or synthesize its contents. As a result the discussion and findings below

follow this progression: Authenticity of the Documents, Authenticity of the Task and Constructed Response Questions and Readability Levels of the Documents.

Authenticity of the Documents

As established previously by Grant et al. (2004) the nature in which documents are presented to a researcher must follow a particular criterion in order for the information to be perceived as unbiased and therefore valid. Similar to the framework utilized by Grant et al. (2004), when assessing the documents within the Global History and Geography Regents DBQ the following was considered:

- Determination as to whether it was a primary or secondary source of information
- Was the source of the document provided
- Was there an indication to the student as to whether or not the document was adapted
- Could the document readily be found by students if given the opportunity
- Was there a variety of documents given that contained differing viewpoints.

This framework for analysis was established so that the researcher could validate the authenticity of the document and was evaluated separately from the constructed response questions that accompany the documents on the test.

The results of the investigations, as expected, varied from document to document and from test to test. On the whole however, nearly all documents that were included in the June Geography History and Geography Regents DBQ over the last five years indicated the source from which the document was taken and whether or not the document had been adapted. In fact, only one document from the June 2009 exam did not indicate from where the document was derived and in every instance an adaptation occurred it was noted to the reader. What was not explicitly stated to the reader however was neither where the adaptation had occurred nor the rationale. The lack of information in regards to the adaptation of the document would largely be

without consequence if the reader was provided the opportunity to examine the original document. Yet, in an effort to obtain objective results, the Regents exam does not provide students with time to research the sources of the documents. Indeed, even test administrators cannot verify the source's validity, as they too are given only the adapted version of the document. If the board of Regents decides to alter this policy as the state of Michigan (Grant et al., 2004) has then they will also need to select sources that can be more readily found through resources available to every student. As illustrated by Table 1 below, only 50% of documents that were adapted could be readily located online, arguably the most prominent and equal avenue of research for students in the digital age. In an attempt to locate all of the documents from the DBQ's, I was only able to find 25 of 51 online, approximately 49%. Again, this is of little dispute with the current testing format as this privilege is not afforded test takers in New York State.

The unavailability of these documents though does allow one to question the validity of the source. Without the ability to view the document in its original context, the reader is forced to trust the test creator that the objectivity of the document has been preserved. This level of trust is rarely found amongst historians, as illustrated within the literature review (Williams, 2003), or students who have been taught to be critical thinkers. To reiterate, for research to be considered an authentic task it must be carried out in the manner of historians, whose skills the students are taught to replicate as indicated by NYSED standards.

Table 2

Adapted Documents and Their Availability Online

Year DBQ was Administered	Documents Adapted	Documents Found Online in Original Form
2007	3	2 (66%)
2008	9	4 (44%)
2009	2	0 (0%)
2010	2	1 (50%)
2011	4	2 (50%)
Totals:	20	10 (50%)

Balancing the sources of the documents was also deemed critical to establishing the DBQ as a valid assessment (Grant et. al, 2004). As demonstrated by Table 2, the Regents exam over the last five years has maintained nearly a 50/50 split between primary and secondary sources. However, students only take one these exams and as illustrated in Table 2 this balance usually favored more secondary sources, particularly in 2009. Upon closer examination, a pattern emerged correlating secondary source documents with material prior to the year 1900 and primary source documents with more contemporary material. Not surprisingly then this pattern was exemplified in the June 2009 Regents, which had students examine societal and economic changes from the Middle Ages, the Industrial Revolution in England, and the Age of Globalization. As two of the ages took place prior to 1900, 8 of the 11 documents included in the DBQ contained content from before the year 1900, six of which were secondary sources. The rationale behind this disparity is not stated explicitly in Regents publications, but a simple reason could be that it is easier to locate secondary sources that were written contemporaneously in English as opposed to primary source documents that maybe hundreds of years old and in

need of translation. While most research relies upon a certain amount of material from secondary sources authentic research revolves around the interpretation of primary sources. Students taking the exam are therefore limited in their capacity to develop an argument or standpoint from analysis of primary materials and are instead subjected to reiterating the standpoints and information that have been gathered by others. These sentiments were echoed by participants Erie and Tom, both of whom witness their students simply copying passages from documents. Tom believes that the state allows students to perpetually answer questions in this manner because copying is all that is asked of them. There is, in Tom's words, "no thinking involved" (cite). This issue is compounded exponentially by not allowing students to validate the sources of any of the documents as stated previously.

Table 3

Number of Primary Source and Secondary Source Documents within the DBQ

Year DBQ was Administered	Primary Source Documents	Secondary Source Documents
2007	4 (44%)	5 (56%)
2008	6 (60%)	4 (40%)
2009	3 (27%)	8 (72%)
2010	5 (46%)	6 (54%)
2011	4 (40%)	6 (60%)
Totals:	22	29 (56.8%)

Given the limitations of a research essay within a timed high stakes assessment it makes sense to tailor the documents and their content to the overall task of the DBQ. With minimal variance, the DBQ follows a simple formula that presents three distinct time periods or areas of

study with three to five documents accompanying each. In an effort to provide students with choice, the test taker is instructed to address two of the three topics in their writing, which translates to 6-8 documents. Consequently, by allowing students three choices the test makers exacerbate the brevity of knowledge that the test taker could ascertain from the documents on the given topic. Conversely, this readily translates into a 5-6 paragraph essay that can be produced within the time constraints of the exam. These conflicting standpoints once again call into question the validity of the DBQ as an authentic task. Most professional researchers would argue that three sources were insufficient for the creation of a strong argument or as a basis of a broad topic. Nevertheless, students must develop their topic with what is provided and their schema.

Authenticity of the Constructed Response Questions and Task

In the words of Williams (2003), “historical research is a process of discovery and construction... Historians construct their story on the basis of evidence by selecting and arranging the facts” to “develop a persuasive argument” (p.11). Contrarily, the evidence discerned from the analysis of the documents proves it would be difficult to construct a formative argument or thesis given the potentially bias presentation of the documents. Thankfully, the Global History and Geography Regents DBQ does not require students to formulate and validate a thesis derived from evidence gathered from the documents. In an effort to discern the level of questions, and therefore answers required, contained within the DBQ, I compared the wording of the questions to Bloom’s Taxonomy levels. The results from each year and each document are found in the appendix of this paper. As illustrated within the chart there are six increasing levels of thought and questioning according to Blooms Taxonomy. Given in increasing order the chart flows as follows: Knowledge (recall data or information),

Comprehension (understanding the meaning, restate in own words), Application (use of a concept in a new way, applies what was learned), Analysis (distinguishes between facts and inferences), Synthesis (creates a whole from assembled parts resulting in new meaning), and Evaluation (make judgments based on evidence). The lowest two levels of the chart, knowledge and comprehension, require the student to simply recall data or reiterate what is already known. In accordance to Table 3 these two categories account for over 50% of the questions included in the DBQ over the last five years. Using simple inference skills students should be able to answer correctly another 34% of questions found within the DBQ. Students are asked less frequently to simply state the facts, knowledge (15%), or synthesize the information into a coherent whole (12%). Never are the students required to evaluate the documents to determine bias, or create an argument, which are the cornerstones of a historian's work.

Table 4

Correlation of Constructed Response Questions and DBQ Task with Blooms Taxonomy Levels

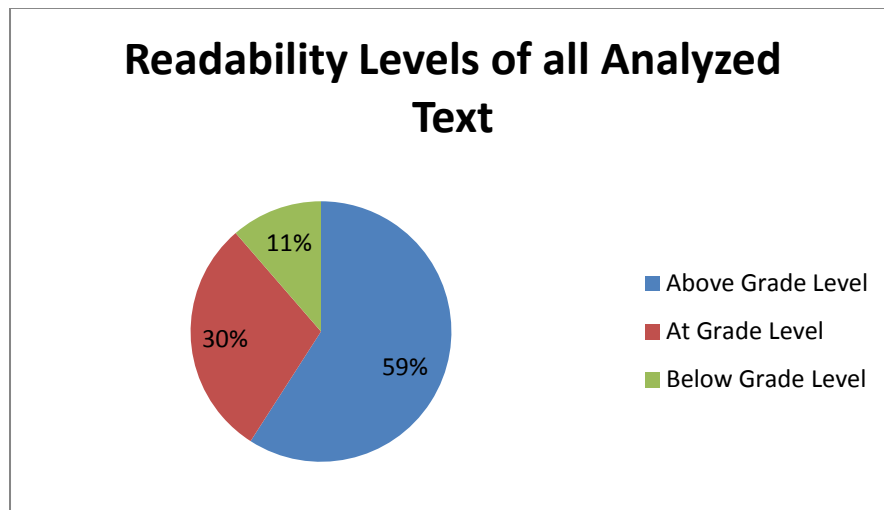
Bloom's Taxonomy Level	Year DBQ was Administered					Totals
	2007	2008	2009	2010	2011	
Knowledge	0	5	3	2	0	10 (15%)
Comprehension	3	2	6	6	7	24 (37%)
Application	0	1	0	0	0	1 (1.5%)
Analysis	9	4	4	3	2	22 (34%)
Synthesis	0	1	1	3	3	8 (12%)
Totals:	12	13	14	14	12	65

These findings mirrored the concerns of the interviewed teachers Erie and Tom. When asked if she believed the Regents DBQ was an authentic task she replied, “No. The kids just copy the documents and don’t really understand what they are writing. It tests their copying skills not their comprehension.” Given the simplicity of the questions included on the DBQ some may argue that all they have to do is copy. Again, this invalidates the DBQ in its current form as an authentic assessment that should mirror “the extent to which students experience questions and tasks under constraints as they typically and ‘naturally’ occur, with access to the tools that are usually available for solving such problems” (Wiggins, 1993, p. 214). Rarely if ever, will students be able to solve real world problems simply by copying the work of others. Additionally, Erie has inadvertently hinted at a much larger are of concern, students’ ability to comprehend the level of text within the documents.

Readability Levels of the Documents

Given the plethora of questions at the comprehension level that are correlated with the documents, the ability to comprehend the text is of utmost importance. Congruently, the work of Johns (2008) dictates that reading materials administered to a student should be within their independent and instructional levels for growth and success. Yet, as illustrated in Figure 1, the readability formulas’ consistently rated the documents to read at levels greater than the levels of the students taking the exam. In all, 44 documents were analyzed, 26 of which were found to be at an average readability level greater than grade 10, 13 were considered at grade level, 9 or 10, and only 5 were written at a grade level lower than 9. These results are briefly summated in table 4 and can be found in their entirety in the appendix of this paper.

Figure 1

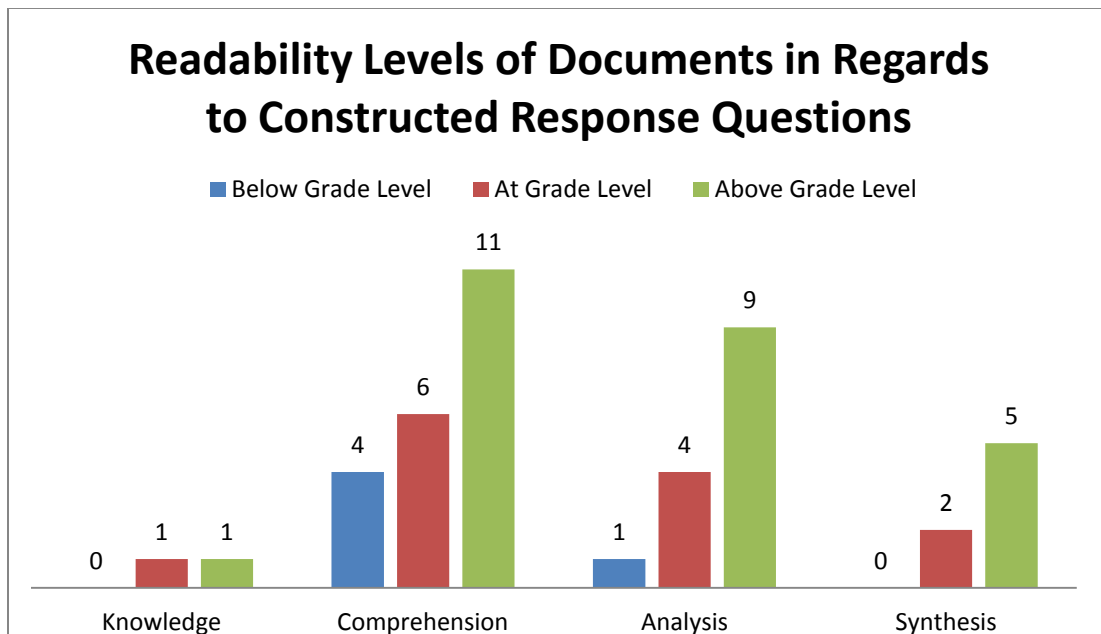


Even without the aid of readability formulas Tom believed that, “Most (of the documents) are too high for actual understanding.” As a result, “students use coping tools to get through. For, (sic) example, they look for the word in the document that matches the word in the question and write whatever follows that word (for their answer).” Erie, also stated that she “think(s) most of my kids can read them (the documents), but this doesn’t mean they understand them.” In summation, students are able to answer questions about passages that they cannot understand because the questions require little more than copying skills to answer. These findings are presented in short by Figure 2. Contradictorily, in accordance to the presented data it could be argued that an almost equal number of questions were asked at the analysis level. However, when put into context these questions required simple inference abilities on behalf of the student. Although the answer was not stated explicitly for these questions, it could be easily filtered from the text assuming the student was able to comprehend the text’s content. The assumption that the student can interpret the text though is also problematic given that 9 of the questions asked at the analysis level were correlated with documents that read at grade levels greater than 10th grade. These contradictions between the levels of the documents and their

accompanying questions lead one to wonder: Why are the majority of questions within the DBQ seeking basic comprehension, if the majority of the documents are written at exceedingly difficult levels for the student population? What is this supposed to prove? Does this make students seem more intelligent because they are capable of answering a question about a document that is written at the college level? Is it because everyone has to take the exam, so the questions are easier to interpret? Are the responses of the students easier to score because there is little room for interpretation? Is the exam more objective because the students cannot draw on their schema to synthesize the new information? What is the rationale? Unfortunately the state offers no explanation in this regard and these are questions that deserve answers.

To make matters worse, after being inundated with documents at their frustrational level students are thrown off by even easily deciphered documents and questions that Tom notes, “are so simple that the students doubt themselves and fumble with the answer.” Here again arises the problematic context in which the Regents is administered. Tom has incidentally noted the emphasis of the individual when portraying his students’ struggles. If students were granted the resources that they customarily utilize and have been taught to employ, students would be able to break down complicated text and discuss their doubts about easier questions. The conclusions of the whole would provide more perspectives for synthesis, evaluation and argument, the hallmarks of historical studies.

Figure 2



As discussed previously, maintaining a balance between primary source and secondary source documents is crucial to creating an authentic task. In this regard the DBQ over the last five years managed to uphold a relative balance with some inconsistencies amongst the years. However, when analyzing the readability levels of the text it became alarmingly clear that a disproportionate amount of primary source documents read well above grade level, as illustrated by Table 5.

Table 5			
<i>Collating Documents' Readability Levels, Blooms Taxonomy Levels of Questioning and Source Type</i>			
Blooms Taxonomy Levels of Questioning	Documents at Readability Level	Primary Source Documents	Secondary Source Documents
Above Grade Level Readability			
Knowledge	1 (4.5%)	1 (100%)	0
Comprehension	11 (50%)	7 (63.6%)	4 (36.36%)
Analysis	9 (40.9%)	4 (44.44%)	5 (55.55%)
Synthesis	1 (4.5%)	0	1 (100%)
At Grade Level Readability			
Knowledge	1 (7.6%)	0	1 (100%)
Comprehension	6 (46.1%)	1 (17%)	5 (83%)
Analysis	4 (30.7%)	2 (50%)	2 (50%)
Synthesis	2 (15.3%)	0	2 (100%)
Below Grade Level Readability			
Knowledge	0	0	0
Comprehension	4 (80%)	2 (50%)	2 (50%)
Analysis	1 (20%)	0	1 (100%)
Synthesis	0	0	0

While a majority of secondary source documents were also written above grade level in terms of readability there was a more favorable balance between above grade level, at grade level and below grade level texts than what was featured amongst primary sources. Indeed, 69% or 11 of 16 primary source documents read at levels above the 10th grade. To reiterate, primary source documents are integral to developing an argument as they inherently must be analyzed, synthesized, and evaluated to determine authenticity. Hindering the analysis of the document however, is the level at which the text is written in comparison to the abilities of the students. In order for students to fully comprehend a text, without the aid of a teacher, peers, or other resources, the materials should be at the students' independent level, as according to Johns (2008) even texts at the instructional level should be scaffolded and mediated by an instructor. Despite this insight, and the crucial role of primary source documents in maintaining authenticity, the primary texts featured on the Global History and Geography Regents DBQ are continuously written above grade level, which for many correlates to the frustrational level. Essentially, when provided the opportunity to engage with primary source materials students are forced to decipher text that they may not be able to comprehend and therefore cannot synthesize the material when responding to the task of the DBQ.

Table 5
Source Type of Documents in Comparison to Readability Levels

Documents Readability Level	Primary Source Documents	Secondary Source Documents
Below Grade Level (9 th Grade and Lower)	2	3
At Grade Level (10 th Grade)	3	8
Above Grade Level (11 th Grade or Higher)	11	10

Again though, the questions of the test appear to indicate that the creators of the exam recognize the difficult level of the primary source texts. When comparing primary source texts that were judged above grade level to the questions associated with them it was found that 63.6% of questions required only comprehension level thinking (see Table 7 in the appendix). This is not to say that answering questions about the main idea of a text can be easily done when the document may be well above the student's reading level. However, as stated previously it does hint at the fact that the creators of the exam may well be aware of the complexities of the text included within the DBQ.

To reiterate, the authenticity of the documents in regards to their presentation, perspectives, and balance of sources was jeopardized as a result of a compilation of factors. As a whole, sources of the documents were noted by the Regents exam but in large part these sources could not be validated through the most common source of material, the internet. Furthermore, the format of the exam disallowed a balanced presentation of perspectives due to constraints on the number of documents pertaining to each topic within the DBQ. The inauthentic nature of the documents was compounded by an imbalance between primary source and secondary source

materials from year to year. Additionally, the questions accompanying each document required little analysis on behalf of the student and never encouraged the formation of an argument based on evidence gathered from the documents. Finally, the readability levels of the texts were consistently above grade level and accompanied by questions that required the students to be able to comprehend the material. These three factors, the inauthentic nature of the documents, the low level questioning and the emphasis on frustrational level text, combine to create a testing context that differs greatly from how students learn, problem solve, interact with text and synthesize information in the classroom and in the ‘real world.’

Implications

The current format of the Global History and Geography Regents DBQ fails to adhere to the Social Studies standards established by NYSED as an authentic task that requires students to “ask and answer analytical questions, take a skeptical attitude toward questionable arguments, acquire and organize information, evaluate data, draw conclusions, and view the human condition from a variety of perspectives” (NYSED, 2011). Congruently, analysis of the data consistently demonstrated that the task of the DBQ does not require students to compose an argument, ascertain bias, in-depthly evaluate data or view a topic through a multitude of perspectives. Though educators are required by the standards to instill these intellectual skills within their students, the Regents exam emphasizes lower level thinking skills intermixed with documents that are more often than not too difficult for students to decode within a limited time frame and without resources that are integral to their synthesizing processes. As illustrated through the examination of the documents taken from the Regents Global History and Geography DBQ and discussion with teachers from the field, the documents are written at levels that are incomprehensible for the average student in the 10th grade. Because of students’

inability to analyze the documents the DBQ inherently loses a strong component of construct validity. The construct validity of the exam is further jeopardized by the format of the exam and the manner in which it is given. By following a cookie cutter formula that promotes objectivity across the dynamic student body taking the exam the test creators inhibit the inquisitive nature of social studies and history in particular. Conclusions from the data indicate that teachers must ensure that their students engage frequently with challenging levels of text and that their students are equipped with the tools and skills necessary to analyze such materials independently.

Teachers then, must also have the knowledge and tools to assess the readability levels of text, so that they can better scaffold the growth of their students. Additionally, in an era of reform and changing standards NYSED would be well advised to consider changing the format and context in which students are assessed, particularly in regards to high stakes testing and inauthentic assessments such as the Global History and Geography DBQ. Given the success of authentic assessment in classrooms across the nation I would advocate strongly for a portfolio type of assessment that demonstrates growth, knowledge and skills over a period of time and across multiple mediums.

There is no doubt that change is eminent as New York State has begun its implementation of the Common Core, a national standard for the core subjects that emphasizes literacy growth and integration in every subject. To date, there have not been any indications as to how this will affect the format of NYS Regents exams or if they will be replaced with a national assessment. However, there are clear expectations in regards to students' literacy development particularly to reading standards which:

“place equal emphasis on the sophistication of what students read and the skill with which they read. Standard 10 defines a grade-by-grade “staircase” of increasing text complexity that rises from beginning reading to the college and career readiness level. Whatever they are reading, students must also show a steadily growing ability to discern more from and make fuller

use of text, including making an increasing number of connections among ideas and between texts, considering a wider range of textual evidence, and becoming more sensitive to inconsistencies, ambiguities, and poor reasoning in texts.” (CCS, 2011, p.8).

More specifically, students are expected to “read and comprehend complex literary and informational texts independently and proficiently” (CCS, 2011, p. 10). In accordance to the texts featured on the Global History and Geography DBQ this emphasis should not be anything new to Global History teachers. What is new is the implementation of this ideal across subjects and at earlier stages of students’ schooling. There are inherent complications however with this quick implementation. High school and even middle school students who have not had exposure to difficult or challenging expository texts for years are suddenly expected to analyze and interpret these materials on high stakes tests without the schema and scaffolding that the next generation of students will benefit from. This disparity in preparedness is already being addressed by Erie who states that “I try to increase reading comprehension and writing by working with the English department a lot, especially now that I am looping with an English teacher.” The marriage of social studies and English came implicitly to Erie who believes that “social studies skills are mostly the same as ELA (English Language Arts) skills.” Erie’s belief coincides precisely with the Common Core’s standards that stress literacy development as a cross curricular activity. In summation, New York State’s endorsement of the Common Core reiterates the emphasis on challenging text as a means of assessment and therefore will likely continue to be included within the Global History and Geography DBQ. Teachers, such as Erie, are already instituting curricular modifications such as co-curricular instruction, to better prepare their students for the demands of high stakes assessment and as the Common Core stated, “career and college readiness” (CCS, 2011, p.8).

As discussed, the use of complex texts cannot be expected to diminish over the ensuing years; for this reason teachers must be able to judge what exactly a complex text is and at what level are their students reading at or accustomed to. Within the literature review and data collection sections of this paper I illustrated the use of three readability formulas and how they can be used to assess the difficulty of a particular text. In congruence with the findings of Burke and Greenburg (2010), my data indicated that each readability test interpreted the document at differing levels. To recap, the Flesch Kincaid typically rates text passages two grades lower than other readability formulas, while the Dale Chall typically swings the results in the other direction stating the document to be approximately two grades higher. The most balanced of the formulas was the Lexile Framework which was not stated in Burke and Greenburg's (2010) analysis but could be concluded when the average of three formulas was compared to the results of the Lexile Framework. All of the formulas are easily accessible, the Lexile Framework requires a username and password but is free, should be used in conjunction with another and employed in replace of an eye test. To better gauge the level of difficulty expected under the new Common Core Standards (CCS) I would recommend assessing the readability levels of the exemplar texts included in the appendix of the CCS. With this knowledge, teachers can mirror the expectations of the CCS while maintaining autonomy within their classroom by using their own resources. If possible, teachers should also try to differentiate the implementation of challenging text to meet the levels of individual students. As advocated by Johns (2008) the pairing of students with instructional level materials best enables growth. There are obvious difficulties in this regard however, as only students who are perceived to be exceptionally low functioning readers are assessed in the middle and high school years. Thus, it may be extremely difficult to know at what levels all of your students are reading. To counter these challenges teachers should receive

continuous professional development to discover innovative and research based methodologies for introducing challenging texts with a diverse student body.

In spite of the negative consequences such as increased dropout rates, and a widening gap between the success of whites and minorities, high stakes assessments continue to be the norm across the nation as it allies to the standards of NCLB. Despite the implementation of the Common Core little has been stated about its effects on the use of standardized tests. Even the idea of a national world history assessment has not been extensively explored; however, Bain and Shreiner's (2005) postulations revealed many inherent difficulties in creating a new assessment or protocol for assessing a world history course. One of the more perplexing issues for the researchers was the incongruent nature of the world history curriculum amongst the states of the nation (Bain & Shreiner). While the Common Core establishes a national standard, Bain and Shreiner (2005) worry that these standards and the ensuing assessment will greatly influence what is and is not taught. As demonstrated by nearly all courses that are concluded with a high stakes exam Bain and Shreiner reiterate the unfortunate truth that, "If we test it, they will teach it" (2005, p. 242).

The influence the test has on the creation of teachers' curriculums cannot be overstated and repeatedly arose when talking to participants Erie and Tom. When asked what her concerns were about the Global History and Geography Regents Erie replied "The kids are tested on such a wide range of topics. As a teacher, I feel like I have to cover everything with little depth so that I can try to get as much knowledge into the kids as possible." Erie's concerns were echoed by Tom who stated "Global History is a huge course. No one, not even me (I've been teaching the course for 10 years), knows all of global history. Given this, what is it that the students are supposed to learn?" It is clear that both Erie and Tom are overwhelmed by the amount of

material that needs to be covered. This sentiment was reiterated later in the discussion as well when I asked each participant if they find themselves planning for the test when developing their curriculum. Both of the participants replied reluctantly that the test had a large impact. Erie stated that she teaches to the “test about 75% of the time,” and that “I HATE this!” Similarly, Tom has “discarded virtually all of my(his) projects in the last ten years” as “they take too long to do well and still be able to cover the material that the Regents require.” The constraining influence the Regents exam exemplified through Erie and Tom’s words is a small example of a much larger problem: high stakes assessment such as the Global History and Geography Regents encourage teaching to the test, promote testing anxiety and create doubt amongst teachers as well as students. To combat this practice Bain and Shreiner advocate an assessment that evaluates a cross section of content approaches that allow schools to construct the framework of their curriculum. For this type of assessment to be instituted the exam must emphasize big ideas and require students to demonstrate higher order thinking skills as outlined in the standards.

Given the inauthentic nature of the DBQ and its disconnect from NYSED standards, evidence from the research of literature, discussions with the participants and the findings of this study I propose a movement towards more authentic assessment. Carmichael, King, and Newman (2009) believe that authentic intellectual work “can serve as guidelines for curriculum, instruction and assessment that extend beyond the basics, and beyond extensive lists of content standards” (p. 43). Indeed, Carmichael et al. (2009) stress that authentic intellectual work that involves “construction of knowledge, through the use of disciplined inquiry, to produce discourse, products or performances that have value beyond school...provide a framework for teaching and assessing any goal that relies on knowledge from an academic or applied discipline” (p.43, 47). In other words, authentic intellectual work should be the basis on which

the foundation of the curriculum is built, not a small part that is used sparingly and when educators ‘have time.’ True, it will take time, even years, to fully recreate a curriculum with these learning goals in mind, but as educators we have a duty to develop our students’ ability to learn not just what they learn.

Conclusions

In order to maintain the confirmability and dependability of this study as defined in the methods section of this paper it is important to note the limitations of this study. This study was conducted with a small sample set, just DBQ’s from one subject across a five year time span. Because of this light sample size the results of this study should not be transferred to other exams or elements of the test without further research. Additionally, the analysis of the constructed response questions and task through the use of Blooms Taxonomy Levels is an imperfect practice that provides for varying interpretations that may be influence by my familiarity with the material, style of questioning and expectations of students. In regards to the readability formulas utilized, it should be stated that I employed only three because of their accessibility but that this is not an exhaustive list and research employing other readability formulas would provide further insight. Also, the participants described within this study both teach at the same school and though they differ in age their opinions should not be reflected on all Global History teachers. In the end, this was a qualitative study that drew its results from a small sample pool. Therefore, the validity of this study would benefit from further inquiry into the Global History and Geography Regents DBQ as well as other DBQs featured on Regents exams within New York State.

As any student of history will note, research brings more questions than answers. Despite, the conclusions gathered from the data I still have several questions regarding the construction of the DBQ and how it is perceived by NYSED. As, the participants Tom and Eric can attest, I am not the only teacher who has noticed the inauthentic nature of the DBQ and the contextual nature of high stakes testing. Yet, NYSED continues to endorse, through its emphasis on scores, assessments that contradict its own standards. Is it simply for federal funds that the state perpetuates this cycle? Are they currently evaluating the effectiveness of alternate assessments? Will the Common Core bring the changes that teachers across the state so desperately desire and our students so desperately need? If the best teaching methods involve authentic intellectual work why are students not assessed in a similar manner? Why are all students held to the same levels when emphasize differentiated learning within our classrooms? These are questions that may not yet have answers, however, if teachers are to prepare all students for career and college readiness they deserve some inquiry.

In an era of high stakes standardized assessment the DBQ featured on the Global History and Geography Regents fits the mold. Though created to reflect the authentic nature of a historian's work the context in which the DBQ is administered stifles nearly all of the aspects that make it valid. More than ever, students' literacies are "necessarily social, "situationally specific" and a "multimodal, multimedial, dynamically changeable enterprise" (Greenhow et al., 2009, p. 250) and yet the DBQ is taken by individual students with no outside resources other than their schema. Additionally, the documents within the exam jeopardize their authenticity by appearing in truncated passages with little or no context and no opportunities for the students to validate the sources or ascertain bias. Furthermore, the documents, particularly primary source documents, the crux of a historian's research, read at a level that mirrors the demands of college

students not 10th grade students. Coincidentally, the majority of documents that are written at readability levels above grade level are paired with questions that are based on student comprehension; a contradiction that seemingly sets students up for failure. To conclude, the DBQ as presented within the Global History and Geography regents does not reflect the inquisitive and collaborative nature of authentic learning as mandated by New York State's Department of Education standards.

References

- Bain, R. B., & Schreiner, T. L. (2005). Issues and options in creating a national assessment for World History. *The History Teacher*, 38(2), 242-271. Retrieved at <http://web.ebscohost.com.pluma.sjfc.edu/ehost/detail?vid=3&hid=112&sid=203745e6-e2ab-4d64-a281-12cd88c1d498%40sessionmgr110&bdata=JnNpdGU9ZWWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=ehh&AN=16494933>
- Burke, V., & Greenburg, D. (2010). Determining readability: How to select and apply easy-to-use readability formulas to assess the difficulty of adult literacy materials. *Adult Basic Education and Literacy Journal*, 4(1), 34-42. Retrieved at <http://web.ebscohost.com.pluma.sjfc.edu/ehost/detail?vid=5&hid=112&sid=203745e6-e2ab-4d64-a281-12cd88c1d498%40sessionmgr110&bdata=JnNpdGU9ZWWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=ehh&AN=59373837>
- Carmichael, D. L., King, M. B., & Newman, F. M. (2009). Authentic intellectual work: Common standards for teaching social studies. *Social Education*, 73(1), 43. Retrieved at <http://search.proquest.com.pluma.sjfc.edu/docview/210625542?accountid=27700>
- Common Core State Standards Initiative (2011). *Common core state standards for English language arts and literacy in history/social studies, science and technical subjects*. Retrieved from Common core website: http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf
- Dodge, A. (2009). Heuristics and NCLB standardized tests: A convenient lie. *International Journal of Progressive Education*, 5(2), 1-17. Retrieved at <http://web.ebscohost.com.pluma.sjfc.edu/ehost/detail?vid=7&hid=112&sid=203745e6-e2ab-4d64-a281-12cd88c1d498%40sessionmgr110&bdata=JnNpdGU9ZWWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=ehh&AN=51837750>
- Fine, M. (2005). High-stakes testing and lost opportunities the New York State Regents exams. *ENCOUNTER: Education for Meaning and Social Justice*, 18(2), 24-29. Retrieved at <http://web.ebscohost.com.pluma.sjfc.edu/ehost/detail?vid=9&hid=112&sid=203745e6-e2ab-4d64-a281-12cd88c1d498%40sessionmgr110&bdata=JnNpdGU9ZWWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=ehh&AN=18019652>
- Forte, E. (2010). Examining the assumptions underlying the NCLB federal accountability policy on school improvement. *Educational Psychologist*, 45(2), 76-88. doi:10.1080/00461521003704738

- Gee, J. P. (2001). What is Literacy? In E. Cushman, E. R. Kingtgen, B. M. Kroll & R.M. (Eds.), *Literacy: A critical source book* (pp. 537-544). Boston: Bedford/St. Martins
- Gerwin, D. (2004). Preservice teachers report the impact of high-stakes testing. *The Social Studies*, 71-74. Retrieved at <http://web.ebscohost.com.pluma.sjfc.edu/ehost/detail?vid=11&hid=112&sid=203745e6-e2ab-4d64-a281-12cd88c1d498%40sessionmgr110&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=ehh&AN=13379168>
- Goleman, D. (2007). *Social intelligence: The new science of human relationships*. New York: Bantam.
- Grant, S. J., Gradwell, J. M., & Cimbricz, S. K. (2004). A question of authenticity: The document based question as an assessment of students' knowledge of history. *Journal of Curriculum and Supervision*, 19(4), 309-337. Retrieved from <http://web.ebscohost.com/ehost/detail?vid=3&hid=122&sid=6ff1cb4b-e93a-494b-b897-b30e15678fad%40sessionmgr4&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=ehh&AN=13755769>
- Greenhow, C., Robelia, B., & Hughes, J. E. (2009). Web 2.0 and classroom research: What path should we take now? *Educational Researcher*, 38(4), 246-259. doi:10.31102/0013189X09336671
- Guba, E.G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology*, 29 (2), 75-91
- Halliday, M. A. K. (1969). Relevant models of language. *Educational Review*, 22, 26-37. (on Blackboard)
- Haycock, K., Lankford, H., & Olson, L. (2004). The elephant in the living room. *Brookings Papers on Education Policy*, 7, 229-263. Retrieved at <http://www.jstor.org/stable/20067270>
- Horn, C. (2003). High-stakes testing and students: Stopping or perpetuating a cycle of failure? *Theory into Practice*, 42(1), 30-41. Retrieved at <http://www.jstor.org/stable/1477316>
- Hursh, D. (2007). Assessing No Child Left Behind and the rise of neoliberal education policies. *American Education Research Journal*, 44(3), 493-518. doi:10.3102/00028311207306764
- Jenkins, H., Clinton, K., Purushotma, R., Robinson, A.J., & Weigel, M. (2006). *Confronting the challenges of a participatory culture: Media education for the 21st century*. Chicago, IL: The MacArthur Foundation. (on Blackboard)
- Johns, Jerry L. (2008). *Basic reading inventory* (10th edition). Dubuque, IA: Kendall/Hunt Publishing Company.

- Kucer, S. (2005). *Dimensions of Literacy: A conceptual base for teaching reading and writing in school settings*. (2nd Ed.). Mahwas, NJ: Erlbaum.
- Lankshear, C., & Knobel, M. (2003). *New literacies: Changing knowledge and classroom learning*. Buckingham: Open University Press. *Chapter 1* (on Blackboard)
- Larson & Marsh. (2005). *Making Literacy Real theories and practices for learning and teaching*. California: SAGE Publications Ltd.
- Loveless, Tom. (2005). Test-based accountability: The promise and the perils. *Brookings Papers on Education Policy*, 8, 7-45. Retrieved at <http://www.jstor.org/stable/20062554>
- Maestri, M. A. (2006). The myth of a multicultural curriculum: An analysis of New York State U.S. History Regents. *The History Teacher*, 39(3), 381-402. Retrieved at <http://web.ebscohost.com.pluma.sjfc.edu/ehost/detail?vid=14&hid=112&sid=203745e6-e2ab-4d64-a281-12cd88c1d498%40sessionmgr110&bdata=JnNpdGU9ZWWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=ehh&AN=21309847>
- Meyers, C. (2007, November 21). Reading for test scores and for life. [Letter to the editor]. *New York Times*, p. A35
- Mills, Geoffrey E. (2011). *Action research: A guide for the teacher researcher*. Boston, MA: Pearson Education, Inc.
- Olssen, M., Codd, J., & O'Neil, A. M. (2004). *Education policy: Globalization, citizenship and democracy*. Thousand Oaks, CA: Sage.
- O'Toole, J. M., & King, R. A. R. (2011). Reading by the numbers: Reconsidering numerical estimates of reading difficulty. *The International Journal of Learning*, 17(10), 181-194. Retrieved at <http://web.ebscohost.com.pluma.sjfc.edu/ehost/detail?vid=16&hid=112&sid=203745e6-e2ab-4d64-a281-12cd88c1d498%40sessionmgr110&bdata=JnNpdGU9ZWWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=ehh&AN=64439282>
- Pestacore, C. (2008). Current events as empowering literacy: For English and social studies teachers. *Journal of Adolescent & Adult Literacy*, 51(4), 326-339. Retrieved at <http://search.proquest.com.pluma.sjfc.edu/docview/216915621?accountid=27700>
- Reich, G. A., & Bally, D. (2010). Get smart: Facing high-stakes testing together. *The Social Studies*, 101, 179-184. doi:10.1080/00377990903493838
- Robertson, S. (2000). *A class act: Changing teachers' work, the state, and globalization*. London: Falmer.
- Rubin, J. (2008). Judging our students-keeping it just: The theorem of intellectual measure. *Journal of College Admission*. 199, 4-11. Retrieved at <http://web.ebscohost.com.pluma.sjfc.edu/ehost/detail?vid=18&hid=112&sid=203745e6-e2ab-4d64-a281-12cd88c1d498>

%40sessionmgr110&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=ehh&AN=31804452

- Scheidet, R. A. (2003). Improving student achievement by infusing a web-based curriculum into Global History. *Journal of Research on Technology in Education* 36(1), 77-94. Retrieved at <http://web.ebscohost.com.pluma.sjfc.edu/ehost/detail?vid=20&hid=112&sid=203745e6-e2ab-4d64-a281-12cd88c1d498%40sessionmgr110&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=ehh&AN=11951186>
- Sipple, J. W., Killeen, K., & Monk, D. H. (2004). Adoption and adaptation: School district responses to state imposed learning and graduation requirements. *Educational Evaluation and Policy Analysis*, 26(2), 143-168. Retrieved at <http://www.jstor.org/stable/3699557>
- Tabb, W. (2002). *Unequal partners: A primer on globalization*. New York: The New Press
- The University of the State of New York (1996). *Social studies resource guide with core curriculum*. Albany, NY: The New York State Education Department. Retrieved from <http://www.p12.nysed.gov/ciai/socst/pub/ssovervi.pdf>
- The University of the State of New York (2011). *The New York state district report card*. Retrieved from NYSED website: <https://reportcards.nysed.gov/files/2009-10/AOR-2010-043200050000.pdf>
- Toolin, R. E. (2004). Striking a balance between innovation and standards: A study of teachers implementing project-based approaches to teaching science. *Journal of Science Education and Technology*, 13(2), 179-187. Retrieved at <http://www.jstor.org/stable/40188780>
- Ward, Thomas (1995). Curriculum, instruction and assessment preliminary draft framework for social studies. Albany, NY: New York State Education Department, 206-207. Retrieved from <http://www.emsc.nysed.gov/ciai/socst/pub/ssframe.pdf/>
- Wiggins, G. (1993). Assessment: Authenticity, context and validity. *Phi Delta Kappan*, 75, 214.
- Williams, R. C. (2003). *The historian's toolbox: A student's guide to the theory and craft of history*. Armonk, New York: M. E. Sharpe.

Table I				
<i>Validating the Authenticity of the Documents in Regard to their Presentation within the DBQ</i>				
DBQ June 2011	Type of Document	Is the source given? Is it adapted?	Could the document be readily found for further investigation or inquiry?	Underlying bias in the presentation of the documents? Do they portray differing views?
Doc. 1	Secondary Source Passage from a Book	Yes Linda Jacobs Altman, <i>Genocide: The Systematic Killing of a People</i> , Enslow Publishers	Could not be found in full online but could be purchased at the publishers website: http://www.enslow.com/ displayitem.asp?type=1 &item=2491	The documents within this DBQ provide differing perspectives for all three genocides. Also, there is a relatively equal balance of primary and secondary sources. Thus, there does not appear to be an underlying bias across the documents.
Doc. 2	Primary Source Memorandum by USSR against Ukrainians	Yes Soviet Archives Exhibit, Library of Congress (adapted)	Found in full here: http://www.loc.gov/exhi bits/archives/k2grain.ht ml	
Doc. 3	Primary Source Speech given at the unveiling of monument to Ukrainian Genocide	Yes Dr. Oleh W. Gerus, “The Great Ukrainian Famine- Genocide,” Centre for Ukrainian Canadian Studies, University of Manitoba, August 4, 2001 (adapted)	Found in full here: http://umanitoba.ca/cent res/ukrainian_canadian/ newsletter/2001/dauphin _monument.html	
Doc. 4	Secondary Source Passage from a book on Pol Pot’s regime	Yes Ben Kiernan, <i>The Pol Pot Regime: Race, Power, and Genocide in Cambodia under the Khmer Rouge, 1975–79</i> , Yale University	Could not be found in full and preview from book could not locate passage http://books.google.com /books/about/The_Pol_P ot_regime.html?id=Mq8 sAcvg-AgC	

Doc. 5	Secondary Source Website passage	Yes “Genocide in the 20th Century: Pol Pot in Cambodia 1975-1979,” <i>The</i>	Found in full here: “Genocide in the 20th Century: Pol Pot in Cambodia 1975-1979,” <i>The History Place</i>
Doc. 6	Primary Source Eyewitness Testimony	Yes Teeda Butt Mam, “Worms from Our Skin,” Children of Cambodia’s Killing Fields, Yale University Press	Was not found in described chapter at website: http://books.google.com/books?id=FjEpaj1F9VoC&q=worms+from+our+skin#v=snippet&q=worms%20from%20our%20skin&f=false
Doc. 7	Secondary Source Transcript of news telecast “Frontline”	Yes “The Triumph of Evil,” <i>Frontline</i> , January 26, 1999	Found in full here: http://www.pbs.org/wgbh/pages/frontline/shows/evil/etc/script.html
Doc. 8	Unknown	Yes Aimable Twagilimana, Teenage Refugees from Rwanda	Could not be found Google returned results related only to Regents
Doc. 9a	Secondary Source Video	Yes Terry George, ed., Hotel Rwanda, Newmarket Press	Could not be viewed online
Doc. 9b	Primary Source Map	Yes UN High Commissioner for Refugees, December 1994	Could not be found online

Table II				
<i>Validating the Authenticity of the Documents in Regard to their Presentation within the DBQ</i>				
DBQ	Type of Document	Is the source given? Is it adapted?	Could the document be readily found for further investigation or inquiry?	Underlying bias in the presentation of the documents? Do they portray differing views?
June 2010	Secondary Source Passage from book	Yes Robin Hallett, <i>Africa to 1875: A Modern History</i> , The University of Michigan Press, 1970	Found online at: http://books.google.com.pk/books?id=q6AMAAAIAAJ&q=Sahara	Document was not truncated
Doc. 2	Secondary Source Passage from book	Yes Philip Koslow, <i>Ancient Ghana: The Land of Gold</i> , Chelsea House Publishers	Could be purchased online but not previewed	
Doc. 3	Primary Source News Article	Yes Ute Schaeffer, "Deutsche Welle reporters on the ground," <i>Down to Earth: News & Views on Desertification</i> , UNCCD, June 2006, Volume 21	Found in entirety at: http://www.unccd.int/publicinfo/june17/2006/docs/Down_to_Earth-2006UN-eng.pdf	Presented as an excerpt
Doc. 4	Secondary Source Textbook	Yes James I. Clark, <i>India: The Subcontinent: India, Pakistan, and Bangladesh</i> , McDougal, Littell & Company	Could not be previewed but could be purchased Published in 1983	

Doc. 5a	Secondary Source Passage from Book	Yes Pomeranz and Topik, <i>The World That Trade Created</i> , M. E. Sharpe, 1999 (adapted)	Preview found here: http://books.google.com/books?printsec=frontcover&vid=ISBN0765602490&vid=ISBN075602504&vid=ISBN0765602504&vid=ISBN0765602490&vid=ISBN	
Doc. 5b	Secondary Source Map	Yes "The West and the Spice Trade," Calliope, Cobblestone (adapted); Mountain High Maps, Digital Wisdom (adapted)	Google only turned up results that related to the Regents	
Doc. 6a	Primary Source Photograph	Yes Priit J. Vesilind, "Monsoons: Life Breath of Half the World," National Geographic, December 1984, Photograph by Steve McCurry	Found in full at: http://stevemccurry.photoshelter.com/image?&_bqG=14&_bqH=eJzL9Qz3LnIsr7R0Lo4IlogI0DV3CTExTIKdy.2MrSwMjQwAGEg6RnvEuxsm5aTn5.SmZeu5hmvH.Qe7.li6w.STTQsqKisyI90LYtUAYuMd_RzsS1Ri3d0DrEtLS4KTK0sSs5Qc48PdnUMcvaI9_V3cbU1UAMb4A4ywL84zM2owdsIz9wAANzSLAc-&GI_ID=	Description of photograph was added to the photo and is not found with the photo at the website given to the left.
Doc. 6b	Primary Source Magazine Article	Yes Priit J. Vesilind, "Monsoons: Life Breath of Half the World," National Geographic, December 1984	Could not be found online through Google nor National Geographic	

Doc. 7	Secondary Source Website information page	Yes http://country-studies.us/russia/24.htm	Found in entirety at: http://countrystudies.us/russia/24.htm Taken from: Glenn E. Curtis, ed. <i>Russia: A Country Study</i> . Washington: GPO for the Library of Congress, 1996.	Small modification to the internet source code was required to find the information. The passage given is two paragraphs from a lengthy synopsis of Russian climate.
Doc. 8	Primary Source Political Cartoon from 1941	Yes Leslie Gilbert Illingworth, November 10, 1941, Library of Wales, Aberystwyth	Found in full at: http://www.cartoons.ac.uk/browse/cartoon_item/anytext=Illingworth%201941?page=151	Adaptations are not noticeable. Title in DBQ is placed above the cartoon
Doc. 9	Primary Source Report by IPIECA	Yes IPIECA, "Conoco in the Russian Arctic: Preserving delicate Arctic ecology by minimizing the development footprint and environmental impact"	Link to the document: http://www.docstoc.com/docs/20209549/The-Oil-and-Gas-Industry-Operating-in-Sensitive-Environments Must sign in with Facebook address, seemed unnecessary	

Table III				
<i>Validating the Authenticity of the Documents in Regard to their Presentation within the DBQ</i>				
DBQ	Type of Document	Is the source given? Is it adapted?	Could the document be readily found for further investigation or inquiry?	Underlying bias in the presentation of the documents? Do they portray differing views?
June 2009				
Doc. 1	Secondary Source Textbook Excerpt Manorial System	Yes Morris Bishop, The Middle Ages, Houghton Mifflin	Presented in Full Found online at: http://www.houghtonmifflinbooks.com/books/bookpreview.html?isbn=061805703X	Only two paragraphs from the beginning of a chapter but they are not altered
Doc. 2	Secondary Source?	Yes “Legacy of the Crusades,” Aramco World	Could not be found except as a source for another document in the June 2005 exam where it was cited as if from a journal titled “Aramco World” published in 1956	
Doc. 3	Secondary Source Textbook	Yes Frances & Joseph Gies, Cathedral, Forge, and Water Wheel: Technology and Invention in the Middle Ages, Harper Perennial (adapted)	<i>Could not be read online but could be purchased</i> ISBN: 9780062016607; ISBN10: 0062016601; Imprint: HarperCollins e-books ; On Sale: 9/21/2010; Format: eBook; Trimsize: ; Pages: 0; \$11.99; Ages: 18 and Up.	Notice it recommends ages 18 and up...
Doc. 4a	Secondary Source Textbook	Yes Farah and Karls, World History: The Human Experience, Section Focus Transparencies, Glencoe McGraw-Hill	Could not be read online but could be purchased	No reviews available
Doc. 4b	Secondary Source Cannot be	No	Could not be discerned	

Doc. 5	Secondary Source Passage from book	Yes Michael Mandelbaum, The Ideas that Conquered the World, Public Affairs	Can be previewed at: http://books.google.com/books/about/The_ideas_that_conquered_the_world.html?id=F7SC2K_oIGoC	Many reviews given, could not find the exact page of text. Entire title of book was not given in source
Doc. 6a	Primary Source Passage from book	Yes Robert Agnew, M.D., "Observations on the State of the Children in Cotton Mills," Manchester, March 23, 1818	Original could not be located Google search pulled up only that which is linked to the Regents exam	
Doc. 6b	Primary Source Quote from a book written by Engels	Yes Friedrich Engels, The Condition of the Working Class in England, Stanford University Press (adapted)	Exact page could not be found in English version http://books.google.com/books/reader?id=-6CQRN4n2zsC&printsec=frontcover&output=reader&source=gbs_atb	
Doc. 7	Secondary Source Excerpt from Magazine for school age children	Yes Herbert Buchsbaum, "Living in a Global Economy," Scholastic Update, March 7, 1997	Only information that could be located through a Google search was related to the DBQ	
Doc. 8	Primary Source Newspaper Article	Yes Associated Press, Syracuse Herald American, June 24, 2001	According to Google the Syracuse Herald American does not exist	
Doc. 9	Secondary Source Excerpt from book	Yes Joseph E. Stiglitz, Globalization and Its Discontents, W. W. Norton & Co., 2003	Can be previewed online at: http://books.google.com/books/about/Globalization_and_Its_Discontents.html?id=geN6MUthHdkC	Author is a nobel prize winner in economics

Table IV				
<i>Validating the Authenticity of the Documents in Regard to their Presentation within the DBQ</i>				
DBQ	Type of Document	Is the source given? Is it adapted?	Could the document be readily found for further investigation or inquiry?	Underlying bias in the presentation of the documents? Do they portray differing views?
June 2008				
Doc. 1	Primary Source Letter to the Editor of New York Times Written by Raphael Lemkin whom defined Genocide	Yes Raphael Lemkin, New York Times, Nov. 8, 1946 (adapted)	Adapted Appears truncated Unable to locate readily online	From what was found the letter as it appears on the document has been shortened but not paraphrased
Doc. 2a	Primary Source Newspaper Article	Yes Irina Lagunina, "World: What Constitutes Genocide Under International Law, and How are Prosecutions Evolving?," Radio Free Europe/Radio Liberty, 09/10/2004	Abbreviated Found online at: http://www.rferl.org/content/article/1054788.html	The full article discusses the evolution of the term genocide and the prosecution of offenders, the question that follows was answered by an interviewed person
Doc. 2b	Primary Source Political Cartoon Genocide	Yes Steve Greenburg, <i>Seattle Post-Intelligencer</i> , March 29, 1999 (adapted)	Adapted (minor changes) http://greenberg-art.com/.Toons/.Toons,%20political/Genocide.html	There are minute changes that do not interfere with meaning. The names of ethnic genocides are typed in the Regents form instead of handwritten and the original appears in color.

Doc. 3	Secondary Source Excerpt from a book on Genocide	Yes Louis Henkin, "Human Rights: Ideology and Aspiration, Reality and Prospect," <i>Realizing Human Rights</i> , St. Martin's Press, 2000	Could not be found online Maybe part of another work: "Realizing Human Rights"	
Doc. 4	Secondary Source Textbook Author composed only the introduction of the 3 rd Edition	Yes Norman J. Vig, "Introduction: Governing the International Environment," <i>The Global Environment: Institutions, Law, and Policy</i> , CQ Press, 2005 (adapted)	Could be purchased online but not viewed	
Doc. 5	Primary Source Newspaper Article	Yes Frank Langfitt, "Desertification," <i>The Post-Standard</i> , May 13, 2002 (adapted)	Could not be found online Browsed the Post Standard Archives and nothing was found	
Doc. 6	Secondary Source Chart of Events given in chronological	Yes "Environmental Milestones," World Watch Institute (adapted)	Adapted Found in entirety at: http://www.worldwatch.org/brain/features/timeline/timeline.htm	Abbreviated and only certain dates and events were chosen from the overall timeline. Valuable information was left out and other
Doc. 7	Secondary Source Timeline Illustrates weapons of mass	Yes "Weapons of Mass Destruction," National Geographic,	Adapted Timeline could not be found online though this article was: http://ngm.nationalgeographic.org	

Doc. 8	Primary Source Political Cartoon Einstein between Pakistan and India	Yes Jeff Danziger, Tribune Media Services, January 4, 2002 (adapted)	Adapted(slightly) Found online at: http://www.danzigercartoons.com/archive/cmp/2002/danziger1183.html	The only adaptation was typing Einstein's words and they were originally handwritten
Doc. 9	Primary Source Article Excerpt	Yes Glaser and von Hippel, "Thwarting Nuclear Terrorism," <i>Scientific American</i> , February 2006	Abbreviated Found in full at: http://www.bnl.gov/news/news/SciAm0206Fishbone.pdf	

Table V				
<i>Validating the Authenticity of the Documents in Regard to their Presentation within the DBQ</i>				
DBQ June 2007	Type of Document	Is the source given? Is it adapted?	Could the document be readily found for further investigation or inquiry?	Underlying bias in the presentation of the documents? Do they portray differing views?
Doc. 1	Secondary Source Diagram (Medieval Manor)	Yes Kime and Stich, <i>Global History and Geography</i> , <i>STARreview</i> , N & N Publishing Company		Presented in full and in context of the task, though not directly stated in the document itself
Doc. 2	Primary Source Written document Tasks of serfs	Yes S. R. Scargill- Bird, ed., <i>Customals of Battle Abbey in the Reigns of Edward I and Edward II (1283- 1312)</i> The Camden Society (adapted)	Adapted Original found at: http://www.archive.org/stream/custumalsofbattl00battrich/page/n19/mode/2up Document taken from pages xiii-xiv	Document was truncated and numerous details were deleted, however portrays the main idea in the Regents form
Doc. 3	Secondary Source Written passage from textbook Economy of Medieval Period	Yes Norman F. Cantor, <i>The Civilization of the Middle Ages</i> , Harper Perennial	Abbreviated Taken from textbook ISBN: 9780060925536; ISBN10: 0060925531; Imprint: Harper Perennial ; On Sale: 6/3/1994; Format: Trade PB; Trimsize: 5 5/16 x 8; Pages: 624; \$18.99; Ages: 18 and Up; BISAC1:HIS000000; BISAC2:HIS010000; BISAC3:HIS037010	

Doc. 8	Secondary Source Newspaper Article Covering effects of China's "Great Leap Forward"	Yes BBC News, Special Reports, China's Communist Revolution	Found online in full at: http://news.bbc.co.uk/1/hi/english/static/special_report/1999/09/99/china_50/great.htm Published originally on October 6, 1999
Doc. 9	Primary Source Transcript of speech given by Deng Xiaoping titled, "We Shall Speed Up Reform"	Yes Deng Xiaoping, Fundamental Issues in Present- Day China, Foreign Language Press, 1987	Found online in republished form at: http://www.scribd.com/doc/55967806/46/COLLECTION S

Table VI					
<i>Readability Levels of the Documents</i>					
DBQ: June 2011	Flesch Kincaid	Okapi (Dale Chall)	Lexile Framework for Reading	Lowest Average of all Three	Highest Average of all Three
Doc. 1	9.4	Index: 9.69 Raw Score: 13-15 th Grade Level	1120L 10 th Grade level	Raw Average – 10.8 10-11 th grade	Raw Average – 11.46 11-12 th grade
Doc. 2	9.3	Index: 10.24 Raw Score: 16 th Grade Level	1240L 12 th Grade Level	Raw Average – 12.43 12 th grade to College Level	Raw Average – 13.1 College Level
Doc. 3	12.9	Index: 10.68 Raw Score: 16 th Grade Level	1130L 10-11 th Grade Level	Raw Average – 12.96 12 th grade to College Level	Raw Average – 13.96 College Level
Doc. 4	9.1	Index: 10.44 Raw Score: 16 th Grade Level	1300L College Level	Raw Average – 12.7 12 th grade to College Level	Raw Score - 13.36 College Level
Doc. 5	9.3	Index: 9.97 Raw Score: 13-15 th Grade Level	1030L 8-9 th Grade Level	Raw Average – 10.1 10 th Grade Level	Raw Average – 11.1 11 th Grade Level
Doc. 6 (88 words)	4.6	Index: 8.24 Raw Score: 11-12 th Grade Level	640L 3 rd Grade Level	Raw Average – 6.2 6 th Grade Level	Raw Average – 7.2 7 th Grade Level
Doc. 7	8.2	Index: 10.11 Raw Score: 16 th Grade Level	1200L 11-12 th Grade Level	Raw Average – 11.73 11-12 th Grade Level	Raw Average - 12.73 12 th Grade to College Level
Doc. 8	8.7	Index: 9.45 Raw Score: 13-15 th Grade Level	1030L 8-9 th Grade Level	Raw Average – 9.9 9-10 th Grade Level	Raw Average – 10.9 10-11 th Grade Level
Doc. 9	10.3	Index: 9.5 Raw Score: 13-15 th Grade Level	1440L College Level	Raw Average - 12.1 12 th Grade to College Level	Raw Average – 13.43 College Level

Table VII					
<i>Readability Levels of the Documents</i>					
DBQ: June 2010	Flesch Kincaid	Okapi (Dale Chall)	Lexile Framework for Reading	Lowest Average of all Three	Highest Average of all Three
Doc. 1	7.8	Index: 10.79 Raw Score: 16 th Grade Level	1480L College Level	Raw Average - 12.26 12 th Grade to College Level	Raw Average - 12.93 College Level
Doc. 2	6.8	Index: 8.44 Raw Score: 11-12 th Grade Level	1190L 9-10 th Grade Level	Raw Average - 8.93 8-9 th Grade Level	Raw Average - 10.26 10 th Grade Level
Doc. 3	6.1	Index: 8.25 Raw Score: 11-12 th Grade Level	1370L College Level	Raw Average - 10.03 10 th Grade Level	Raw Average - 11.03 11 th Grade Level
Doc. 4	6.9	Index: 7.63 Raw Score: 9-10 th Grade Level	1260L 11-12 th Grade Level	Raw Average - 8.96 8-9 th Grade Level	Raw Average - 10.3 10 th Grade Level
Doc. 5 (abbreviated to 174 words)	6.4	Index: 7.99 Raw Score: 9-10 th Grade Level	1650L College Level	Raw Average - 9.46 9 th Grade Level	Raw Average - 10.46 10 th Grade Level
Doc. 6	N/A	N/A	N/A	N/A	N/A
Doc. 6b	8.9	Index: 10.35 Raw Score: 16 th Grade Level	1200L 11-12 th Grade Level	Raw Average- 11.96 11-12 th Grade Level	Raw Average - 12.96 12 th Grade to College Level
Doc. 7	8.8	Index: 8.98 Raw Score: 11-12 th Grade Level	1430L College Level	Raw Average - 10.93 10-11 th Grade Level	Raw Average - 11.93 11-12 th Grade Level
Doc. 8	N/A	N/A	N/A	N/A	N/A

Doc. 9	7.1	Index: 11.51 Raw Score: 16 th Grade Level	1310L College Level	Raw Average – 12.03 12 th Grade Level	Raw Average – 12.7 12 th Grade to College Level
--------	-----	---------------------------------------------------------------	------------------------	-----------------------------------------------------------	---------------------------------------------------------------------------

Table VIII					
<i>Readability Levels of the Documents</i>					
DBQ: June 2009	Flesch Kincaid	Okapi (Dale Chall)	Lexile Framework for Reading	Lowest Average of all Three	Highest Average of all Three
Doc. 1	7.9	Index: 9.02 Raw Score: 13-15 th Grade Level	830L 5 th grade	Raw Average – 8.63 8-9 th Grade Level	Raw Average – 9.96 9-10 th Grade Level
Doc. 2	3.3	Index: 9.29 Raw Score: 13-15 th grade level	1420L College Level	Raw Average – 9.76 9-10 th Grade Level	Raw Average – 11.1 11 th Grade Level
Doc. 3	7.6	Index: 10.10 Raw Score: 16 th Grade Level	1430L College Level	Raw Average – 12.2 12 th Grade Level	Raw Average – 12.86 12 th Grade to College Level
Doc. 4	N/A	N/A	N/A	N/A	N/A
Doc. 5	10.5	Index: 9.81 Raw Score: 13-15 th Grade Level	1380L College Level	Raw Average – 12.16 12 th Grade to College Level	Raw Average – 13.5 College Level
Doc. 6	10.1	Index: 7.15 Raw Score: 9-10 th Grade Level	1130L 10-11 th Grade Level	Raw Average – 9.7 9-10 th Grade Level	Raw Average – 11.03 11 th Grade Level
Doc. 6b (86 words)	3.6	Index: 5.98 Raw Score: 5-6 th Grade Level	840L 5 th Grade Level	Raw Average – 4.53 4 th Grade Level	Raw Average – 5.53 5 th Grade Level
Doc. 7	5.0	Index: 7.58 Raw Score: 9-10 th Grade Level	920L 8-9 th Grade Level	Raw Average – 7.33 7 th Grade Level	Raw Average – 8.66 8-9 th Grade Level
Doc. 8	11.5	Index: 9.79 Raw Score: 13-15 th Grade Level	1500L College Level	Raw Average – 12.5 12 th Grade to College Level	Raw Average – 13.83 College Level
Doc. 9	8.8	Index: 8.61 Raw Score: 11-12 th Grade Level	1200L 11-12 th Grade Level	Raw Average – 10.26 10 th Grade Level	Raw Average – 11.6 11-12 th Grade Level

Table IX					
<i>Readability Levels of the Documents</i>					
DBQ: June 2008	Flesch Kincaid	Okapi (Dale Chall)	Lexile Framework for Reading	Lowest Average of all Three	Highest Average of All Three
Doc. 1	9.7	Raw Score: 16 th grade level , Index 11.59	1740L College Graduate	Raw Average – 13.9 College Level	Raw Average – 14.56 College Level
Doc. 2	8.5	Raw Score: 16 th Grade Level	1460L College Level	Raw Average – 12.5 12 th Grade to College Level	Raw Average – 13.1 College Level
Doc. 3	11.6	Index: 12.35 Raw Score: 16 th Grade Level	1340L Grade 12	Raw Average – 13.2 College Level	Raw Average – 13.86 College Level
Doc. 4	8.7	Index: 9.35 Raw Score: 13-15 th Grade Level	1260L Grade 11-12	Raw Average – 10.9 10-11 th Grade Level	Raw Average – 12.56 12 th Grade to College Level
Doc. 5	10.7	Index: 8.95 Raw Score: 11-12 th Grade Level	1260L Grade 11-12	Raw Average – 10.9 10-11 th Grade Level	Raw Average – 12.23 12 th Grade Level
Doc. 6	16.2	Index: 12.94 Raw Score: 16 th Grade Level	1260L Grade 11-12	Raw Average – 14.4 College Level	Raw Average – 15.4 College Level
Doc. 7	N/A	N/A	N/A	N/A	N/A
Doc. 8	N/A	N/A	N/A	N/A	N/A
Doc. 9	10.2	Index: 10.92 Raw Score: 16 th Grade Level	1410L College Level	Raw Average – 13.06 College Level	Raw Average – 13.73 College Level

Table X					
<i>Readability Levels of the Documents</i>					
DBQ: June 2007	Flesch Kincaid	Okapi (Dale Chall)	Lexile Framework for Reading	Lowest Average of all Three	Highest Average of all Three
Doc. 1	N/A	N/A	N/A	N/A	N/A
Doc. 2	N/A	N/A	N/A	N/A	N/A
Doc. 3	6.4	Index: 8.9 Raw Score: 11-12 th Grade Level	1070L 8-9 th Grade Level	Raw Average – 8.46 8-9 th Grade Level	Raw Average – 9.8 9-10 th Grade Level
Doc. 4	N/A	N/A	N/A	N/A	N/A
Doc. 5	N/A	N/A	N/A	N/A	N/A
Doc. 6	6.2	Index: 8.15 Raw Score: 11-12 th Grade Level	1910L College Graduate Level	Raw Average - 11.06 11 th Grade Level	Raw Average - 12.06 12 th Grade Level
Doc. 7	5.6	Index: 9.44 Raw Score: 13-15 th Grade Level	1370L College Level	Raw Average - 10.53 10 th Grade Level	Raw Average - 11.86 11-12 th Grade Level
Doc. 8	10.3	Index: 8.97 Raw Score: 11-12 th Grade Level	1150L 9 th Grade	Raw Average – 10.1 10 th Grade Level	Raw Average – 11.1 11 th Grade Level
Doc. 9	7.0	Index: 8.16 Raw Score: 11-12 th Grade Level (shortened to 190 words)	1110L 9 th Grade Level	Raw Average – 9 9 th Grade Level	Raw Average – 10 10 th Grade Level