Literacy, Meritocracy, and the National Assessment of Educational Progress: the Problem of Standardized Testing in America

---

A Thesis

Presented to

The Division of History and Social Sciences

Reed College

In Partial Fulfillment

of the Requirements for the Degree

Bachelor of Arts

Joseph Aaron Wasserman

May 2009

Approved for the Division

(Anthropology)

Ben Peacock

Table of Contents

Introduction **1**
Outline of the Chapters 1
Chapter 1: Standardized Tests in America **3**
Introduction 3
Institutional Structures Surrounding the National Assessment of Educational Progress 3
National Assessment Governing Board 3
National Center for Education Statistics 4
Contractors 5
The No Child Left Behind Act of 2001 6
History of Standardized Tests 6
Anthropological Critiques of Testing 7
Chapter 2: Theorists and Ethnographers of Education and Social Reproduction **9**
Introduction 9
Theorists of Education and Social Reproduction 9
Althusser 9
Bourdieu and Passeron 10
Bowles and Gintis 10
Foucault 11
Groups: Reification, Essentialization, and Comparison 12
Ethnographers of Education and Social Reproduction 17
Shirley Brice Heath 18
Paul Willis 19
Jay MacLeod 20
Reinvigorating Reproduction Theory 22
Chapter 3: The National Assessment of Educational Progress **23**
Introduction 23
General Background 23
Theoretical Grounding 23
Contextualization 24
'Use information from the text to support your answer': What Counts as Literacy on NAEP 25
Specificity 25
Stated Generalizations 26

Superficiality 29
Decontextualization 30
Speaking to Authority 32
Configurations of Relations to Text, Relations to Knowledge 35
Experiential Bias 39
Science Fairs 39
When Experience Harms 41
Conclusion 42
Conclusion 45
Literacy and Ideologies of Standard 45
What Might Be Done 46
Connections 46
Suggestions 47
Research Projects 47
Appendix A: Student Background Questionnaire 49
Appendix B: NAEP 2007 8th Grade Reading Passage 51
Appendix C: NAEP 2005 12th Grade Reading Passages 53
Bibliography 61

List of Figures

Figure 1. Trend in fourth-grade NAEP reading average scores and score gaps, by selected racial/ethnic groups (NCES 2007c: 11) 15

Figure 2. Grade 4 NAEP Reading Item map (NCES 2007c: 21) 16

Figure 3. ‘Partial’ student response to the 2007 8th grade NAEP reading test 26

Figure 4. ‘Unacceptable’ student response to the 2007 8th grade NAEP reading test 27

Figure 5. ‘Partial or surface comprehension’ student response to the 2005 12th grade NAEP reading test. 28

Figure 6. ‘Essential’ student response to the 2005 12th grade reading test. 28

Figure 7. ‘Essential’ response to the 2005 12th grade NAEP reading test. 30

Figure 8. ‘Full comprehension’ response to the 2005 12th grade NAEP reading test. 31

Figure 9. ‘Partial or surface comprehension’ response to the 2005 12th grade NAEP reading test. 31

Figure 10. ‘Partial or surface comprehension’ response to the 2007 8th grade NAEP reading test. 32

Figure 11. ‘Little or no comprehension’ response to the 2007 8th grade NAEP reading test. 33

Figure 13. ‘Little or no comprehension’ response to the 2005 12th grade NAEP reading test. 35

Figure 14. ‘Full comprehension’ response to the 2007 8th grade NAEP reading test. 36

Figure 15. ‘Full comprehension’ response to the 2007 8th grade NAEP reading test. 37

Figure 16. ‘Partial or surface comprehension’ response to the 2007 8th grade NAEP reading test. 37

Figure 17. ‘Little or no comprehension’ response to the 2007 8th grade NAEP reading test. 37

Figure 18. ‘Acceptable’ response to the 2007 8th grade reading test 39

Figure 19. ‘Acceptable’ response to the 2007 8th grade reading test 40

Figure 20. ‘Unacceptable’ response to the 2007 8th grade reading test. 40

Figure 21. ‘Unacceptable’ response to the 2007 8th grade reading test. 40

Figure 22. ‘Little or no comprehension’ response to the 2005 12th grade NAEP reading test. 42

Abstract

Standardized tests are ubiquitous in any number of institutions in the United States, but they most pervade the educational system. As such, they are an important area for serious inquiry. Unlike IQ tests, however, anthropologists have spared standardized tests critiques of bias and implicit racism. I argue that the reading and writing tests of the National Assessment of Educational Progress implicitly construct a standard of literacy that favors the socioculturally and historically specific histories of socialization that tend to be associated with dominant classes at the expense of the equally specific and contingent histories of other groups. This implicit standard is revealed by a close reading of test questions, student responses, scorer comments, and scoring guides. By conceiving of test scores as being caused by essentialized qualities inherent in groups, aggregate test scores articulate with a meritocratic ideology to legitimate and reproduce racialized hierarchical social and economic relations. To see how test scores can reproduce and legitimate these unequal class relations, it is necessary to draw on reproductionist theories that explain class outcomes after schooling as strongly influenced by class origin and legitimated by the process of schooling itself.

Introduction

Standardized tests of various sorts pervade the lived experiences of many in the United States. IQ scores are frequently used as convenient indicators of 'intelligence', other tests are used to determine fit and unfit drivers, and exams are given to evaluate potential employees. But in the U.S., standardized tests are most prevalent in the education system. They are used to evaluate and track individual students, to measure the performance of schools or districts, to sort students into institutions of higher education or receive college credit for Advanced Placement high school classes, and even to assess the state of students' knowledge nationwide.

The ways in which standardized test questions, responses, and scores are generally conceptualized and put to use, however, are severely problematic. I argue that standardized tests reproduce and legitimate hierarchical class relations and dramatic socioeconomic inequalities. They obscure class-related cultural capital, relations to text and to knowledge, and socialization practices, as well as articulate with meritocratic ideologies to legitimate and reproduce racialized hierarchical class relations. By naturalizing test scores as a direct reflection of a student's knowledge or aptitudes and essentializing these qualities as inherent in the groups of which they are a member, individuals and institutions ignore the socioculturally and historically contingent nature of students' answers. This 'emblematization' obscures the variety of ways in which students have been socialized in school, at home, and elsewhere into habitual practices and relations to texts and to knowledge. Moreover, the practices and relations to text and to knowledge that are valued on standardized tests tend to be those associated with 'mainstream', dominant groups.

Outline of the Chapters

In Chapter 1, I lay out the institutional organization of the bureaucracies and agencies surrounding the National Assessment of Educational Progress (NAEP) to situate NAEP in the larger field of United States educational policy and testing in general. Next, I sketch a brief history of standardized testing in the United States before elaborating earlier anthropological critiques of IQ tests. As part of these American anthropologist's anti-racist projects, they critique the racial biases of IQ tests and the unfounded assumptions backing popular notions of IQ.

In Chapter 2, I cover aspects of a number of theorists and ethnographers of education that are central to my discussion of standardized tests, focusing on

reproductionist accounts of classrooms and education. These authors provide theories and concrete examples of both the reproduction and the legitimation of hierarchical class relations by the educational system.

In Chapter 3, I provide an analysis of publicized NAEP reading and writing test questions, student responses, scorer evaluations, and scoring guides to draw out the particular standard of 'literacy' that is implicit in NAEP. I conclude by using theories of social reproduction as a lens through which to explore the ways in which NAEP test results articulate with meritocratic ideologies to reproduce and legitimate racialized hierarchical class relations.

Chapter 1: Standardized Tests in America

Introduction

I have three goals for this chapter. First, I sketch the institutional roles and constitution of the various organizations related to the National Assessment of Educational Progress (NAEP). For this section, it is important to keep in mind that I am presenting almost entirely what is the public face of these organizations—all of this material has been gathered from their respective websites. Second, I present a brief history of standardized testing in the United States. Finally, I situate my critique of standardized literacy tests in terms of earlier anthropological critiques of IQ tests.

Institutional Structures Surrounding the National Assessment of Educational Progress

National Assessment Governing Board

According to their website (NAGB n.d.-b), the National Assessment Governing Board (NAGB) is responsible for:

- Selecting subject areas to be assessed
- Developing appropriate student achievement levels
- Developing assessment objectives and test specifications that produce an assessment that is valid and reliable, and are based on relevant widely accepted professional standards
- Designing the methodology of the assessment
- Developing guidelines for reporting and disseminating results
- Developing standards and procedures for regional and national comparisons
- Approving all cognitive and noncognitive NAEP items
- Taking appropriate actions needed to improve the form, content, use, and reporting of results
- Planning and executing the initial public release of National Assessment of Educational Progress reports

Because of these broad responsibilities, they have control not only over the form and content of NAEP, but also influence published reports on test results. The bipartisan Board is composed of 26 individuals, none of whom can serve more than four years: two governors, two 'state legislators', two chief state school officers, one private school administrator, one state school board member, one local school board member, three

teachers, three ‘testing experts’, two ‘curriculum specialists’, two principals, one superintendent, the director of the Department of Education’s Institute of Education Sciences, two parents, two members of the ‘general public’, and one ‘representative of the business community’, all of whom are appointed by the U.S. Secretary of Education (NAGB n.d.-c). Conspicuously absent from this board are students themselves and social scientists who may be more conscious and more critical of the ways in which NAEP might articulate with social structures and common-sense ideologies.

As specified in the No Child Left Behind Act of 2001 ('No Child Left Behind Act of 2001' 2002), the only portions of NAEP that states must administer are the reading and mathematics tests for 4th and 8th graders, which are given every other year. All other subjects, including geography, writing, and science, are optional. According to the NAGB website, ‘all students who participate in NAEP do so on a voluntary basis.’ (NAGB n.d.-c). Even so, it is difficult to imagine that students are fully aware of their options for opting out of NAEP testing in a hierarchical classroom setting. The website continues, ‘NAEP is forbidden by law to maintain or report information on individual students or schools’ (NAGB n.d.-c). Results and data therefore must be aggregated by groups, and are not kept for either individual students or for particular schools. This fact has dramatic consequences not only for the kind of information the National Center for Education Statistics (NCES) reports about NAEP test results, but also for the analytical moves and claims they make in their reports, as discussed in Chapter 2.

Nowhere on the NAGB website is there a mention of scorers, how they are trained, who they are, or what they do. This seems like part of their project of naturalizing NAEP test scores as a direct reflection of student ‘literacy’, unmediated by an interpreter. When discussing methodology (e.g. NAGB n.d.-a), the focus is on sampling and content development, not scoring.

<u>National Center for Education Statistics</u>

After NAGB develops content frameworks, NCES writes the actual test questions that appear on NAEP. Before administration, however, these test questions ‘undergo extensive field and pilot testing, gathering actual responses from students’ (NCES 2008a). NCES also administers, scores, analyzes, and reports the results of NAEP as *The Nation’s Report Card* (e.g. NCES 2007c).

One section on the NCES website—to which, not incidentally, I could only navigate after contacting them for assistance because it is unintuitively accessible only from the ‘Sample Questions’ page—briefly describes information relevant to scorer training and concrete scoring practices (NCES 2005b). To prepare the scoring guides that

serve as references for scorers on how to rate different responses, the NAEP standing committee selects student responses as examples of particular scoring categories. If the examples selected to appear on their website (analyzed extensively in chapter 3) are any indication, this selection process is highly problematic and is likely to influence scoring dramatically. To qualify as a test scorer, one must have at least a bachelor's degree. Ironically, scorers have to pass a 'scorer placement test' before they are hired. Training involves instruction from unspecified 'content and scoring experts', reviewing the aforementioned scoring guides, and scoring practice papers. Finally, trainees must assign scores to a set of papers that are consistent with those assigned by NAEP 'content and scoring experts' at least 70% of the time.

Once scorers begin their job, there are several processes that ensure their agreement. When scorers cannot score 'accurately' and 'consistently' (i.e. without disagreeing frequently), they are retrained or removed from scoring particular items. The 'accuracy' and 'consistency' of scoring is measured by double-scoring a number of tests, random checking of scores, and statistical monitoring of scorer ratings.

Contractors

NCES work with a variety of contractors for a number of purposes. From 2003-2007, the years for which I analyze NAEP, they contracted with the following: Educational Testing Service (ETS), American Institutes for Research (AIR), Pearson, Westat, Inc., Government Micro Resources, Inc., Hager Sharp, and Human Resources Research Organization (NCES 2008b). ETS—known for their administration of standardized tests like Advanced Placement (AP), SAT, Graduate Record Examinations (GRE), and Test of English as a Foreign Language (TOEFL)—coordinated the ensemble of contractors, ensured the validity of results, and designed test questions for the writing, reading, mathematics, and science NAEP tests. To ensure the validity of test results, they designed field and pilot tests and analyzed statistical reports.

AIR is an organization for behavioral and social science research. For NAEP, they developed the background variables requested from all test-takers, 'cognitive items', and scoring rubrics. They also reviewed test questions, performed pilot tests of questions and scoring rubrics, and 'assist[ed]... in the training of scorers' (NCES 2008b). AIR do not elaborate on their training role on their website (AIR). They also played a central role in analyzing test results, explaining results, and identifying patterns such as so-called 'achievement gaps'.

Pearson published the test itself and distributed it to test administrators. They also developed 'training and scoring materials' and scored the tests. Among other

responsibilities, Westat, Inc. contributed to statistical analyses by collecting data, selecting representative samples, and preparing 'sample weights for NAEP assessments'. Hager Sharp is a PR firm that created all NAEP products and dissemination plans for test results. Because all of these firms are such major players in educational testing, publishing, and research, it is likely that at least some of the problems plaguing NAEP can be found in other standardized tests as well. Though beyond the scope of this thesis, the tests produced by ETS—AP, SAT, GRE, TOEFL—seem like fruitful objects for further analysis, particularly as they figure so prominently in the post-secondary education system of the United States.

The No Child Left Behind Act of 2001

In order to qualify for Title I funding under the No Child Left Behind Act of 2001 (NCLB), which authorized $25 billion of appropriations for 2007 ('No Child Left Behind Act of 2001' 2002 Sec. 1002), states must participate in the biennial 4th and 8th grade reading and math NAEP tests, paid for by the federal government. Although *administration* of NAEP is required to receive Title I funding, no rewards or punitive sanctions are decided by NAEP *results*. Instead, states are required to administer *additional* mathematics and reading assessments to 3rd-8th graders for these purposes (NCES 2005a). Still, results on NAEP are used to evaluate and legitimate educational policy (e.g. U.S. Department of Education 2008), especially concerning closing racial 'achievement gaps' (e.g. U.S. Department of Education 2005). It is this legitimating role that makes the standard of literacy that is implicit in NAEP, as well as interpretations of NAEP test results, so problematic.

History of Standardized Tests

In 1905, Alfred Binet and associates published the Binet-Simon intelligence scale, from which current intelligent tests still draw most of their evaluative tasks. Although this particular scale was developed in France to distinguish so-called 'retarded' from normal children, it was soon expanded to classify children by age levels. By 1916, Lewis Terman had translated the Binet-Simon scale into English and adapted its content to suit the United States, intending to use it primarily to classify mentally deficient children and adults. When the United States entered World War I, these tests were latched onto as a means for promoting officers and sorting millions of recruits into specific roles and positions. Instead of using Binet's age level scale, the psychologist heading the development of tests for the Army, Robert Yerkes, promoted the more familiar point

scale (Kaufman 1990: 4-7). As the United States economy industrialized, schools were increasingly modeled after assembly-lines and factories, in which students progressed from one grade to the next with the explicit goal of producing students with standardized sets of knowledge, as measured by standardized tests. Furthermore, standardized tests made it easy to track children into different types of curricula as they progressed through the various levels of the educational system (Falk 2000: 3-6).

Anthropological Critiques of Testing

As part of their more general anti-racist projects, early 20th century anthropologists critiqued IQ tests and the purposes to which IQ test results were put from a number of perspectives. Ruth Benedict and Gene Weltfish (1943) sought to dispel common-sense, essentialized ideas about racial intelligence in their popular book, *The Races of Mankind*. They claimed that '"Northern Negroes" had a higher IQ than "Southern Whites"'(Price 2003: 3) as a result not of qualities inherent to members of those groups, but different histories (Benedict and Weltfish 1943: 18). Otto Klineberg, one of Franz Boas's students, made a similar claim, arguing that results of intelligence tests are a consequence not of a property inherent in racial categories, but of environmental influence (Klineberg 1935). Another popular anthropologist, Ashley Montagu, critiqued the notion of race in general in his book, *Man's Most Dangerous Myth: The Fallacy of Race* (Montagu 1947). The point of all of these critiques is that IQ is not an indicator of some quality inherent to a racial group, but the consequence of multiple socioculturally and historically specific factors, including socialization.

More recently, anthropologists have been strongly critical of Richard Herrnstein' and Charles Murray's *The Bell Curve* (Herrnstein and Murray 1994). For example, among the critiques made in a six-author review article (Alland Jr. et al. 1996) is the claim that intelligence tests reflect not 'intelligence' but economic and social stratification. But as David Price has pointed out as recently as 2003 (Price 2003), other standardized tests have received little critical anthropological attention. But IQ tests and standardized literacy tests are similar in significant ways, and are therefore both deserving of scrutiny. Just as IQ tests were used to sort recruits into roles during WWI, so are standardized tests used to sort individuals into educational tracks and particular jobs by suggesting career tracks, influencing hiring practices, and influencing their access—or lack thereof—to particular kinds of higher education. What is more, both aggregate IQ scores and aggregate standardized test scores provide a framework for the legitimation of racialized socioeconomic inequality. Because standardized test scores are increasingly

being used to measure the 'educational progress' of racial, ethnic, and gender groups in the U.S. and to mete out punitive sanctions under NCLB, it is pressing to explore what these tests actually measure. Ignoring standardized tests is a great disservice to the groups whose social and economic subordination is reproduced and legitimated via test scores. Part of the reason for this neglect seems to stem from anthropology's move as a discipline away from structural theories of social reproduction. It is to a discussion of these reproductionist accounts of socioculture that I now turn.

Chapter 2: Theorists and Ethnographers of Education and Social Reproduction

Introduction

This chapter situates standardized tests and the National Assessment of Educational Progress (NAEP) in terms of theories and ethnographies of the school, education, and social reproduction. I point out ways in which standardized tests can legitimate and reproduce socioeconomic hierarchies, as well as suggesting some alternative ways of reading standardized test results. Finally, I suggest that anthropology should reinvigorate reproductionist theoretical perspectives.

Theorists of Education and Social Reproduction

Althusser

In 'Ideology and Ideological State Apparatuses', (Althusser 2001), Althusser considers the school the foremost Ideological State Apparatus (ISA) in capitalist society. In his theory, the body of institutions that are ISAs are distinct from Repressive State Apparatuses, in that they function not by force and coercion, but instead by 'ideology', defined as 'not the real relations which govern the existence of individuals, but the imaginary relation of those individuals to the real relations in which they live' (Althusser 2001: 111). The function of the school *qua* ISA, then, is to inculcate ideologies that reproduce the class structure into students. Accordingly, these ideologies are appropriate to the positions in the relations of production that students occupy when they are 'ejected' from the education system, and it is by this class-appropriateness that ideology reproduces class relations. For Althusser, the consequence is that the education system ensures the reproduction of exploitative capitalist relations of production by producing groups who have ideologies appropriate to 1) the exploited, 2) the agents of exploitation, 3) the agents of repression, and 4) the 'professional ideologists' who legitimate these relations of production. By being ideologically framed as non-ideological, the school and its agents generally perform this inculcation uncontested and unquestioningly. In this view, then, to reproduce the labor force it is *necessary* that some students perform poorly on standardized tests that would otherwise potentially qualify them for other occupations. I will show that, like the school, standardized tests both contribute to the reproduction of

class structures and are framed as non-ideological, or *natural*, which inhibits critical examination and reflection on these tests.

Bourdieu and Passeron

In their simultaneously descriptive and highly theoretical work on the French education system and education systems in general, *Reproduction in Education, Society, and Culture* (Bourdieu and Passeron 1990), Bourdieu and Passeron examine the processes by which class origins strongly influence class outcomes after schooling. They argue that children enter school with a set of resources including linguistic, cultural, and social capital; economic and power relations; an 'ethos', i.e. dispositions towards school, values, authority, &c.; expectations of access to, success in, and the effects of school; and relations to language and culture, all of which are a consequence of their initial class membership, in a broad sense of 'class' as categorical group (a meaning which I follow throughout). According to their analysis, all of these class-originating resources, dispositions, and relations are 'retranslated' through activity in the school into criteria that are evaluated as purely *academic*, rather than *class*, criteria. That is, when schools evaluate students on criteria that are purportedly learned *in* the school, these evaluations actually measure the resources and dispositions that students have *before* entering school, and what is more, these resources (like 'linguistic capital') enable only some students to maximally acquire pedagogic knowledge. As students progress higher in the educational system, the connections between academic criteria and their class origins become further removed. These retranslated criteria legitimate class outcomes by making them seem like the products of purely academic successes or failures instead of matters of class origin. Along the way, students internalize their perceived chances of success in the workforce and in the education system. By the time they are ejected from the school system, they have internalized a habitus and a perspective on their 'objective probability' of economic success that is appropriate to their position in the relations of production. As I will show, standardized tests like NAEP function very similarly as they retranslate class criteria into purely 'academic' criteria. In the case of NAEP reading and writing tests, this process is made possible by constructing a monolithic standard of 'literacy' that is conceived of as universal and value-free.

Bowles and Gintis

Despite basing their analysis on extensive statistical analysis, Samuel Bowles' and Herbert Gintis's *Schooling in Capitalist America* (1976) is so structural and abstracted from the level of lived reality that it is more accurately considered theoretical

than descriptive. In their analysis of the United States educational system, they examine how schooling is involved in the legitimation and reproduction of inequality and repressive hierarchical relations of production, coming to conclusions very similar to those of Bourdieu and Passeron. The educational system is involved in what Bowles and Gintis term the 'technocratic-meritocratic ideology' that legitimates inequality. Because their statistics indicate that both grades and IQ contribute only little to economic success, they claim that these potential indicators of 'merit' are used only symbolically to legitimate structures of inequality. They point to the functionalist explanation of hiring (e.g. Davis and Moore 1945), in which differential reward is a consequence of the importance of a role and the scarcity of personnel to fill that role, as the epitome of this ideology. Furthermore, they argue that continual interaction with the school—which is already conceptualized and perceived as strongly linked to economic rewards—on the part of students and parents reinforces a spurious connection between cognitive qualities and economic success. Finally, by meting out frequent punishments for failure in school, students' aspirations are lowered.

In addition to legitimating hierarchical relations of production, the school reproduces this economic structure by developing in students 'the types of personal demeanor, modes of self-presentation, self-image, and social-class identifications which are the crucial ingredients of job adequacy' (Bowles and Gintis 1976: 131). Because the hierarchical relations students encounter in the workplace are analogous to those encountered in the school, socialization into social relations in school prepares students for the social relations of work. Translated into the language of Bourdieu, they claim that the school cultivates students' habitus such that they are socialized into the attitudes and social relationships that will be required of them in the workplace.

Foucault

As presented by Foucault in his book, *Discipline and Punish* (Foucault 1995), the examination, in combination with the profusion of written documents that fix individuals in relation to each other based on their scores, makes two interrelated consequences possible:

> firstly, the constitution of the individual as a describable, analysable object, not in order to reduce him to 'specific' features, as did the naturalists in relation to living beings, but in order to maintain him in his individual features, in his particular evolution, in his own aptitudes or abilities, under the gaze of a permanent corpus of knowledge; and, secondly, the constitution of a comparative system that made possible the measurement of overall phenomena, the description of groups, the characterization of collective facts, the

> calculation of the gaps between individuals, their distribution in a given 'population'. (Foucault 1995: 190)

By differentiating individuals along a single scale, examinations allow not only the comparison of individuals to each other, but also the comparison of individuals to quantitative standards that, if unmet, indicate deviance and deficiency. Moreover, examinations allow objective (in the sense that they were inscribed by individuals as incontrovertibly physical artifacts) qualities to be assigned to groups. Although Foucault emphasizes the distribution and comparison of individuals, a crucial consequence of these objective qualities is the ability to compare *groups* by fixing them, their individual members having been aggregated, in particular relations to each other along a hierarchized quantitative scale. Once on such a scale, an entire *class* of people can be marked deviant or deficient, *ab*normal. The National Center for Education Statistics (NCES) makes precisely these moves with the results of NAEP in their published material.

Groups: Reification, Essentialization, and Comparison

Understanding the formation, description, and comparison of *groups* is critical for understanding the force of NAEP test scores. In this case, all groups are constituted by an act of self-reporting by test-takers in a questionnaire completed before the test itself (NCES 2007b)[1]. At this moment, individuals reify their identity and group membership by fixing themselves into these categories with an oval filled in by number 2 pencil. For the purpose of their published reports, NCES focus on economic, gender, and racial categories. Economic classes are determined by eligibility for free or reduced-price lunch, while gender is a 'simple' choice between male and female. The racial categories from which students can select on the 2007 Reading test are 'White', 'Black or African American', 'Asian', 'American Indian or Alaska Native', or 'Native Hawaiian or other Pacific Islander'. A separate question asks whether the student is 'Hispanic or Latino'. In addition to the complicated and inconsistent conflation of race and place, using these self-reported racial identities as the foundation for descriptive categories is unusual because NCES make no explicit claims as to what constitutes these racial categories. Indeed, exhorting the student to 'fill in **one or more ovals**' thankfully recognizes that they are not mutually exclusive categories. Even so, despite making no claims as to the positive criteria for membership in one race or another, by using self-reporting, NCES only postpone the process of essentialization to the point at which they aggregate test results.

---

[1] One page of which is reproduced in Appendix A.

Moreover, by aggregating the scores of these self-reported groups for legal reasons and for the purpose of statistical analysis and comparison, they reify these groups and make an implicit claim that each student who identifies himself as a particular race shares some qualities in common that make them all members of that group. Furthermore, once scores have been aggregated by race, class, or gender, NCES proceed to *essentialize*—that is, treat a quality as inherent to a group—these categories by assigning to them an average score. This average score then becomes the quality that members of a particular category have in common. Individual differences in histories, background, socialization, and even in score on NAEP are elided in favor of monolithic categories associated with single identifying scores. Which is not to say that one could read back from an individual's score to learn that individual's race, class, or gender with certainty. If indeed group differences in average scores are *not* statistically significant (it is unclear from the published figures whether this is so), it would even be impossible to make a *probabilistic* guess as to group membership based on an individual's score. However, whether or not aggregate scores are significantly different between groups, the mere fact of assigning qualities to groups makes it *ideologically* reasonable to read back from an individual's score to their group. Perhaps more importantly, it is also possible to use a class' newly inscribed quality, their score that is ideologically a causal consequence of their quality of literacy and that thereby defines that class of people, to explain and legitimate the actions and situations of individuals in a given class. For example, a group might be ideologically conceived of as 'stupid', which explains their stupid behavior, which explains why they all have such low-paying jobs.

This kind of explanation relies on a *naturalized* interpretation of aggregate test scores, which necessitates essentialization as a precondition. A naturalized interpretation of a sign, such as a test score, is taking a sign that can be interpreted as sociocultural or ideological and interpreting it as natural. By 'natural' I mean that the sign-form (e.g. the test score) is interpreted as standing in a necessary causal relationship to the referent of the sign (the student who took the test), such that the characteristics of the referent (the student's 'literacy' level) cause the sign-form to take on characteristics as a consequence. To continue this example, a naturalized interpretation of an individual student's test score would look something like this:

Say a student's test score is 90 (the sign-form). In a naturalized interpretation, the sign-form (the score) is taken as a direct, necessary causal consequence of the student's (the referent) characteristics, in this case her literacy level. That is, she physically came into contact with the test and *caused* her score by inscribing her quality of literacy through her responses. Conceiving of all members of a class in an essentialized manner—

e.g. treating a racial category like 'Asian' as having in common some sort of property—permits reductive and naturalized interpretations of test scores based on group membership. Scores are thereby conceived of as *caused* by the qualities essential to an individual's class. For example, claiming that an individual has received a certain score on NAEP *because of* her race, or gender, or class misrecognizes the contingencies of her score in favor of an explanation by ideological classificatory essence, some quality which is conceived of as inherent in a class of people and as causing their manifest behavior. An individual's score, however, is *not* actually a necessary causal consequence of her literacy, but the contingent outcome of a history of socialization, life experiences, and interaction with texts, knowledge, peers, teachers, and others. As such, the test score is not a natural, cause-and-effect sign of her literacy, but instead merely a suggestion of these historical sociocultural relations.

In their report on the 4th and 8th grade 2007 NAEP reading test, *The Nation's Report Card: Reading 2007* (NCES 2007c), NCES perform each of the steps explained above. In their discussion of group test results, nowhere do they mention the origin of their classifications. By omitting the fact of self-reporting in their publications, they encourage misrecognition of these groups as somehow more 'natural' than they in fact are—they seem to come from nowhere. Racial groups are assigned average scores, broken down by year, in a full-age figure (NCES 2007c: 10). In 2007, the average reading scores for the reported groups were White 231, Black 203, Hispanic 205, Asian/Pacific Islander 232, American Indian/Alaska Native 203. They immediately proceed to 'calculate the gaps' (Foucault 1995: 190) between groups in graphics as in figure 1. Although there is a difference between the average scores of groups, it is not clear from the published figures that these differences are statistically significant. To most casual observers, then, what look in the graphs like real group differences could in fact likely be accounted for by randomness. Regardless, there is almost certainly a large amount of overlap between the individual scores in each group. On the next page, a graph shows the 'score gap' between White and Black students, reproduced in figure 1. The starred data points indicate statistically significant difference from the analogous point in 2007, but by taking for granted that the difference between groups is statistically significant, they fail to make this point explicit.

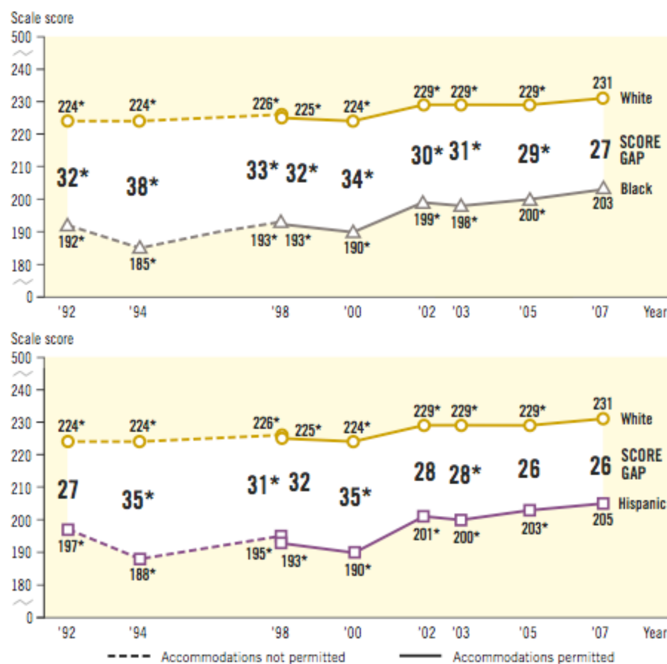


Figure 1. Trend in fourth-grade NAEP reading average scores and score gaps, by selected racial/ethnic groups (NCES 2007c: 11)
* Significantly different ($p < .05$) from 2007.

Furthermore, NCES develop an ‘item map’ as part of the *Nation’s Report Card* (NCES 2007c: 21) that both assigns ranges of scores to three performance categories—Advanced, Proficient, and Basic—as well as links literacy skills to particular scores. Any scores below 208 are classified as sub-‘Basic’ literacy, i.e. they are deficient, deviant, and abnormal according to their distribution on this hierarchical scale. Only White and Asian/Pacific Islander groups, scoring on average 231 and 232 respectively, are ‘Basic’ according to the scale. Though NCES never explicitly make this claim, a comparison of this scale to the scores of Black, Hispanic, and American Indian/Alaska Native groups suggests that, according to the criteria of NAEP, these racial groups are somehow ‘deficient’, ‘deviant’, and ‘abnormal’. Because this scale so epitomizes Foucault’s description of fixing by quantitative measurement, comparison along hierarchies, and normalizing judgments, it is reproduced in its entirety in figure 2.

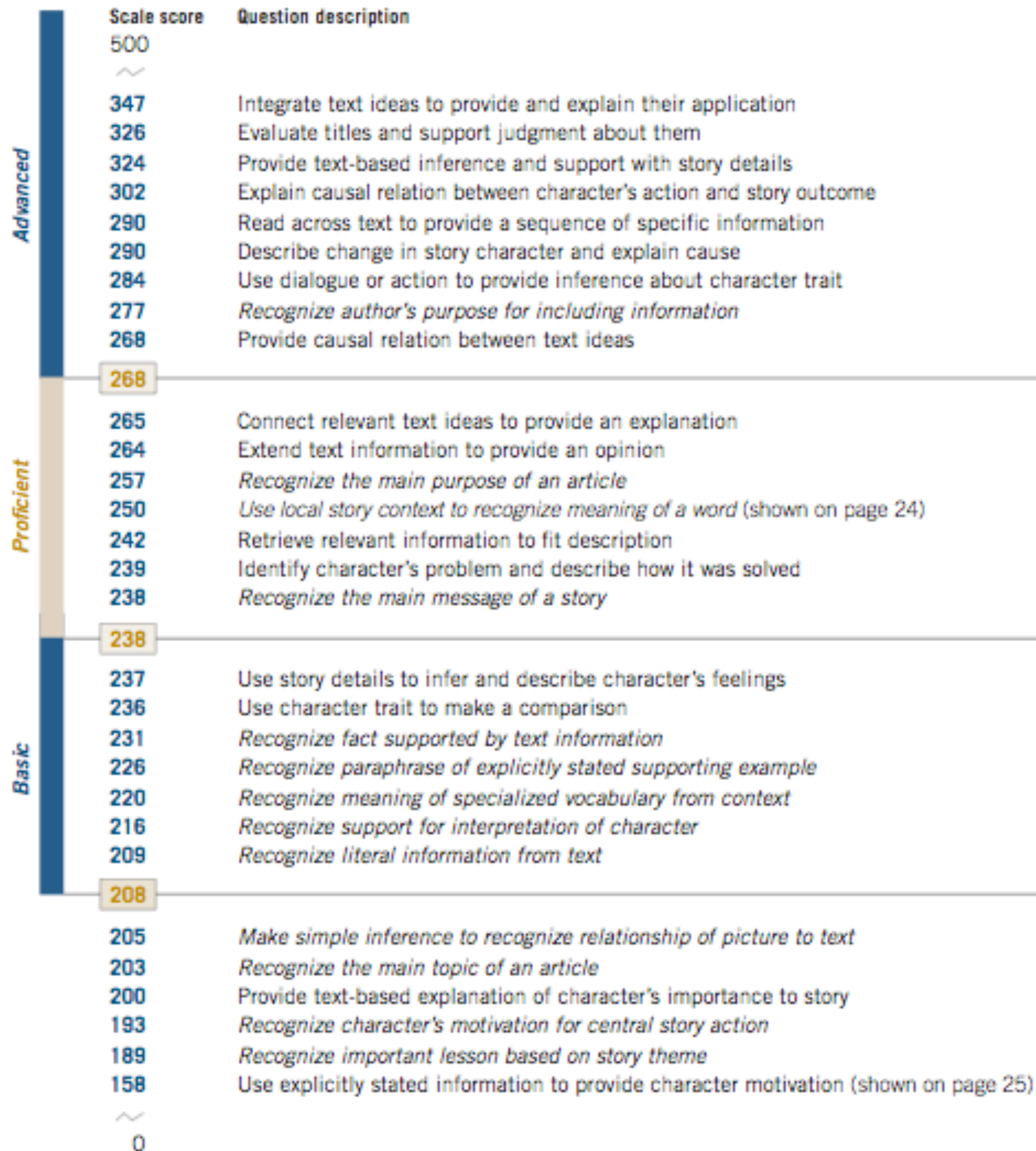


Figure 2. Grade 4 NAEP Reading Item map (NCES 2007c: 21)

NOTE: Regular type denotes a constructed-response question. *Italic* type denotes a multiple-choice question. The position of a question on the scale represents the average scale score attained by students who had a 65 percent probability of successfully answering a constructed-response question, or a 74 percent probability of correctly answering a four-option multiple-choice question. For constructed-response questions, the question description represents students' performance rated as completely correct. Scale score ranges for reading achievement levels are referenced on the map.

Earlier in *The Nation's Report Card*, NCES claim that 'as the key that allows access to many forms of knowledge and information, reading literacy is a skill critical to learning' (NCES 2007c: 4). In light of the comparison of groups above, this statement potentially has the racist implication that because Black, Hispanic, and American Indian/Alaska Native groups have, on average, a sub-'Basic' quality of literacy, they therefore lack not only knowledge and information that White and Asian/Pacific Islander groups have, but even the ability to learn that knowledge and information. Alternatively, I suggest that there are systematic ways in which members of these groups have histories that typically lead them to knowledge and information, but also to configurations of relations to text and to knowledge, that do not appear as 'literate'. Finally, each reified and essentialized racial group's newly-inscribed quality of literacy, an average number compared to the average number of each other racial group, can be used to legitimate the unequal circumstances of these groups: the stratified social positions of these groups are analogous to their aggregate score on NAEP, which score is a reflection of their quality of literacy. So, the qualities of literacy that have been newly assigned to groups that have been essentialized by the process of aggregation and reporting of NAEP scores, particularly in this case racial groups, have ideological explanatory power. That is, the tendency for members of certain groups to be in structurally subordinate social and economic classes can now be explained as an effect of their low level of 'literacy'.

Ethnographers of Education and Social Reproduction

There is a whole body of anthropological literature on education[2], but this section focuses on works that relate most to social reproduction. Unlike the structurally-focused reproduction theories discussed above, ethnographers of the lived experience of literacy learning, institutional education, and social reproduction tend to focus more on dialectics between agency and structure (e.g. Heath 1983, MacLeod 1987, Willis 1981). Shirley Brice Heath (1990) focuses specifically on relations to texts and oral performance, and

---

[2] For good overviews, see Collins, J. 1995. Literacy and Literacies. *Annual Review of Anthropology* 24:75-93; Pelissier, C. 1991. The Anthropology of Teaching and Learning. *Annual Review of Anthropology* 20:75-95 Wortham, S. 2008. Linguistic Anthropology of Education. *Annual Review of Anthropology* 37:37-51; Yon, D. A. 2003. Highlights and Overview of the History of Educational Ethnography. *Annual Review of Anthropology* 32:411-429.

shows how the home-based socialization practices of two small, lower-class communities conflict with the expectations of mainstream schoolteachers. Paul Willis (1981), while explicitly focused on social reproduction, intervenes in structural theories and explores the agency of 'the lads', who partially penetrate school ideologies but nevertheless reproduce their class of origin through their practices of resistance. Jay MacLeod (1987) examines how structural forces get mediated through life histories to inform ideological conceptions of the world and orientations toward school and dominant meritocratic ideologies.

Shirley Brice Heath

Shirley Brice Heath, in her ethnography of communication, *Ways with Words* (Heath 1983), describes the language socialization practices of two small working class communities embedded in a North Carolinian Piedmont city. She argues that white working class children of Roadville are socialized into a submissive relation to knowledge (cf. Nespor 1987). For example, they are taught to tell stories that are purely factual, chronological retellings of events. As preschoolers, they are asked to label things that appear in the books that their caretakers read to them, although after the age of three this active participation in reading stories subsides, and children are expected to sit and listen passively. The black working class children of Trackton, on the other hand, are encouraged to tell fanciful fictions that blossom from an initial germ of truth to extol one's virtues and gain attention on the plaza that is the central stage for Trackton social life. Trackton children read for functional purposes, like interacting with the mailman and shopping for groceries, but text is inextricably bound to its context: when the cereal brand name 'Kellogg's' is presented in small-capitals, in contrast to its usual looping red script, children cannot recognize the word (Heath 1983: 193). For Trackton children, reading is not about extracting either decontextualized semantico-referential meaning or a 'deep', abstract concept or theme from a text. Instead, they are socialized into reading as a means for everyday, practical ends like grocery shopping. Around the plaza, reading texts like magazines and newspapers is a wholly social event, where they are read aloud and audience reactions are encouraged.

Once in school, both Roadville and Trackton children encounter challenges as their linguistic and textual practices are not those expected and valued by their teachers. At first Roadville children excel, being able to answer information-seeking questions. But once their creativity, interpretation, and recontextualization of information are sought, they struggle. At home, they are socialized into memorization and routine, chronological retellings of events, for which they are rewarded in lower grades but which are not as

valued in the higher. Trackton children, on the other hand, have difficulty providing answers to simple information-seeking questions, and respond with creativity that is only valued at higher grades. By the time they reach those grades, however, they tend to have been sorted out of the system on a failure-track.

The patterns in these students' responses to standardized test questions would likely be striking. Roadville children would probably be able to answer information-seeking questions well and support their answers with information from the text. When required to be more interpretive and creative, however, it is probable that they would be unable to do so. In contrast, Trackton children would struggle with these information-seeking questions and would have difficulty with the decontextualized and a-social nature of NAEP questions and scoring criteria. As such, both groups of working class children would likely receive relatively low scores. As I will show, submissive relations to texts, focusing on the contextual use of written words, and a social orientation toward texts all cause problems for students taking NAEP.

Paul Willis

In his seminal ethnography of education, *Learning to Labor*, Paul Willis (1981) describes the counter-school culture of 'the lads', a working class group of male students in England. In his explanation, because the lads 'partially penetrate' their actual conditions and relations of production, they opt out of the education system by spending their school time as they please, cutting class and 'having a laff'. In contrast to the theories of social reproduction of Althusser and Bourdieu, the lads are not inculcated or interpellated *by the school* into working class subjects. Instead, through their agentive practices of resistance that are a product of their realization that school provides them nothing with which to alter their class telos, the lads reproduce their own class position. Standardized tests like NAEP are precisely the kind of school activity that the lads would attempt to subvert to their own enjoyment. Whereas their test responses and scores would be interpreted as a product of their *literacy*, they could be more rightly interpreted as an active expression of their class identity that devalues intellectual activity and champions manual labor.

In addition to the insight Willis's ethnography suggests for a potential misinterpretation of test responses and scores, it also has some implications for the effects of ideology. The lads are more prepared for shop floor culture than are their school-conforming peers, or 'ear'oles'. Besides valuing manual over mental labor, being prepared for physical and verbal toughness and aggression, and engaging in the kinds of sexually charged banter that is common in the workplace (perhaps calling each other's

girlfriends ‘shipoopis’ (DaCosta 1962)), the lads are not subject to the same grand disillusionment that the ear’oles suffer—that the qualifications they strove for and valued in school will do little for their professional advancement and that the menial jobs they will be able to get will do little to satisfy their senses of self. In such a situation as the lads’, children have no incentive to practice the skills valued on NAEP, which are perceived of as useless in the workplace. In this case, it should be no surprise that partially penetrating the meritocratic ideology that literacy and school-granted qualifications lead to economic success (when in fact they do not, as demonstrated by ear’oles’ lack of economic success) leads students to opt out of the educational system. Unfortunately, the lads’ valuation of physical over mental labor prevents them from viewing their position in the class structure as an exploitative relation of dominance, and thus prevents them from having a truly critical class consciousness that considers how reaffirming their value system reproduces the conditions of their domination. Moreover, Bourdieu and Passeron (1990) would re-analyze this partial penetration of ideology as internalizing their ‘objective probability’ of academic success and the degree to which such success could contribute to their eventual economic success.

Jay MacLeod

In Jay MacLeod’s longitudinal study, *Ain’t No Makin’ It* (MacLeod 1987), the Hallway Hangers—a group of white lower-class boys—are in a situation strikingly similar to Willis’s lads. Both groups opt out of the educational system, at least in part because of, in Willis’s words, a ‘partial penetration’ of dominant ideology: that the system is stacked against their favor when it comes to achieving economic success. This kind of opting out has been termed by Ann Ferguson (2000) ‘disidentification’, a process in which particular groups of students—namely African American boys in her work—‘actively distance and separate themselves from school as a desirable and authoritative object of identification while simultaneously embracing alternative subject positions as a means for becoming visible and gaining recognition in the social world’ (Ferguson 2000: 97). Whereas the lads choose their jobs to affirm their values of manual labor and masculinity (while devaluing mental labor and femininity), however, the Hallway Hangers are *resigned* to taking whatever jobs they can get and often blame their failure on reverse racism that favors their black peers. Just as the lads stop short of a critical class consciousness by valuing masculine labor, so too do the Hallway Hangers by blaming their failure on reverse racism, despite an ideological commitment to group solidarity that could be the basis for a counterhegemonic worldview. Thus, though both groups’ lowered aspirations lead them to lower-class jobs, as described by both Bourdieu

and Passeron (1990) and Bowles and Gintis (1976), the processes by which these low aspirations arise and the ways in which they are experienced vary.

Unlike the lads and the Hallway Hangers, the Brothers—a group of black lower-class boys—participate fully in the school system, in stark contrast to Bourdieu and Passeron's model of reproduction that would expect them to have internalized their 'objective probabilities' of economic success. Instead, they buy into a meritocratic ideology of achievement wherein diligence and school success are thought to lead to economic rewards as they attempt to compete for scarce jobs. Nevertheless, they run into problems in school comparable to the students in Heath's ethnography (see Heath 1983): different cultural capital and unknown teacher expectations create trouble for the Brothers' academic achievement. Even so, they all successfully graduate from high school. The seductiveness of the dominant meritocratic ideology is apparent when MacLeod concludes the first part of his study with the prediction that the Brothers' diplomas will in fact help them secure jobs. Instead, the economic outcomes for the Brothers, as revealed in the second part of the study, are comparable to those of the Hallway Hangers. If they are employed, they almost universally have unstable, low-paying, menial jobs.

Ultimately, MacLeod explains the class outcomes of the Brothers as a consequence of social and economic structures that are mediated by culture and agency. MacLeod attributes the unsuccessfulness of the Brothers primarily to their lack of access to informal networks that would allow them to secure jobs and to their class habitus, which is judged negatively by potential employers. Whereas at least one of the Hallway Hangers is able to use a friend's connection to get a supervisory position—which simultaneously cuts off any chance of advancement for the workers under him—the Brothers do not have access to networks of privilege, however meager, that would allow them to do the same. This situation contradicts meritocratic ideologies where hard work and individual abilities are the only things valued in hiring decisions. To the contrary, both being able to produce appropriate cultural forms (i.e. cultural capital) and having an informal network that can lead to job opportunities are critical to economic success. Focusing instead on test results obscures both of these crucially important factors in the reproduction of class.

MacLeod's study clearly describes how individuals' experiences of structural constraints vary widely and are mediated by a number of factors, including race, class, and gender. But as his sensitive and in-depth data and analysis show, it is impossible to predict how an individual will react to social structures and ideologies because subjectivity and identity cannot be neatly extricated from the histories of their production.

This insight implies that it is in fact impossible to interpret any test response accurately without a fully-fledged understanding of an individual's complex historical relations to social and economic structures and the refraction of these structures through an individual's class, racial, and gender identities. And, therefore, that any test score, whether for an individual or an aggregated group, is essentially meaningless—in that is impossible to learn something necessarily true about either an individual from his score or about a group from their aggregate score, because it completely obscures all of these factors.

Reinvigorating Reproduction Theory

Theorists and ethnographers of social reproduction suggest that the school is a powerful site for the reproduction and legitimation of hierarchical social and economic relations of production. In more recent treatments of classrooms and the school, as James Collins points out in a review article on 'Social Reproduction in Classrooms and Schools' (Collins forthcoming), anthropologists have neglected social reproduction theory because it seems too deterministic. Instead, contemporary anthropologists have privileged individual and group agency and performativity. But neglecting structure and ideology is not only analytically myopic, it dangerously suggests that practices of resistance have more power to effect structural changes than they might actually have. As many ethnographers have shown (e.g. MacLeod 1987, Willis 1981), practices that seem like resistance can nevertheless reproduce social relations. Even some recent identity-centric analyses demonstrate reproductionist effects (e.g. Bettie 2003) without, however, explicitly theorizing social reproduction. Celebrating resistance, agency, and identity while ignoring their relations to structure and ideology, then, can give the false impression that subordinate groups are both in control of their social positions and not in need of allies—academic or otherwise. But anthropology as a discipline has a long political history of championing causes of the oppressed, as in early 20th century anti-racist projects. To continue this mission, anthropologists must not neglect to theorize dialectics of structure, agency, and ideology that reproduce social relations. Theories of social reproduction should therefore be reintegrated into contemporary anthropological theory both to account more fully for observed phenomena and to avoid the potentially dangerous political implications of celebrating the agency of subordinate groups.

Chapter 3: The National Assessment of Educational Progress

Introduction

General Background

In combination, the reading and writing portions of the National Assessment of Educational Progress are intended to measure the literacy levels of 4th, 8th, and 12th graders in the United States. The data from these tests is used, among other things, to indicate the successes of No Child Left Behind (NCLB) and to compare progress across states and major ethnic groups. But 'literacy' is not such a monolithic entity as the publications of the National Assessment Governing Board (NAGB) and National Center for Education Statistics (NCES) make it out to be. By drawing on several ethnographic case studies, I suggest that how students are socialized in classrooms and elsewhere into particular relations to texts and to knowledge are highly variable, especially across class and racial groups. Through a close examination of National Assessment of Education Progress (NAEP) sample questions[3], sample responses, scorer comments, and other published material, I show that NAEP selects for a very particular standard of literacy that consistently favors the experiences, socialization practices, and relations to text and to knowledge that are typically associated with mainstream, white, upper-middleclass students in the United States.

It is important to note that my analysis relies on the public face of these organizations. Because NCES maintains tight control over the distribution and publication of test questions, my analysis is limited to those that they have already made public via their website (NCES 2007a). As I will demonstrate, even these questions are severely problematic. Because they are presumably picked for their representativeness, the remainder of NAEP is probably just as problematic, if not more so.

Theoretical Grounding

My process of close examination is strikingly similar to that of NAEP scorers. We both are engaged in evaluations of texts. We both relate test questions, student responses,

---

[3] Many questions on NAEP reading tests prompt students to respond to relatively large texts that they have read in their test booklets. The relevant texts are reproduced in Appendices B and C.

scorer evaluations, and scoring guidelines to each other, albeit in a dramatically different fashion. Part of my critique of NAEP scoring guidelines is that the interpretations of student responses they promote are not the only interpretations possible. As such, it is important to keep in mind that my interpretations of NAEP test questions, student responses, scorer comments, and the relations among them are *also* not the only ones possible. Privileging my interpretations as absolute neglects the processual nature of semiosis and makes it too easy to simply replace the NAGB's naturalized interpretation with another. Unlike NAGB, however, I am not engaged in a project of naturalization. As used by NCES, an NAEP test score is conceived of as directly indicating the literacy level of an individual student. Scoring responses through criteria that are independent of particular scorers and scoring practices elides the contingency of every student's response: their relations to texts, to knowledge, to authority, to ideologies, and the entirety of their diachronic experience including socialization, history of using texts in context, test-taking in general, and indeed their whole life. Reducing written responses overflowing with these kinds of contingencies and specifics to a numerical score that is the product of a committee's guidelines is part of the construction of a universalistic, homogeneous standard of 'literacy'.

Moreover, interpreting a test score as a necessary causal consequence of a student's literacy level is a convenient fiction that requires as a precondition this kind of monolithic, standardized literacy. This is the case because using a single scoring scale limits the characteristics of the test-taker supposedly reflected in their score to a single one as well. The NAGB in fact uses two scales, one for 'writing' and one for 'reading', both of which are understood to measure aspects of literacy. As concerns NCLB, the evaluative focus is on the NAEP math and reading tests, which schools are required to administer biennially to qualify for Title I funding. Nevertheless, the reading and writing tests measure only two aspects of a single literacy, not an array of literacies. My project, on the other hand, revolves around revealing biases, making strange NAEP and related texts, and pointing up the contingencies behind student responses and scorer judgments in order to show how NAEP systematically favors the experiences, relations to texts and to knowledge, and literate practices into which white upper-class students tend to be socialized.

Contextualization

There is currently an especially pressing need to critique standardized tests on the grounds of bias. Because of NCLB's reliance on standardized tests to impose punitive sanctions on states and districts, test scores can be used too easily to justify racist social

engineering policies that further entrench social hierarchies and inequalities. Furthermore, heavy reliance on test scores, particularly by policy makers, masks the real grounds of social and economic inequality while simultaneously legitimating the same hierarchies. The kind of critique that I am undertaking here, then, follows in the tradition of early 20th century critiques of IQ tests by anthropologists like Franz Boas, Ruth Benedict, and others.

As I show, NAEP scoring partly explicitly but primarily implicitly values a very particular kind of literacy. For example, using specific examples from the text is valued over using more general examples, though the value of direct quotation is greater as support for an opinion than for other types of responses. Instead of considering the use of text in context, scorers are always looking for students to appeal to the decontextualizable semantico-referential content of any text. Ignoring the contextual use of texts has the consequence of invalidating all socially-oriented textual practices. Emotional appeals, like telling a friend how their behavior makes the writer feel, are seen as illegitimate persuasive devices. Except when prompted by a test question, responses that include social critiques—as of consumerism—tend to fare poorly, while judgments of responses that are framed by NAEP as addressed to authority figures like Martin Luther King, Jr. problematically do not take into account students' variable relations to authorities. Moreover, throughout many evaluations runs a theme of 'literal' vs. 'interpretive', 'superficial' vs. 'underlying', scorers valuing the latter of each pair over the former. When scorers encounter responses that do not fit these criteria, the result is a negative evaluation and a lower score for that individual student. Arguably, however, few students are socialized into a particular configuration of relations to text and relations to knowledge that enable them to perform highly according to the scoring guidelines of NAEP, and it is precisely these practices that typically characterize the histories of socialization of upper-class, white, 'mainstream' children.

'Use information from the text to support your answer': What Counts as Literacy on NAEP

Specificity

Many test questions on NAEP have a common form, a question followed by explicit instructions to use information from the text to support whatever answer is given. At the very least, these instructions withhold from test-takers *what sort* of information from the text counts as evidence. NAEP scoring frequently privileges the specific over

the general. This hierarchy of evidence is clear in the scoring guide for a question on the 2007 8th grade reading test (see Appendix B for the text to which this question refers), ‘Choose two things Ellie Lammer did and explain what those things tell about her. Use examples from the article to support your answer’ (NCES 2007a). The responses to this question are sorted into four categories: ‘extensive’, ‘essential’, ‘partial’, and ‘unsatisfactory’. According to the scoring guide, extensive responses must ‘provide at least two *specific text-based examples* of things that Ellie did’, while essential responses in contrast ‘may provide a *generalization* about Ellie’s actions *without providing specific examples from the article* (e.g., Ellie Lammer dealt with the meter problem)’ (NCES 2007a, my emphasis). This hierarchization of what counts as evidence remains hidden from students throughout the test, however, and therefore distinctly advantages those students for whom providing specific textual evidence is already a part of their repertoire.

Figure 3. ‘Partial’ student response to the 2007 8th grade NAEP reading test.
‘Ellie lammer is a carful, truethful, person that does not want to be cheated and she also helped people save money from something they did not do.’

The sample response in figure 3 and the scorer comments to this response are telling. The scorer comments that this ‘partial’ response ‘provides general statements about Ellie’s character but does not focus on Ellie’s specific actions’ (NCES 2007a). If instead this response had included specifics, it likely would have been scored at least as ‘essential’. This bias in favor of specificity over generalization is problematic not only because it is indicated nowhere in the test instructions, but also because it is inconsistently applied. As will become clear in the next two sections on ‘stated generalizations’ and superficiality, there are instances in which scorers look for a deeper, underlying generalization about texts that go beyond specific textual evidence.

Stated Generalizations

Besides systematically privileging specific examples over generalizations as forms of evidence, NAEP has a more complex relation to the textual practice of citing a sentence from a text as a ‘stated generalization’, which Jan Nespor (Nespor 1987: 43)

defines as ‘explicit summaries of the meaning of the text laid out in fully formed sentences [in which students’] own paraphrases or interpretations of main ideas [are] not acceptable’. The notion of ‘stated generalization’ seems to presuppose a submissive, uninterpretive relation to texts, since only sentences directly out of the text-to-be-summarized count, not original writing on the part of students. Socialization into this submissive relation to texts puts students at a disadvantage for questions that are evaluated for appropriately deep interpretation. This practice, according to Nespor, is restricted to and associated with the context of reading for and in a school setting. If the same kind of practice were being taught in primary and secondary schools, it would be unsurprising to find attempts to provide a ‘stated generalization’ on NAEP. Moreover, if this were the *only* kind of textual interpretation learned by some students in their classrooms, as is the case in Nespor’s research, these students would struggle with many NAEP questions. The published sample responses include several examples of direct quotation that typically, though not exclusively, yield low scores. For example, in response to a question on the 2007 8th grade reading test (NCES 2007a), ‘Why did Ellie’s meter project attract so much attention? Explain why, using information from the article.’, one test-taker writes:

Ellie meter project spread fast, within a
few weeks Ellie got a call from local
politician Diane woolley

Figure 4. ‘Unacceptable’ student response to the 2007 8th grade NAEP reading test. ‘Ellie meter project spread fast, within a few weeks Ellie got a call from local politician Diane Woolley.’

As the scorer comments, this ‘response reproduces a sentence from the text that suggests that Ellie’s meter project was getting attention but does not explain why’ (NCES 2007a). This response may be a (failed) attempt to provide a stated generalization about the attention that Ellie’s meter project attracts. Without having learned other ways of approaching texts, however, this student may have had to resort to a stated generalization as their only way of relating to texts in the school. Similarly, when asked to ‘explain the narrator’s feelings about the grandmother’ in a reading passage from the 2005 12th grade reading test (see Appendix C for the full texts to which the questions and responses on the 2005 12th grade reading test refer), one student directly quotes the text:

she was twenty, green-limbed and raw, and So was this century.

Figure 5. ‘Partial or surface comprehension’ student response to the 2005 12th grade NAEP reading test.
‘She was twenty, green-limbed and raw, and so was this century.’

Despite the scorer’s false claim that this response ‘provides a description of the narrator at the beginning of the story, not her feelings about the grandmother’ (NCES 2007a), it is in fact a description of the grandmother given by the narrator, and potentially another attempt at finding a stated generalization of the narrator’s ‘feelings about the grandmother’ on the part of the student.

Not all responses that are largely direct quotations are scored so poorly as these two, however. In response to the question, ‘What is one message that [Langston] Hughes’s poem[, ‘Harlem’ (Hughes et al. 1951),] and [Martin Luther] King’s letter [(King 1963)] seem to have in common? Give examples from the letter and poem to support your answer.’, amidst paragraph-length responses, one student writes:

"justict too long delayed is justice denied." "what happens to a dream deferred?" Both are telling us that a dream deferred is a dream lost.

Figure 6. ‘Essential’ student response to the 2005 12th grade reading test.
‘ “justice too long delayed is justice denied.” “What happens to a dream deferred?” Both are telling us that a dream deferred is a dream lost.’

The main difference between this relatively successful instance of quotation and the earlier ones seems to be their fitness as stated generalizations. While the first explains *that* Ellie’s project attracted attention instead of *why,* and the second describes the grandmother instead of the narrator’s feelings toward her, the final quotations have greater potential to serve as stated generalizations about the messages of their respective texts. Furthermore, in contrast to most other NAEP questions, according to the scoring

guide this question seeks ‘a more universal or interpretive underlying theme’ (NCES 2007a), e.g. a generalization. It is not immediately clear from the prompt, however, that the commonality between the two works needs to be an ‘underlying theme’ for a response to receive the highest credit. Even so, I would suggest that stated generalizations generally fail to account for the entire text, especially in kinds of texts that lack a particular organizational structure that includes topic sentences and theses. So while this textual practice can achieve relatively high scores on NAEP, it is generally valued less highly by NAEP scorers than by the teachers in Nespor’s case study (Nespor 1987), partially because it cannot accomplish its attempted end—direct quotations often make for poor summaries.

Superficiality

There is a recurrent tension in scoring criteria between ‘literal’ or ‘superficial’ responses and ‘underlying’ or ‘interpretive’ responses. When it occurs, it is the underlying theme, the interpretive insight that receives the highest marks. James Collins (1996), in his contrastive study of low-ranked and high-ranked reading groups—which are stratified by class—in an urban third-grade classroom, shows how teachers use information-seeking questions that do not build on student responses with the low-ranked group and more dialogical, interpretive questions with the high-ranked group. This different treatment stems from a greater need for control in the low-ranked group because of more frequent internal and external interruptions of class work. This kind of tracking, therefore, entrenches these students’ perceived insufficiencies, and dooms them to failure in terms of the standards externally imposed upon them by teachers and tests. Were these students to encounter interpretive questions on NAEP, they would likely struggle.

Often, the NAEP reading test questions that seek a ‘deeper’, more interpretive answer leave this criterion implicit, and it is only revealed through scoring and scorer comments. As Collins’s study suggests, interpretive questions like ‘Some people say that [in ‘Days of Oak, Years of Salt’] the grandmother’s statement “I’m keeping in the eyes of your time” contains the key to what the story means. Do you agree or disagree with this statement? Explain why, using evidence from the story’ on the 2005 12th grade reading test (NCES 2007a) would systematically favor upper-class children who are already socialized into interpretive relations to texts. It is suggestive that 20% of all students omitted a response to this question entirely, a much higher omission rate than on any other question. If students were unprepared to present their interpretation of texts, this question would be overwhelming.

Yes, I do think this statement holds the key to
what the story means. "I'm keeping in
the eyes of your time" is saying that
shell be with her forever. The
grandmother wants the grandaughter
to know that she will always be
in her soul even after she is gone.

Figure 7. 'Essential' response to the 2005 12th grade NAEP reading test.
'Yes, I do think this statement holds the key to what the story means. "In keeping in the eyes of your time" is saying that shell be with her forever. The grandmother wants the grandaughter to know that she will always be in her soul even after she is gone.'

The scorer's comments on the response in figure 7 are instructive, commenting that this 'response provides an interpretation of the meaning of the statement but *does not explicitly connect it to the underlying theme of the story*' (NCES 2007a, my emphasis). The response is judged as 'essential' because it adequately interprets the passage but, according to the scorer, only on a superficial level. The scorer's critique of this answer is remarkably similar to part of my critique of NAEP: just as the student keeps implicit the connection between her argument and her evidence, so NAEP keeps implicit the standard of 'literacy' by which they evaluate student responses.

Decontextualization

NAEP privileges responses that emphasize the decontextualized semantico-referential content of texts. That is, explicit predications about things, divorced from context. This is particularly clear when looking at the evaluation of responses to a question on the 2005 12th grade reading test, 'What is the guide to Metro's Fares and Passes [reproduced in Appendix C] supposed to help you do?' For this question, scorers clearly value responses containing decontextualized information about the text over responses about a text's use in context.

The guide to Metro's Fares and Passes is supposed to
help you understand how to get around the city by
Knowing what the fares are, what is needed to
travel between Metrobus & Metrorail, and help
people enjoy their visit more.

Figure 8. ‘Full comprehension’ response to the 2005 12th grade NAEP reading test.
‘The guide to Metro’s Fares and Passes is supposed to help you understand how to get around the city by knowing what the fares are, what is needed to travel between Metrobus & Metrorail, and help people enjoy their visit more.’

The guide to Metro's Fares & Passes is
supposed to help you get around while you
are visiting the city.

Figure 9. ‘Partial or surface comprehension’ response to the 2005 12th grade NAEP reading test.
‘The guide to Metro’s Fares & Passes is supposed to help you get around while you are visiting the city.

Both of these responses include the purpose of the text in context, ‘to help you get around while you are visiting the city’, but the response in figure 8 adds two decontextualized details about the guide, that it is supposed to help you ‘by knowing what the fares are, [and] what is needed to travel between Metrobus & Metrorail’. The answer in figure 9, in contrast, is *exclusively* about using the text in context.

This student’s emphasis on text-in-context is similar to that of the black working-class children from Trackton in Shirley Brice Heath’s ethnography of communication, *Ways with Words* (Heath 1983). In fact, these children and their caretakers seem to sometimes privilege context over the text itself. In practice, when cereal brand name ‘Kellogg’s’ is presented in small-capitals, in contrast to its usual looping script, children cannot recognize the word (Heath 1983: 193). That explicit metapragmatic discourse about the importance of context is common in Trackton is a further indication of its value. As one Trackton woman explained (Heath 1983: 105, original emphasis)

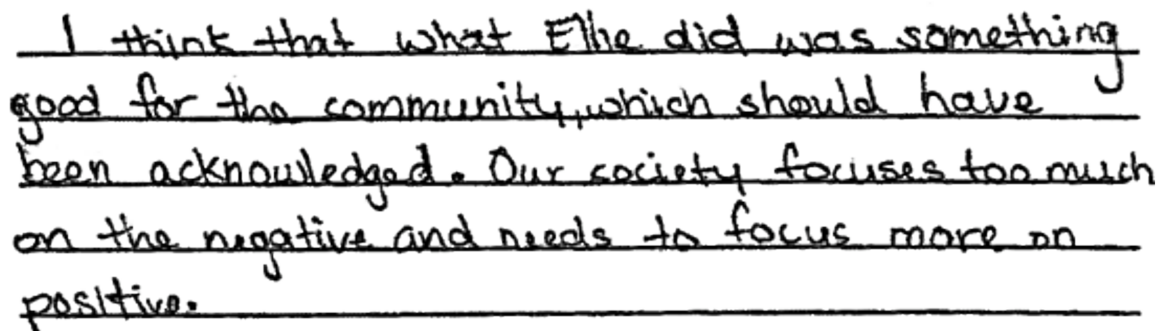
> Ain’t no use me tellin’ ‘im [a child]: ‘Learn dis, learn dat, what’s dis? what’s dat?’ He just gotta learn, gotta know; *he see one thing*

one place one time, he know how it go, see sump'n like it again, maybe it be de same, maybe it won't.

In Trackton, children are socialized into literate practices that foreground the importance of context for the interpretation of any text. If a student were focused this much on context, specifying the decontextualized semantico-referential content of a text would not seem nearly as important as test scorers take it to be. It is dangerous to make generalizations from such a local ethnography, but it seems safe to say that there is more emphasis on social use of texts in non-mainstream socialization practices.

<u>Speaking to Authority</u>

As Caroyln Marvin observes of NAEP in the introduction to her examination of the 1984 writing test (Marvin 1988), NAEP—unlike other forms of mass media—is a medium through which students can (and indeed *must*) answer back to authority. If students were to take the test as a forum to voice their opinions and complaints to authority figures (e.g. the school, teachers, administrators, standardized test writers, or whomever else students imagine read their responses), then they would not perform the 'literate tasks' expected of them, but would often instead engage in very different kinds of communication. In response, scorers would subtract points from these answers that do not fit well into their rubrics. Alternatively, for prompts that request students to frame their answers as directed toward an authority figure, divergent relations to authorities on the part of students affect their responses.



I think that what Ellie did was something good for the community, which should have been acknowledged. Our society focuses too much on the negative and needs to focus more on positive.

Figure 10. 'Partial or surface comprehension' response to the 2007 8th grade NAEP reading test.
'I think that what Ellie did was something good for the community, which should have been acknowledged. Our society focuses too much on the negative and needs to focus more on the positive.'

Some what because even if she did do something good, some kids do even better things, and they don't get credit for it. so why should she?

Figure 11. ‘Little or no comprehension’ response to the 2007 8th grade NAEP reading test.
‘Somewhat because even if she did do something good, some kids do even better things, and they don’t get credit for it. So why should she?’

The responses in figures 10 and 11 to a question on the 2007 8th grade reading test, (NCES 2007a), ‘Do you think Ellie should have become a celebrity because of what she did? Use information from the article to explain why or why not.’, are prime examples. In both, students direct social critiques to their imagined readers. The first is scored as exhibiting ‘surface comprehension’ for using a generalization rather than a specific as support. But as support for the student’s critique that ‘our society focuses on the negative and needs to focus more on the positive’, she claims that doing good deeds rather than misdeeds should be a criteria for celebrityhood, and that therefore Ellie deserves to be a celebrity. Using evidence from the article to support a critique of the constitution of ‘celebrity’ as a category seems counterintuitive. The second reply is more of a critique of the unfairness of many good deeds going unrecognized, but because it does not include ‘information from the article’, it does not fit in the scorer’s categories.

Only when prompted to be critical is this mode acceptable. For example, a critical piece is marked as ‘skillful’ in response to a prompt on the 2007 12th grade writing test (NCES 2007a),

> Many people in this country care more about the way they look and having a lot of nice things, like cars and clothes, than about the problems of society. They do not pay enough attention to problems such as crime, poverty, or unemployment.
>
> Do you agree or disagree with this statement? Write an essay in which you explain your opinion.

This prompt already contains a dissenting voice, which claims that people are more interested in consumption than societal problems, and it is therefore acceptable for students to respond critically. The ‘skillful’ response in figure 12 (only the first half is displayed here) includes a critique of media portrayals of happiness and success as overly material and superficial, consumption as a panacea, and group identity as more important than material overabundance.

When asked what a person desires
for his/her life, their answer could
easily be generalized as a material
item? Most people I know would
answer that they want a big
beautiful home, to be rich and
famous, or own the car of their
dreams In their minds, they believe
that by achieving this want, they
would achieve happiness. I am not
excluded from this accusation. At
first glance and as a hated human
instinct, to have more is to be more.
Media, I believe, plays a big part
in this issue. If you're not skinny,
you're not beautiful; if you don't have
a huge home, you're not admired; and
if you're famous, your invincible.
These are, sadley, the things that
cross our minds first and foremost.

Figure 12. First half of a 'skillful' response to the 2007 12th grade NAEP writing test.

On the other hand, when students are asked to speak directly to authority figures in their responses, their relations to the spectrum of people in similar positions cannot help but inform these responses. Marvin points out the unusualness of a question in the 1984 NAEP writing test that asks students to persuade their principal to change a school rule. As she explains, 'where different social rules govern encounters with authority for different groups of children, these are likely to condition their strategies of request' (Marvin 1988: 75). But moving from the systemic to the particular, student's imaginings

of their idiosyncratic principals likely informs not only their strategies, but also what falls into the realm of the appropriate.

Similarly, a question on the 2005 12th grade NAEP reading test (NCES 2007a) has students speaking to a figure who commands the utmost authority and respect, Martin Luther King, Jr: ‘If King were alive today, what question would you most want to ask him about his views of civil disobedience that he has not already answered in the letter [‘Letter from Birmingham Jail’ (King 1963)]? Explain why you chose this question.’ Because students frequently directly quote parts of the prompt in their replies, one student’s addition of a title indicating deference, ‘Dr. King’ for the prompt’s ‘King’, suggests immense deference and admiration. Another response, which receives no credit, defers entirely to Dr. King’s authority.

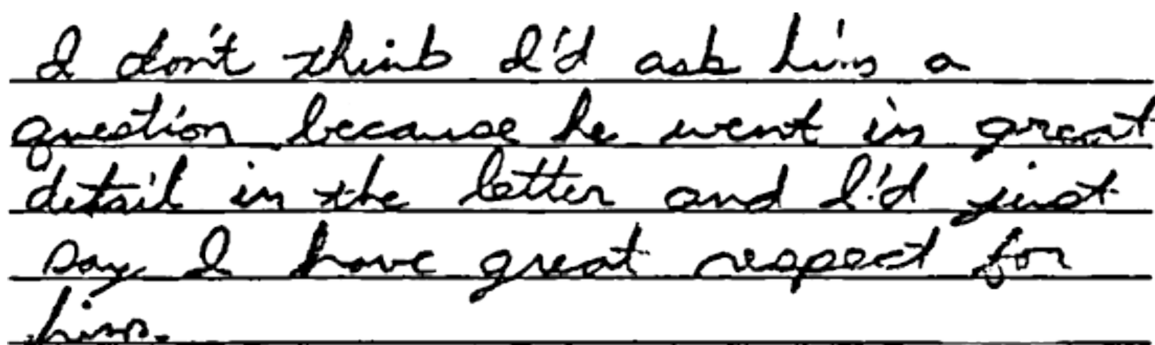


Figure 13. ‘Little or no comprehension’ response to the 2005 12th grade NAEP reading test.
‘I don’t think I’d ask him a question because he went in great detail in the letter and I’d just say I have great respect for him.’

This answer explicitly includes the student’s respect for Dr. King, and seems to prioritize this relationship of respect, deference, and admiration over any desire to ask him a question. Any attempt to interpret this response in terms of an NAEP scoring guideline completely disregards and misses any of the meaning within this statement. There are any number of purely speculative histories that could potentially account for this sort of answer, including socialization into a profound respect for authority figures in general, civil rights leaders more specifically, and Dr. Martin Luther King, Jr. in particular.

Configurations of Relations to Text, Relations to Knowledge

The description of and instructions for NAEP are fraught with unquestioned assumptions and seeming contradictions that suggest a distinct advantage for children from upper class families. When describing the ‘Making Reader/Text Connections’

aspect of reading questions in the *Reading Framework for the 2007 National Assessment of Educational Progress* (NAGB 2006), NAGB claim that 'all student responses must be text based to receive full credit. NAEP does not ask students about their personal feelings'. Immediately following this description, however, they give as an example, 'Would you have *liked* to live in colonial times? Use information from the text to support your answer.' (NAGB 2006: 8, my emphasis). This is clearly a feeling-seeking question, which is only mitigated by the instruction to 'use information from the text', a combination that challenges students to sort out which part of the prompt is more important. Questions like this are endemic throughout NAEP. A question that elicits emotional reactions from students on the 2005 12th grade reading test (NCES 2007a) asks students to 'Describe your reaction to reading the story ['Days of Oaks, Years of Salt', reproduced in Appendix C], and what was it about the story that made you react that way?' Emotional words in responses include 'sweet', 'sad', 'touching', and 'cried'. Whether NAGB consider this question to 'ask students about their personal feelings', it nevertheless elicits emotional responses. On a similar question that asks for an opinion in the 2007 8th grade reading test (NCES 2007a), 'Do you think Ellie should have become a celebrity because of what she did? Use information from the article to explain why or why not.', only 49% of answers were scored as 'full comprehension' for supporting their response 'with specific text-based information about Ellie or her accomplishments', 26% 'partial or surface comprehension' for supporting their opinion only with 'vague generalizations', and 23% 'little or no comprehension' for not using 'text-based information' as support (NCES 2007a). For a purportedly 'easy' question, very few students received full credit. For another 'easy' question on the same test that is framed not in terms of opinion but causality, 'Why did Ellie's meter project attract so much attention?', 88% received full credit.

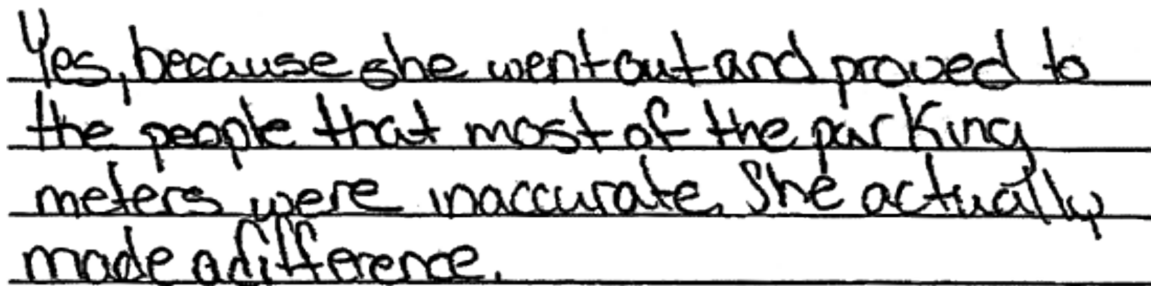


Figure 14. 'Full comprehension' response to the 2007 8th grade NAEP reading test. 'Yes, because she went out and proved to the people that most of the parking meters were inaccurate. She actually made a difference.

Yes, because if she hadn't done what
she did, many people still may have
been getting cheated by parking
meters & recieving parking tickets
today.

Figure 15. 'Full comprehension' response to the 2007 8th grade NAEP reading test.
'Yes, because if she hadn't done what she did, many people still may have been getting cheated by parking meters + receiving parking tickets today.'

Yes because she helped lots of people.

Figure 16. 'Partial or surface comprehension' response to the 2007 8th grade NAEP reading test.
'Yes because she helped lots of people.'

Yes and no because it was
great what She did but she
didn't need to be on the late
Show. I think that was to
much.

Figure 17. 'Little or no comprehension' response to the 2007 8th grade NAEP reading test.
'Yes and no because it was great what she did but she didn't need to be on the late Show. I think that was to much.'

These sample responses demonstrate that appropriate evidence merely has to accompany an opinion to receive full credit. Students here are not required to connect their evidence to criteria for celebrityhood. The response in figure 14 is specific about Ellie's achievements: she proved the inaccuracy of the meters, while that in figure 15, though not directly about the text, nevertheless specifies another of Ellie's achievements, preventing other people from being cheated and getting parking tickets. Figures 10, 11, 16, 17, are all vague about Ellie's accomplishments, describing them as doing 'something

good for the community', 'help[ing] lots of people', 'great', and 'do[ing] good'. Fully addressing the prompt requires shifting gears from answering a test question that prompts for a personal opinion with 'Do you think...?' upon encountering additional instructions to 'use information from the article to explain why or why not'. Scoring well on this kind of prompt, then, would necessitate following instructions that seem to contradict the spirit of the original question. This kind of question contradicts its seemingly feeling- and personal opinion-seeking nature with instructions to 'use information from the article', which suggests a different mode of response.

Jean Anyon's case study (Anyon 1981) that describes five different schools that vary by their class composition provide a potential explanation for varying performance on this kind of question. She defines social class as 'a series of relationships to several aspects of the process in society by which goods, services, and culture are produced' (Anyon 1981: 4). In the working-class schools, the majority of students' fathers hold unskilled or semiskilled jobs, while the parents of students in the middle-class school are generally highly skilled workers or hold managerial positions. Parents of students in the affluent professional school are usually doctors, executives, designers, or (as the name suggests) other affluent professionals, and students' parents at the executive elite school are most often vice presidents or higher of financial firms and multinational corporations.

She describes how only students at the executive elite school are socialized into a relation to knowledge in which 'knowledge results not from personal activity or attempts to make sense, but from following rules of good thought, from rationality and reasoning' (Anyon 1981: 31). Moreover, unlike the affluent professional school children who perceive themselves as having an active role in the creation of knowledge, most executive elite children, when 'asked where knowledge comes from...said, "from past experience" or from "tradition," or "other people" ' (Anyon 1981: 29), which suggests that they take a relatively passive role toward knowledge, and are therefore more likely to go to the provided text for answers. Because of the unique combination of using official modes of analysis and privileging exterior knowledge, these upper class children are likely to consistently perform well not only on problems requiring textual support, but on questions in which explicit instructions oppose the tenor of the rest of the question as well, like 'Would you have *liked* to live in colonial times? Use information from the text to support your answer.' (NAGB 2006: 8, my emphasis). This suggests that it is a matter of having, in Bourdieu's terms (Bourdieu 1977), not more or less 'cultural capital', but a particular configuration of relations to text and knowledge that can only be compared holistically with other configurations.

Experiential Bias

Some of the questions on NAEP presuppose familiarity with the kinds of experiences upper- and middle-class students are more likely to have than working-class students. Marvin (1988) highlights this situation in her examination of the 1984 NAEP writing test. As she points out regarding a question prompting students to write a note to a friend with instructions for caring for a pet, some studies have suggested that pet ownership and interaction are organized racially, with whites more likely to own and spend more time caring for pets than blacks (e.g. Merdrich 1982, Simmons 1985; cited in Marvin 1988). Whites, then, would be more likely to be able to provide the level of detail sought by NAEP scorers. Although a similar point could be made about including a rail and bus fare booklet as a textual object on NAEP, as discussed in the 'Decontextualization' section above (how, for example, would somebody from a rural area with little or no experience of cities approach it?), I focus here on science fairs and a poem by Langston Hughes.

Science Fairs

One question on the 2007 8th grade reading test (NCES 2007a) asks, 'Do you think Ellie's meter project was a "good science-fair project"? Support your opinion with information from the article.' Answers reflecting prior knowledge about science fair projects score more highly on this question.

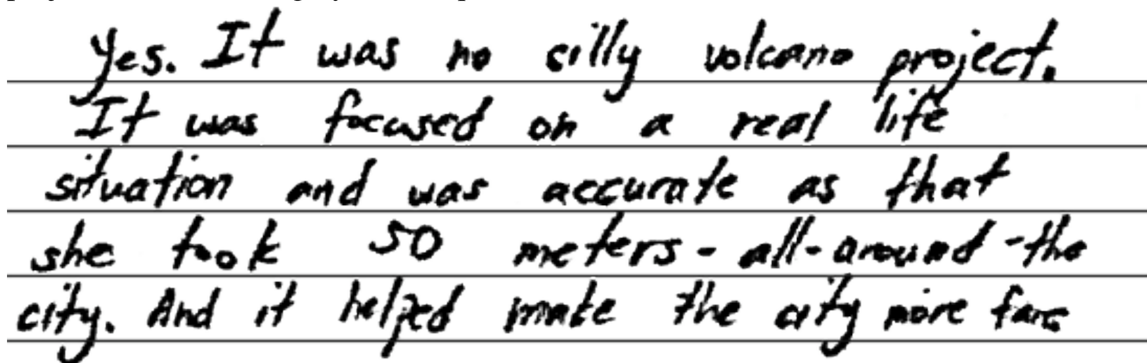


Figure 18. 'Acceptable' response to the 2007 8th grade reading test.
'Yes. It was no silly volcano project. It was focused on a real life situation and was accurate as that she took 50 meters-all-around-the city. And it helped make the city more fare.'

Figure 19. 'Acceptable' response to the 2007 8th grade reading test.
'I do think it was a "good science-fair project. This is because she did her research, spent her time working on it, and gave detailed reports. It also informed the city of what was going on.'

Figure 20. 'Unacceptable' response to the 2007 8th grade reading test.
'No, I do not think Ellies meter project was a good science-fair project because it doesn't seem very scientific to me. Although, I do think it was an amazing project, I just dont think it was ment for the science fair.'

Figure 21. 'Unacceptable' response to the 2007 8th grade reading test.
'I think it was because she worked hard and long to finish it.'

The first 'acceptable' response, in figure 18, uses a specific detail of Ellie's behavior, measuring 50 meters, as support for the opinion that her project was good for a science-fair. The 'acceptable' answer in figure 19, however, is vague and could describe any science-fair project when it claims that 'she [Ellie] did her research, spent her time working on it, and gave detailed reports'. The similarity between this 'acceptable'

response and the ‘unacceptable’ response in figure 21, ‘I think it was because she worked hard and long to finish it’, is significant. While they both describe generic projects, the scorer describes only the latter as ‘a generalization that could be made about any project’. Indeed, but the only criteria that sets the former apart is that it is a generalization that could be made about any *science fair* project. It is not specific to Ellie’s particular project—it is odd that the scorer rates the response in figure 19 so highly, because it is both general and not explicitly connected to the text. It seems, then, that the scorer takes for granted the test-taker’s knowledge of science fairs. If the student writing the ‘unacceptable’ response had little or no experience with science-fair projects (and it is easy to imagine that this might be the case), he would simply be oblivious to which aspects of Ellie’s activity are most pertinent specifically to science-fair projects, while the effort Ellie clearly put into her project would be striking to anybody without such specialized knowledge. It is also not difficult to imagine both ‘effort’ and ‘completion’ on a scoring rubric for some science fair project. In Anyon’s case study of five schools (Anyon 1981), both the affluent professional and executive elite schools use science materials that engage students in hands-on experimentation. Similar science projects are not mentioned for either the working class or the middle-class schools. It seems probable that students at schools which prepare them for independent investigation through such science projects would be more likely to have science fairs, or at least to have more rigorous criteria than merely hard work and completion, and the students at these schools tend to come from higher class backgrounds. If this situation can be generalized—and it seems like it could be, because only better-funded schools can afford such science kits—then only upper-class students would have the experience necessary to answer this question satisfactorily.

When Experience Harms

There are moments when relating seemingly relevant personal experience actually harms scores. In response to a prompt on the 2005 12th grade reading test (NCES 2007a), ‘Explain how the poem [‘Harlem’ (Hughes et al. 1951)] relates to your personal experience, or to something you have read or known about,’ one student writes:

I have read lots of poems by
Langston Hughes, some of them are
hard to understand. you need to put yourself
in the poem and figure it out inside
out.

Figure 22. ‘Little or no comprehension’ response to the 2005 12th grade NAEP reading test.
‘I have read lots of poems by Langston Hughes, some of them are hard to understand. you need to put yourself in the poem and figure it out inside out.’

Despite the scorer’s comment that this ‘response provides a personal opinion about Langston Hughes’ poetry, but never refers to the poem or connects it to personal knowledge or experience’, this response clearly *does* connect the act of reading and interpreting this poem by Langston Hughes to others he has read and endeavored to interpret in the past. But this response in unacceptable to scorers because it is unrelated to what is described semantico-referentially in the poem.

Conclusion

It should by now be clear that when it is claimed that NAEP measures ‘literacy’, it is not a common-sense notion of ‘literacy’ as the ability to read and write. Instead, this literacy is a very specific thing that could be summarized as an interpretive relation to the decontextualized semantico-referential content of texts. Moreover, the constitution of this particular literacy is not transparently available from any one source: it cannot be extracted from the questions on NAEP, nor student responses, nor scorer comments, and certainly not from their scoring guidelines. On the contrary, the implicit standard of ‘literacy’ in NAEP can only be determined through a close examination of the relationships between these test questions, student responses, scorer evaluations and comments, and scoring guides. Because the nature of the literacy valued by NAEP scorers is almost wholly implicit, it has been necessary to examine these texts closely to tease out its specific features. The configuration of relations to texts and knowledge that NAEP constructs as ‘literacy’ consists of an interpretive relation to texts in which students are socialized not only into analyzing texts for their themes, but crucially into following the rules for interpretation and understanding laid out by official authorities,

including the use of specific examples to support arguments. Evidence to support an argument must be a specific reference to the decontextualized semantico-referential content of text, though not necessarily a direct quotation. In fact, direct quotations are usually insufficient answers because NAEP seeks interpretations that are not immediately available in the text. Questions that ask about the theme of texts require a significantly more interpretative perspective. Instead of valuing students' understanding of the pragmatic, contextual use of texts, NAEP only accepts responses that refer to a text's decontextualized semantico-referential content.

A persistent theme throughout this chapter has been that the textual practices and orientations NAEP scorers consider most 'literate' are those into which upper-class, 'mainstream' children tend to get socialized and inculcated, and that other classes of people are negatively marked in comparison to this standard. If it is the case that these practices and relations are, at least in part, a product of para-school socialization—and Heath, MacLeod, and Willis (Heath 1983, MacLeod 1987, Willis 1981) all testify to this situation to varying degrees—then NAEP effectively 'retranslates' a class-rooted ethos and habitus into academic criteria of 'literacy', just as described by Bourdieu and Passeron in *Reproduction* (Bourdieu and Passeron 1990). Instead of being treated as dependent on contingent and variable processes of socialization and inculcation into relations to texts and knowledge, NAEP treats literacy as a monolithic standard (cf. Silverstein 1996). This standard, however, is by no means natural or universal; instead, following Bourdieu and Passeron, it is the very habitus of the dominant class.

By linking school success and its measures, like standardized test scores, to economic success via meritocratic ideologies such as those described by Bowles and Gintis (1976), NAEP and other test scores simultaneously legitimate and reproduce a hierarchical class structure. Test scores legitimate hierarchy by constructing a framework within which socioeconomic class can be explained as a consequence of ability or merit; the lower the score, the lower in the hierarchy the individual. In the case of NAEP aggregate scores, they can be most thoroughly used to legitimate *racial* inequality, because this is the category most fully differentiated on the student questionnaire. The argument would go something like this: on average, Whites receive higher standardized test scores than Blacks. This disparity is a consequence of the distance between these races' qualities of literacy. As such, Whites *deserve* their higher average social class than Blacks, and vice versa. Their abilities are unequal. Because society rewards individual ability, at least according to meritocratic ideologies, it is only natural that their social positions are unequal.

But this legitimation of hierarchical social relations is terribly problematic, not least because it treats inequality as caused by the essential qualities of racial groups, in this case their quality of literacy. Anything that might be called an individual's or a group's 'literacy', however, is really a judgment of contingent, variable, and socioculturally specific histories of socialization and inculcation against the equally contingent and specific histories of relations that tend to be associated with the dominant class. The difference is that the history of socialization associated with the dominant class is ideologically conceived of as a norm, the value-free universal standard against which all others are (necessarily *negatively*) contrasted.

At the same time, test scores reproduce hierarchical class structures as individuals internalize their and their class's scores not only as an 'objective' measure of their worth, but also as an indicator of their life chances. It would be falling into the meritocratic trap to claim that this functions by influencing how hard students work in school, whereby those who receive poor scores decide that they are unlikely to reap economic rewards via institutional education, and therefore opt out of the system. As MacLeod (1987) shows, other factors like connections to networks of privilege importantly influence chances of economic success. Nevertheless, individual and group scores are recognized and, in Bourdieu' and Passeron's (1990) framework, are internalized as 'objective life chances', which alters an individual's expectations, behavior, and habitus. Part of this process would be called by Althusser (2001) the inculcation of ideologies that are appropriate to an individual's position in the relations of production. So although test scores do not directly cause hierarchical social relations, they are nevertheless a component of the dialectical motor that reproduces and legitimates inequality.

Conclusion

Literacy and Ideologies of Standard

The reading and writing tests of the National Assessment of Educational Progress implicitly construct a monolithic standard of 'literacy' that can only be extracted by a close analysis of the relations between test questions, student responses, scorer comments, and scoring guides. This standard is not broadly the ability to read and write, but a specific interpretive relation to the decontextualized semantico-referential content of texts. Individual scores are naturalized as caused by an individual's quality of 'literacy'—conceived of in essentialized terms as inherent in their status as a member of a certain class of people—rather than their contingent, socioculturally specific histories of socialization. Hierarchies of aggregated group scores serve to legitimate and reproduce hierarchical social and economic class relations by articulating with a meritocratic ideology. In this ideology, socioeconomic class positions are conceived of as a consequence of meritorious qualities like literacy that are inherent to different classes of people, which justifies unequal class relations.

This situation has many parallels to the 'culture of Standard' explored by Michael Silverstein (1996). In the culture of Standard in the United States, the history of a language that could otherwise be conceived of as a struggle over language and prestigeful varieties is conceptualized as teleologically functional. Ideologically, the form of the Standard language has developed because it is a maximally precise denotational variety. Moreover, what he calls 'folk-extensionalization', or the matching of linguistic varieties to class-associated qualities, is nearly identical to what I call 'essentialization', while what he calls 'folk-intensionalization', or conceiving of these varieties as being *caused by* these qualities, is what I call 'naturalization'. In combination, extensionalization and intensionalization—or essentialization and naturalization—have the consequence of treating utterances (or test scores) as emblems of group membership. That is, an individual's group membership can be read back from their practice because this practice is conceived of as being caused by the qualities that an individual has by virtue of their class membership. The consequence of these two factors—emblematization and the ideological displacement of struggle to function—is that any non-Standard variety is conceived of *in terms of Standard.* And, because they must by definition be less ideally precise, these varieties are therefore necessarily negatively marked (as, e.g., stupid, lazy, unsophisticated, &c.). Just so with standardized tests: any response that does not match

the Standard is perceived not as different but as *deficient*. And, what is more, the groups of people, e.g. 'races', that receive low aggregate scores are perceived as having inherent, negatively marked qualities that cause them to score poorly. As Silverstein points out, because the Standard language is perceived as accessible and acquirable by all, not mastering Standard language is seen as a choice that reflects an individual's (or a class's) laziness, stupidity, &c. The same could be said for literacy: not mastering the practices of Standard literacy is seen as a consequence of an individual's (or class's) laziness, stupidity, &c.

Because standardized tests can so easily be used to justify a racist status quo and racist social engineering policies, it is surprising that anthropologists have left standardized tests uncritiqued, especially as IQ tests have been a historically important object for anthropological scrutiny. But because the discipline has shifted its focus away from reproductionist approaches and toward perspectives that emphasize performativity and agency, there has been little recent engagement with these problems. This thesis has been an attempt to redress this injustice by denaturalizing and pointing up the contingency of the reading and writing tests of NAEP, revealing the standard of literacy implicit in this test, and highlighting the points at which this test articulates with dominant meritocratic ideologies and socioeconomic structures.

What Might Be Done

Connections

In thinking about the problems of NAEP and the uses to which test scores are mobilized, I have struggled to come up with positive alternatives, much less any solutions. Though there is no immediate panacea for the problems of standardized testing laid out in this thesis, merely an awareness of these problems could help mitigate some of their potentially negative effects. Still, a number of potential improvements, connections, and further research projects have suggested themselves. If the problem of standardized tests is that they legitimate and reproduce hierarchical social structures, then it seems as though class-based society would be at the root of this problematic. At this point, revolution for a wholly egalitarian society seems unlikely, unfeasible, and perhaps impossible. Indeed, as suggested by the writings of Antonio Gramsci, the struggle is merely over *which* class is to be dominant (see Gramsci 2000). Nevertheless, there are aspects of the relations of production in late capitalism that seem to be most strongly implicated by standardized tests. The kinds of bureaucratic organizations that have such

great influence over policy and flow of wealth currently rely on technocratic means of measurement and evaluation, the standardized test and its concomitant experts being a prime example. If naturalized readings of test results are taken for granted, these tests certainly have the merit of efficiently and concisely providing information about individuals and groups. On the other hand, the current emphasis on producing individuals with standardized sets of knowledge, particularly of the relatively inflexible kind implicit in NAEP, seems to be a poor match for 'post-Fordist' economies and perhaps more suited for industrial production. If what is valued under post-Fordism is flexibility and teamwork, then limiting literacy education to inflexible relations to knowledge and asocial relations to texts seems inappropriate.

Suggestions

Standardized tests, then, could be improved along these lines by evaluating not how 'literate' an individual is, but by breaking down the category of literacy into pluralistic relations to text and to knowledge. That is, the test would not measure some sort of innate quality of ability, like 'literacy', but instead indicate a likely configuration of types of relations to text and to knowledge. When interpreting the results of this test, it would be crucial to keep in mind that the results would in no way be a necessary consequence of these relations. At the very least, such a test would go some way toward deconstructing a notion of 'literacy' as a universal, neutral standard, thus weakening the potential for test scores to legitimate hierarchical social and economic class relations.

Research Projects

Two research projects seem to have the most immediate potential for elucidating the problems laid out in this thesis. The first, of the organizations and institutions surrounding the test enumerated in Chapter 1; the second, of the training and practice of scorers themselves, on whose own texts I have heavily relied for my analysis in Chapter 3. Ethnography of these institutions could flesh out the interests, ideologies, structural constraints, and concrete practices of actors like test creators, statisticians, and government officials, as well as those of other stakeholders in the standardized testing project.

An ethnographic study of test scorer training, socialization, knowledge, and practice could also be enlightening. Like the difficulty of rationalizing scorer judgments with given answers without analyzing multiple texts, their relations, and histories of socialization, might part of the problem of scorer interpretation be scorers' situatedness and their necessarily limited perspective and knowledge? In examining how this situated

knowledge affects scorer judgments, a useful starting point might be Eve Sedgwick's insightful understanding of the ways in which ignorance can be powerful (Sedgwick 1990). Are scorers so immersed in their norms that they are unable to recognize resistance and therefore classify it otherwise, e.g. as 'illiteracy'? Or are they structurally constrained from realizing their understanding of student test responses?

Appendix A: Student Background Questionnaire

SECTION 3

Section 3

In this section, please tell us about yourself and your family. The section has 11 questions. Mark your answers in your booklet.

1. Are you Hispanic or Latino? Fill in **one or more ovals.**
   - No, I am not Hispanic or Latino.
   - Yes, I am Mexican, Mexican American, or Chicano.
   - Yes, I am Puerto Rican or Puerto Rican American.
   - Yes, I am Cuban or Cuban American.
   - Yes, I am from some other Hispanic or Latino background.

2. Which of the following best describes you? Fill in **one or more ovals.**
   - White
   - Black or African American
   - Asian
   - American Indian or Alaska Native
   - Native Hawaiian or other Pacific Islander

Appendix B: NAEP 2007 8th Grade Reading Passage

Kid Fights Cheater Meters and Wins!

The true story of a girl with a stopwatch and a bag of nickels who uncovered a local parking scandal and helped change the laws of her state . . .

Ellie Lammer wasn't trying to spark a revolt, she just wanted a haircut. That was in the fall of 1997. Ellie was 11 years old at the time, and she was getting her tresses trimmed in her hometown of Berkeley, California. When Ellie and her mom returned to their car, they found a parking ticket stuck to the windshield. It didn't seem possible: Less than an hour earlier, Ellie had pumped an hour's worth of coins into the meter. But now the needle was at zero, and Ellie's mom owed $20.

Feeling cheated, Ellie dropped another nickel in the meter and twisted the knob. The needle clicked over to the four-minute mark. Ellie stared at her watch while her mom watched the meter. Less than three minutes later, all of the time had expired. There it was: proof that they'd been cheated. The city tore up the ticket when Ellie's mom complained about the meter.

But the experience left Ellie wondering how many other meters were inaccurate. Six months later, she decided to find out. She'd been looking around for a good science-fair project—and that meter in Berkeley still bothered her. So armed with a bag of nickels and a stopwatch, she hit the streets.

Ellie didn't have the time or money to test every meter, so she focused on a sample of 50 meters located in different parts of the city. To avoid inconveniencing motorists, she did her research after 6 P.M. and on Sundays, when the meters were not in use. She put in eight minutes' worth of nickels in each meter, then measured how much time it really gave.

The results were not pretty. Ellie's findings suggested that more than nine out of every ten meters in the city were inaccurate—and that every fourth parking meter was running out of time too quickly. With 3,600 parking meters in the city, that meant a lot of undeserved tickets. As Ellie wrote in her science-project report, "I learned which meters cheat you and which meters cheat the City of Berkeley. But I learned that almost all meters cheat someone, so beware."

When the science fair rolled around, Ellie presented her findings with computer-generated charts and graphs. Her classmates weren't very interested in her project. "It's

not like they have to drive a car or put money in a parking meter," she explains. But her project was a huge hit with parents. More than 50 of them lined up that night to share their own parking-meter horror stories with Ellie.

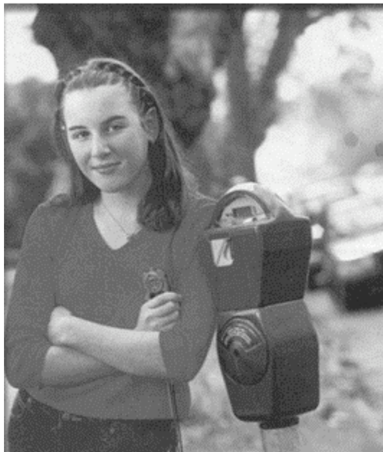
After that, word about Ellie's meter project spread fast. Within a few weeks, Ellie got a call from local politician Diane Woolley. At the time, Berkeley was considering replacing its meters with more accurate digital ones. Ellie shared her findings at city hall, and the politicians were impressed. "We don't get reports this thorough when we pay consultants hundreds of thousands of dollars," one remarked. Based on Ellie's study, they decided to purchase 2,000 new meters.

The California state legislature also decided to crack down on cheater meters. After Ellie presented her findings, they enacted "Lammer's Law," which requires California's 26 counties to test the accuracy of parking meters. Any meter found to be inaccurate must be fixed or dismantled.

California Governor Pete Wilson signed the law on November 1, 1998. At the time, he commented, "Ellie's ingenuity and dedication has earned her the gratitude of those Californians who've dug through their purses and pockets in search of exact change to feed the meters, only to return to find their cars bearing the dreaded green envelope of a parking ticket."

Ellie became a celebrity. She was in newspapers all over the country and featured on local television news during the summer and fall of 1998. CNN did a story about her. She was even a guest on the Late Show with David Letterman. "It was kind of a weird moment of being a celebrity," she says.

Ellie, who's now an eighth-grader at Martin Luther King Middle School, is proud of the work she's done. But she doesn't see meter monitoring as her life's work: "Right now I don't mind being known as the parking-meter girl, but I'm sure that later in life I'll want something different."



© 2000 by Consumers Union of U.S., Inc.
Yonkers, NY 10703-1057, a nonprofit organization.
Reprinted with permission from Zillions: Consumer Reports® for Kids for educational purposes only. No commercial use or reproduction permitted.
www.ConsumerReports.org.

Appendix C: NAEP 2005 12th Grade Reading Passages

Days of Oaks, Years of Salt

LUCIENNE S. BLOCH

My grandmother walked most of the way from a little town near Graz, in Austria, to London. She was twenty, green-limbed and raw, and so was this century: both of them restless, unshackled, upheaved from an ancient order of things into a world whose recent peace was more tentative than convincing.

Of course she did not walk alone; there were, still, vestigial proprieties in operation. Her brother, senior by a couple of significant years, accompanied her: two dark-eyed travelers seeking roomier futures than the ones they stood to inherit at home. Leaving behind three younger sisters and a widowed mother, they strolled toward the possibilities that an uncle, well settled in a woolens business in London, might provide. They carried everything on their backs, food and shoes and such, the goodbyes. At night they slept in fields, in barns when the weather turned. They picked up crumbs of new languages, mouthfuls to get by on. There is no record of this legendary journey apart from the remembered and recounted one; no documentary diaries, no franked passports, no railway or steamship ticket stubs, no hotel bills, no souvenir photographs or trinkets, no many-creased maps. Did it happen, as told? I believe so. I always believed so, although I knew the reports had been altered by the time they reached me, embroidered, translated, aggrandized, I supposed. Even so, I swallowed them whole, lured and hooked like a trout by a glitteringly fabulous fly. The adventure of it!

Taking a southerly route—longer, warmer, certainly more picturesque—my grandmother and her brother climbed into Italy through the Carnic Alps where frontiers weren't as strict as they could have been. They walked across the top of Italy, each step lighter than the one before it, springier, down to Genoa, where they followed the seductive curve of the Riviera to Marseilles, then made their way across the bottom of France to Bordeaux to board a ship for the final leg of their leisurely journey.

Upon seeing the Mediterranean and its shores for the first time, my grandmother was so amazed she took to singing, in the streets particularly. She didn't sing for money; they had all the cash they needed wrapped in handkerchiefs in their rucksacks. She sang for the pure joy of adding her note to those that hovered, purling and trilling, in the pellucid sea air. Making a musical offering to gods whose existence she hadn't even suspected, she sang folk songs in the dialect of her girlhood. Her voice, small, untrained, may have moved a heart or two. In Antibes, singing on a boulevard planted with flowering laurels, she was sketched by a man sitting on the terrace of a cafe. It could have been Matisse, we like to think; the dates and place are right. The man showed her the sketch but he did not give it to her.

My grandmother arrived in London about seven months after she commenced walking. Her cheeks were flushed, tomato-red, despite the rough Channel crossing. Long ropy muscles snaked down her legs to her narrow feet. Between them, she and her brother had gone through five pairs of what they claimed were sturdy boots, and through something less tangible, not measurable in distance covered or time elapsed. "Why did you walk? Why didn't you go on trains?" I asked her once when I was nine or so and liked the mechanics of events to be fleshed out so I could grasp them more tightly.

"I was too beauty for men in irons," she answered. "Only stars could have my shining." She was said to be "somewhat" senile, a vague qualifier for an already vague condition. But I could usually catch the drift of her scattered words. She caught my more regular ones. We understood each other.

Soon after reaching London, my grandmother made what must be seen as a brilliant match, acquiescing to arrangements set in motion by her uncle prior to her arrival. Was this match to her liking? Did her likings matter? These are conjectures. The fact appears to be that a future was perceived and undertaken by a woman whose legs may have been stronger than her spirit and whose song, it is possible, was silenced. I know what she told me, repeatedly.

"I was my dream under a lock of petals," she used to say, pointing to her wedding portrait in the snapshot album we looked at together week after week on the Saturday afternoons of my childhood; pictures were the safety net for what fell from her memory's difficult trapeze act. "Seven times I swanned around my stranger, then the glass broke awake to weeping. Salt in the mouth was my sadness to come."

Sadness? Was that the destination of her high adventure or only a stopping place, a marriage's way station?

There was no sadness in my grandmother when I saw her weekly. Or else I was too young to recognize what I saw, a fadedness of sorts, but one I felt was due to a lack of color rather than of cheer. The three rooms of her apartment were done in a variety of whites. Alabaster, ivory, off-white, cream-white, and eggshell puddled into custards on the walls and upholstery, at the silk-swagged windows, on the painted tables and bureaus and kitchen cupboards. Even the rugs on the floors were pallid, washed over the years into what was no more than a thin reminder of beiges and blues. She was blanched too: snowy hair, chalky powdered face, starched white lace and linen blouses, pearly teeth she constantly took out of her soft oystery mouth to amuse me, herself also. She'd hand me the wet dentures and say something like, "Jewels to be is on the tongue. Try me on." We laughed and laughed as I tried to clamp her false teeth between my lips like Halloween vampire fangs. All that whiteness she lived in wasn't cold, wasn't bleak; it didn't chill our times together. We played cards. We baked cupcakes. We knitted wispy mohair mufflers for the entire family. We studied the single photo album she brought to this country, and she told me stories prompted by the pictures. "In the days of oaks," she'd begin; that was her habitual opening phrase.

In my own days of oaks, Granny, there were questions I might have asked you but didn't think of then. One, especially one question haunts me now, about the one photograph you kept on your bedside table to look at all the time, not just once a week when I came to visit you and we pored over the album for clues to remembering. The photograph I want to know about, the one you didn't hide between the tooled leather covers of a book that was further hidden in a drawer between layers of your silky white underwear, is of a person you seldom mentioned to me, a man I never knew because he died in the blitz before I was born.

My grandfather struts on a seaside esplanade, straw-hatted, wearing a snappy striped blazer. His stance is jaunty. He looks extremely pleased, although there isn't a smile below his mustache. His chin points toward his left shoulder, a birdlike tilt of the head. One hand grips a silver-headed walking stick, the other is tucked into the pocket of his white flannel pants. He is a tall slim man casting a sharp pencil-slim shadow on the

paved promenade. At a distance behind him, behind a wrought-iron railing, a pier stretches across the pebbled beach and stilts into the sea. There is some kind of pavilion at the end of the pier above the water, a roofed but open-sided structure. It could have been Brighton, in August perhaps. The picture must have been taken very early in the morning, given the look and angle of his shadow. There aren't any other people in the picture, no other strollers on the broad esplanade, no children squatting at the sea's curly edge. Even in the old and faded photograph, the summer morning light is so splendid and immense it fills the image and its subject with bright importance.

What I want to know is this, Granny: Where were you? Why aren't you on his arm as in all the other vacation snaps in the album, smiling at the photographer approaching and inviting you both to pose, please? What was it about this picture you're not in that made you keep it out? Did it remind you of something you wouldn't talk about even when I asked you the questions I could then? Was that your salty sadness: his self-importance? Did he shine so sharply, absolutely, right in your eyes, dazzling you into arranging for a conspicuous absence of yourself, paling your intense promising colors until they were out of season for you? Did he white you out even then?

Dying, my grandmother's determination was vivid again; her courage as fresh as young grass. I hadn't ever seen her so lofty, almost imperious; death was a dirty penny she wouldn't stoop for. I was summoned from college to her sickroom, at home, to collect what she insisted on passing to me in person, making a physical gesture that resonated far louder and clearer than any testamental paper bell could. We had already said some of our farewells a month earlier when I was home on Christmas break, but certain matters had to be postponed until the last possible minute. She was in bed dozing, waiting for me, face powdered and cheeks rouged as though for a pleasanter outing. My kiss woke her. I couldn't see the sickness below her skin, the sly cells chewing through bone, excavating an insidious one-way tunnel. She still looked intact to me; only her dark eyes were worn, sunk deep in their sockets like eight balls dropping for end shots. I plumped up her pillows, propped her to a sitting position, and sat down on the edge of her bed. My mother left the room to take a nap, make some coffee or calls, go for a walk, get away from her mother-in-law's deathbed for the short time I was there to spell her.

"Eyes, darling eyes," my grandmother greeted me, "don't water me now, I'm for drying. Don't fear such dust. I'm keeping. I'm keeping in the eyes of your time."

I wasn't afraid, but I was crying.

She opened the drawer of her night table, took out a handful of jewelry, almost flung it in my lap, dismissing it disdainfully, such absurd little things: two gold necklaces, a diamond-studded wristwatch, a string of yellowed pearls, two rings that will never fit my thicker fingers. I thanked her. "Bauble me not!" she commanded.

Then we got down to business. She reached into the drawer for the snapshot album we passed so many afternoons with and presented it to me delicately, reverently, her thin arm floating like a ballet dancer's toward a partner, her proud head nodding up and down: yes, yes. I moved to her side, leaned back on the pillows with her, our knees bent up to form a book rest. Then we did what we'd always done, turned the pages one by one. Only this time we did it in silence because, she said, "the words cooked away before me."

Slowly, slowly, we turned the pages until she fell asleep. I sat in a chair by her bed for a while, holding my album, listening to her breathe, listening for the small song her bones, hollowed by disease, were whistling again.

Copyright 1986 by Lucienne S. Bloch. All rights reserved.

Your guide to Metro's Fares & Passes

A simple fare to anywhere.

Washington Metropolitan Area Transit Authority
600 Fifth Street, N.W., Washington, D.C. 20001
INFORMATION: 202/637-7000 (TDD 202/638-3780)
www.wmata.com

Helpful Metro phone numbers

Adopt-A-Stop 202/962-1639
Information about adopting a Metrobus stop in your neighborhood to help keep our community beautiful.

Bike-On-Rail Program 202/637-1328

Customer Assistance 202/637-1328
Register comments, complaints and suggestions.
Online www.wmata.com (Talk to Metro)
e-mail csvc@wmata.com

Customer Information 202/637-7000
Metrobus/Metrorail schedules and fares. Information about parking, senior citizen ID cards and more. Opens at 6 a.m. weekdays, 8 a.m. on weekends. Closes at 10:30 p.m. daily.
TDD 202/638-3780
Online www.wmata.com (The Ride Guide)
e-mail TheRideGuide@wmata.com

Elevator Outages/Metrorail Disruptions................................ 202/962-1212
Recorded information.

Group/Convention Sales 202/962-1326

Guaranteed Ride Home 1-800/745-RIDE
Must register in advance for a ride home in case of unexpected personal emergency or unscheduled overtime.

ID Cards for Riders with Disabilities 202/962-1245
TDD 202/628-8973

Lost and Found 202/962-1195

MetroAccess
Reservations 301/562-5360
Regional paratransit service
TDD 301/588-8186
Toll Free 1/800-523-7009
Eligibility 301/562-5361
Complaints 202/637-0128

MetroPool/Metrochek 202/962-1326

Metro Mobility Link.................... 202/962-6464
Transit information in Braille.

TripFax 202/962-1420
Metrobus and Metrorail timetables faxed to any metropolitan area location.

Transit Police 202/962-2121
Emergencies only

Web Site .... www.wmata.com

Wheelchair Lift-equipped Metrobus Service 202/962-1825

NA02100018 50.248 11/2000

Metro fare replacement policy

Metro is not responsible for lost farecards, passes or tokens/tickets. If you lose your farecard or pass while riding Metrorail, the maximum fare will be charged based on the time you exit. Farecards or passes that fall onto the tracks or into elevator shafts, etc., are considered lost.

Passes can only be exchanged *before* the pass period begins. No exchanges are made after the pass period begins. You may exchange unused farecards or passes and unopened Metrobus tokens or commuter tickets by mail and at Metro sales offices at Metro Center, the Pentagon and Metro Headquarters (600 Fifth Street, N.W.). Additional cash may be required to make the exchange at a Metro sales office because replacement farecards are only available in preset denominations.

If your farecard or pass doesn't open the faregate, show it to the station manager who will check it to see if it's damaged. If it is damaged, the station manager will give you a fare adjustment envelope. Retain the stub with the receipt number and drop the postage-paid, self-addressed envelope in any mailbox. A replacement farecard will be mailed to you. For an immediate replacement during Metro sales office hours, take the envelope to a sales office.

For more information about exchanges of tokens, tickets, damaged farecards or passes, please call:

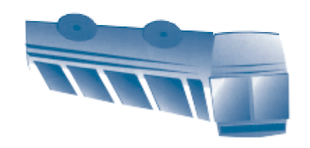
Metro Customer Assistance
202/637-1328

25807-006143 • Y73M.2

A simple fare to anywhere!

Welcome to Washington, D.C.! This guide to Metro's fares and passes will help you to get around while you are visiting the city. With Metro's simple fares, easy-to-use transfers, value-priced passes, and *SmarTrip* cards, you will save money and enjoy hassle-free travel between Metrobus and Metrorail.

By understanding how to get around on Metro's buses and trains, you will have more time to enjoy all the sites in our nation's capital.



Metrobus fares

$1.10 Regular routes

$2.00 Express routes
C11 ▸ N7 ▸ P13 ▸ P17, P19 ▸ W13 ▸ 11Y ▸ 17A, 17B ▸ 17G, 17H, 17K, 17L, 17M ▸ 18E, 18G, 18H, 18P ▸ 29E, 29G, 29H, 29X

Metrobus senior citizen and disabled rider fares

50¢ On all routes at all times for senior citizens 65 years of age and older, and for people with disabilities in possession of a WMATA ID or Medicare card. An additional photo ID card may be requested.

Special Metrobus fares

50¢* Fairfax County, Virginia
2W ▸ 3W ▸ 3Z ▸ 12C ▸ 12D ▸ 12E ▸ 12F ▸ 12L ▸ 12M ▸ 12R ▸ 12S ▸ 18R ▸ 18S ▸ 20F ▸ 20W ▸ 20X ▸ 20Y ▸ 24T

**25¢ fare is in effect on certain routes as long as state funding permits.*

60¢** Anacostia, District of Columbia
94 ▸ 95 ▸ A2 ▸ A3 ▸ A4 ▸ A5 ▸ A6 ▸ A7 ▸ A8 ▸ M20 ▸ W1 ▸ W2 ▸ W6 ▸ W8 ▸ W9 and southbound B2, P1, P2, P6 and 90 at stops at or south of Good Hope Road.

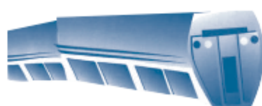
***Transfers are not issued on these routes except on payment of the full $1.10 fare.*

Metrobus school fares in the District of Columbia

A school token or ticket is good for any Metrobus fare within the District of Columbia only. Only eligible students may buy and use school tokens and tickets. Students may check with their school for program eligibility information. Rail-to-bus transfers also are free within the District of Columbia and at Silver Spring and Capitol Heights stations for eligible D.C. school students.

Children ride free

On Metrobus and Metrorail, up to two children under age 5 may travel free with a fare paying customer.



Metrorail fares

Each rider must have a farecard or pass to ride Metrorail. Fares range from $1.10 to $3.25, and are based on when and how far you travel.

Regular fares are in effect
Monday through Friday
5:30 to 9:30 a.m.
3 to 7 p.m.

Reduced fares are in effect
at all other times.

Senior citizens and people with disabilities who have Metro ID or Medicare cards may buy specially encoded $3 and $10 farecards that deduct one-half the regular fare for each trip not to exceed $1.60. Station-to-station fares are posted in every station at the kiosk and beneath the large Metrorail map display. Farecards *except Senior/Disabled and D.C. student farecards* may be purchased in stations at farecard vending or Passes/Farecards machines. Senior/Disabled farecards are available at Metro sales offices (locations are listed inside this brochure).

For more information, call:

Metro Information
202/637-7000
(TDD 202/638-3780)
Weekdays, 6 a.m. to 10:30 p.m.
Weekends, 8 a.m. to 10:30 p.m.

Plan your trip online at:

www.wmata.com (The Ride Guide)
e-mail
TheRideGuide@wmata.com

HARLEM

Since Langston Hughes began publishing poetry in 1921, when he was only nineteen years old, his productive career spanned more than four decades. "Harlem" was published in 1951, as part of a collection called *Montage of a Dream Deferred*.

Harlem

What happens to a dream deferred?

Does it dry up
like a raisin in the sun?
Or fester like a sore—
And then run?
Does it stink like rotten meat?
Or crust and sugar over—
like a syrupy sweet?

Maybe it just sags
like a heavy load.

Or does it explode?

© 1959. Langston Hughes. Reproduced with permission from Random House Incorporated.

A Letter from a Birmingham Jail *

(Description of Text)

In 1963 Dr. Martin Luther King Jr. brought a campaign for civil rights and nonviolent resistance to Birmingham, Alabama. During the demonstrations many people were arrested, and King chose to go to jail rather than obey a court order to end the public protests. While in solitary confinement, King responded to a letter sent to him by eight clergymen who urged him to cancel the demonstrations and allow the courts to negotiate the issue of civil rights.

In his response, King writes about the condition of segregation in Birmingham and justifies his cause there. He cites that “injustice anywhere is a threat to justice everywhere” to support the need for civil rights demonstrations to continue despite a court order. In doing so, King lays out a distinction between just and unjust laws in which he explains that citizens are morally obligated to resist laws that impede freedom.

King goes on to explain how he has set out to bridge the racial divide in the United States by using nonviolent protest. Instead of the “do-nothingism” of the complacent or the violent acts of hatred advocated by some, King suggests that all citizens become “extremists for the cause of justice” and demonstrate their right of civil disobedience. By doing so, Martin Luther King Jr. set the tone of the pacifist movement for civil rights for generations to come.

* [NAEP] did not receive copyright permission to put the text from the student booklet on the website. The full text can be found in King, Martin Luther Jr. (1964). Why We Can't Wait. New York: Harper & Row.

Bibliography

AIR. n.d. American Institutes for Research: Projects & Clients - Education Statistics Services Institute. <http://www.air.org/projects/projects_essi.aspx>. Accessed 2/11/2009.

Alland Jr., A., M. L. Blakey, C. L. Brace, A. H. Goodman, S. Molnar, J. P. Rushton, V. M. Sarich, and A. Smedley. 1996. The Eternal Triangle: Race, Class, and Iq. *Current Anthropology* 37:S143-S181.

Althusser, L. 2001. 'Ideology and Ideological State Apparatuses,' in *Lenin and Philosophy and Other Essays*, pp. 85-126. New York: Monthly Review Press.

Anyon, J. 1981. Social Class and School Knowledge. *Curriculum Inquiry* 11:3-42.

Benedict, R., and G. Weltfish. 1943. *The Races of Mankind. Public Affairs Pamphlets, [85]*. New York: Public Affairs Committee, inc.

Bettie, J. 2003. *Women without Class: Girls, Race, and Identity*. Berkeley: University of California Press.

Bourdieu, P. 1977. *Outline of a Theory of Practice. Cambridge Studies in Social Anthropology ; 16*. Cambridge ; New York: Cambridge University Press.

Bourdieu, P., and J. C. Passeron. 1990. *Reproduction in Education, Society, and Culture*, 1990 / edition. *Theory, Culture & Society*. London: Sage in association with Theory Culture & Society Dept. of Administrative and Social Studies Teesside Polytechnic.

Bowles, S., and H. Gintis. 1976. *Schooling in Capitalist America : Educational Reform and the Contradictions of Economic Life*. New York: Basic Books.

Collins, J. 1995. Literacy and Literacies. *Annual Review of Anthropology* 24:75-93.

—. 1996. 'Socialization to Text: Structure and Contradiction in Schooled Literacy,' in *Natural Histories of Discourse*. Edited by M. Silverstein and G. Urban, pp. 203-228. Chicago: University of Chicago Press.

—. forthcoming. Social Reproduction in Classrooms and Schools. *Annual Review of Anthropology* 38.

DaCosta, M. 1962. 'The Music Man,' pp. 151: Warner Bros. Pictures.

Davis, K., and W. E. Moore. 1945. Some Principles of Stratification. *American Sociological Review* 10:242-249.

—. 2008. Mapping America's Educational Progress 2008. <http://www.ed.gov/nclb/accountability/results/progress/nation.html>. Accessed 4/25/2009.

Falk, B. 2000. *The Heart of the Matter : Using Standards and Assessment to Learn*. Portsmouth, NH: Heinemann.

Ferguson, A. A. 2000. *Bad Boys : Public Schools in the Making of Black Masculinity. Law, Meaning, and Violence*. Ann Arbor: University of Michigan Press.

Foucault, M. 1995. *Discipline and Punish : The Birth of the Prison*, 2nd Vintage Books edition. New York: Vintage Books.

Gramsci, A. 2000. *The Antonio Gramsci Reader : Selected Writings, 1916-1935*. New York: New York University Press.

Heath, S. B. 1983. *Ways with Words : Language, Life, and Work in Communities and Classrooms*. Cambridge [Cambridgeshire] ; New York: Cambridge University Press.

—. 1990. 'The Children of Trackton's Children: Spoken and Written Language in Social Change,' in *Cultural Psychology: Essays on Comparative Human Development*. Edited by J. W. Stigler, Richard A. Shweder, and Gilbert Herdt, pp. 496-519. Cambridge: Cambridge University Press.

Herrnstein, R. J., and C. A. Murray. 1994. *The Bell Curve : Intelligence and Class Structure in American Life*. New York: Free Press.

Hughes, L., E. Pound, P. Taylor, H. H. Smythe, and M. M. Smythe. 1951. 'Harlem,' in *Montage of a Dream Deferred*. New York: Holt.

Kaufman, A. S. 1990. *Assessing Adolescent and Adult Intelligence*. Boston: Allyn and Bacon.

King, J., Martin Luther. 1963. 'Letter from Birmingham Jail.'

Klineberg, O. 1935. *Negro Intelligence and Selective Migration*. New York: Columbia University Press.

MacLeod, J. 1987. *Ain't No Makin' It : Leveled Aspirations in a Low-Income Neighborhood*. Boulder, Colo.: Westview Press.

Marvin, C. 1988. Attributes of Authority: Literacy Tests and the Logic of Strategic Conduct. *Communication* 11:63-82.

Merdrich, E. A., and e. al. 1982. 'The Serious Business of Growing Up: A Study of Children's Lives Outside School.' Berkeley: University of California Press.

Montagu, A. 1947. *Man's Most Dangerous Myth : The Fallacy of Race*. New York: Columbia Univ. Press.

NAGB. 2006. 'Reading Framework for the 2007 National Assessment of Educational Progress.' Edited by U. S. Department of Education. Washington, DC.

—. n.d.-a. National Assessment Governing Board – Policies. <http://www.nagb.org/policies/pl-index.htm>. Accessed 2/9/2009.

—. n.d.-b. National Assessment Governing Board – What We Do – How the Board Works. <http://www.nagb.org/what-we-do/board-works.htm>. Accessed 2/9/2009.

—. n.d.-c. National Assessment Governing Board – Who We Are – Overview. <http://www.nagb.org/who-we-are/overview.htm>. Accessed 2/9/2009.

NCES. 2005a. Important Aspects of *No Child Left Behind* Relevant to Naep. <http://nces.ed.gov/nationsreportcard/nclb.asp>. Accessed Nov 14, 2008.

—. 2005b. Naep – Item Scoring Process. <http://nces.ed.gov/nationsreportcard/contracts/item_score.asp>. Accessed 3/15/2009.

—. 2007a. Naep Questions. <http://nces.ed.gov/nationsreportcard/itmrls/startsearch.asp>. Accessed 11/6/2008.

— 2007b. 'Reading Student Background Questionnaire.' U. S. Department of Education.

— 2007c. 'The Nation's Report Card: Reading 2007.' U. S. Department of Education.

— 2008a. 'An Introduction to Naep.' U. S. Department of Education.

—. 2008b. Naep – the History of Naep Partners. <http://nces.ed.gov/nationsreportcard/contracts/history.asp>. Accessed 2/11/2009.

Nespor, J. 1987. The Construction of School Knowledge: A Case Study. *Journal of Education* 169:34.

'No Child Left Behind Act of 2001.' 2002. Public law *107-110*.

Pelissier, C. 1991. The Anthropology of Teaching and Learning. *Annual Review of Anthropology* 20:75-95.

Price, D. H. 2003. Outcome-Based Tyranny: Teaching Compliance While Testing Like a State. *Anthropological Quarterly* 76:715-730.

Sedgwick, E. K. 1990. *Epistemology of the Closet*. Berkeley: University of California Press.

Silverstein, M. 1996. 'Monoglot "Standard" In America: Standardization and Metaphors of Linguistic Hegemony,' in *The Matrix of Language: Contemporary Linguistic Anthropology*. Edited by D. a. M. Brenneis, pp. 284-306. Boulder, CO: Westview Press.

Simmons. 1985. 'Simmons 1985 Study of Media and Markets,' vol. P-24. New York: Simmons Research Bureau.

U. S. Department of Education. 2005. How No Child Left Behind Benefits African Americans. <http://www.ed.gov/nclb/accountability/achieve/nclb-aa.html>. Accessed 4/25/2009.

Willis, P. E. 1981. *Learning to Labor : How Working Class Kids Get Working Class Jobs*, Morningside edition. New York: Columbia University Press.

Wortham, S. 2008. Linguistic Anthropology of Education. *Annual Review of Anthropology* 37:37-51.

Yon, D. A. 2003. Highlights and Overview of the History of Educational Ethnography. *Annual Review of Anthropology* 32:411-429.