Policies of Test Centers and Jurisdictions and GED Candidate Test Performance

GED Testing Service® Research Studies, 2009-6

GED®
TESTING SERVICE

A Program of the American Council on Education®

GED Testing Service®
One Dupont Circle NW, Suite 250
Washington, DC 20036-1163
(202) 939-9490
Fax: (202) 659-8875
www.GEDtest.org

Policies of Test Centers and Jurisdictions and GED Candidate Test Performance

Amanuel Medhanie and Margaret Becker Patterson

GED Testing Service®
A Program of the American Council on Education®

Acknowledgments

This research was supported by a graduate research internship sponsored by the GED Testing Service, a program of the American Council on Education. A more technical version of this paper will be presented by the first author at the 2010 annual meeting of the American Educational Research Association in Denver, Colorado. Contact information for the first author is: Amanuel Medhanie, Department of Educational Psychology, 56 East River Road, University of Minnesota, Minneapolis, MN 55455, medha001@umn.edu. Contact information for the second author is: Margaret Patterson, Director of Research for GED Testing Service, margaret_patterson@ace.nche.edu. The authors thank three anonymous peer reviewers for their helpful comments on the technical version of the paper. We thank Wei Song and Jizhi Zhang of GED Testing Service for their review of and comments on the final paper.

Abstract

The economic and employment outlook for individuals without a high school diploma is bleak. For many of these individuals, passing the General Educational Development (GED) Test is the first step in competing in the increasingly demanding job market. GED test-taking policies vary across test centers and jurisdictions, and have the potential to affect several outcomes related to the GED credential, such as passing the test or preparedness for postsecondary education. However, little is known about this relationship. This study examines the relationship between GED policies and performance on the content areas and the GED Test as a whole.

The model that best fits the data, including test center- and jurisdiction-level predictors, explained approximately 15 percent of the variability in overall GED standard scores, which varies by content area, ranging from 10 percent for Language Arts, Reading to 17 percent for Science, suggesting that the importance of different variables differs as a function of the content area.

The results suggest that candidates of various backgrounds are at a disadvantage even after controlling for relevant candidate factors. The most consistent finding associated with test centers was that the gap in performance between African-American and white candidates was smaller in test centers that required the Official GED Practice Test (OPT). However, this requirement did not decrease the performance gap between Hispanic and white candidates.

In summary, although there is variability in overall GED standard scores and content area scores associated with the test center and jurisdiction levels, only one of the predictors at these levels—centers open all months of the year—helped account for this variation. Additional implications are discussed.

Introduction

There is currently a dearth of research on the impact of General Educational Development (GED) testing policies on test performance (American Council on Education [ACE], 2009a; Tyler, Murnane, & Willett, 2000). Jurisdictions (e.g., states, provinces, and insular areas) and GED test centers frequently establish their own policies beyond the minimum GED Testing Service policy requirements, and a lack of knowledge about the effects of such policies can adversely affect GED candidates, underrepresented populations in particular. Policies that affect the extent to which candidates develop human capital through the GED testing process, and increase the level of preparedness of GED credential recipients for postsecondary education, are likely to affect testing outcomes (including the scores candidates receive) and subsequent benefits from obtaining the credential. Although GED testing policies are likely to affect test performance, the nature of this relationship is unclear (ACE, 2008; Tyler, Murnane, & Willett, 2000). This study contributes empirical evidence to address the question of the role of GED testing policies on test performance.

The role of education in the economic outcomes of individuals in the United States cannot be overstated. The labor market continues to increase demand for highly skilled, highly educated individuals. For example, in 2006, of the 30 fastest growing jobs, 22 required at least some postsecondary education. Also, of the 30 fastest declining jobs, 28 required no postsecondary education (U.S. Department of Labor, Bureau of Labor Statistics [BLS], 2007a). Individuals with lower levels of education also are at an increased risk of unemployment. This relationship is suggested by data from the Bureau of Labor Statistics (BLS, 2008), which state that the unemployment rate in 2008 was highest for those without a high school diploma (9 percent) and decreased as education level increased. The relationship between education and economic outcomes does not stop with employment status. In 2007, the BLS (2007b) stated that individuals with lower levels of education are more likely to be classified as “working poor,” or someone who spent 27 weeks or more working or looking for work and still fell below the official poverty level. Hispanics and African Americans were more likely to fall into this category relative to whites and Asians.

In response to the increasing demand on educational attainment by the labor market, evidence suggests that the educational attainment of the labor force has been increasing over time (BLS, 2007c). The percentage of 19- to 25-year-olds in postsecondary education also has been increasing in the United States over time (NCES, 2007). Although many individuals aspire to obtain higher levels of education, there are still those who do not enroll.

The percentage of 25- to 29-year-olds that had at least a high school diploma or equivalent has been increasing since 1971 and was at about 88 percent in 2008. However, there are differences in attainment rates by ethnicity: Hispanics and African Americans had the lowest rates (68 percent and 88 percent, respectively) and Asians and whites had the highest (96 percent and 94 percent, respectively; U.S. Department of Education, National Center for Education Statistics [NCES], 2009). One characteristic of this phenomenon is that African Americans and Hispanics are less likely to complete high school. Although the percentage of people who do not complete high school decreased for all racial/ethnic groups from 1990 to 2005, this percentage

was higher for African Americans and Hispanics in 2005 (19 percent and 42 percent, respectively) than for Asians and whites (12 percent and 10 percent, respectively; NCES, 2007). This disparity suggests that Hispanics and African Americans are more likely to have poor economic outcomes as a function of high school graduation status.

One option for individuals who drop out of high school is to prepare for and successfully complete the GED Test, which is designed to assess skills and knowledge from four years of high school study in five content areas (ACE, 2009b). Although the GED credential is a viable solution for high school dropouts, there are questions regarding the economic benefits of the GED credential relative to a high school diploma (Cameron & Heckman, 1993; Tyler, 2005).

Despite disagreements about the economic benefits of the GED credential, 50 percent of candidates in 2008 reported that they took the GED Test for employment reasons (ACE, 2009a). Although evidence is mixed, some studies suggest that GED credential holders are at an advantage compared with high school dropouts without the credential. For example, a study by Murnane, Willett, and Tyler (2000) showed that GED credential holders earned, on average, 13 percent more than dropouts without the GED credential after controlling for years of schooling. However, the authors also found that the biggest benefit was for GED credential recipients that had lower cognitive skills, while those recipients with higher cognitive skills did not benefit.

Although some studies (such as Cameron & Heckman, 1993) have found that the economic benefits of the GED credential are small or nonexistent, Tyler (2001) found that differences in wages increased at a faster rate over time for GED credential recipients compared with non-recipients. Smith (2003) pointed out that some researchers suggest that the lower economic attainment of GED credential holders relative to high school diploma holders may be partly attributed to the lack of socialization GED credential holders experience, compared with high school diploma holders (e.g., lack of conformity to rules, or punctuality). This lack of socialization may be associated with problems following rules in school that many dropouts reported, especially because the GED credential process does not require socialization to the same extent as in high school.

A consistent pattern in the research on the economic benefits of acquiring the GED credential shows that many of the economic outcomes for GED holders are more positive than for high school dropouts without the credential but more negative than high school diploma holders (Boesel, Alsalam, & Smith, 1998; Song & Hsu, 2008).

Although the specific reasons for the economic benefits of the GED credential remain unknown, there are theories about how the credential affects these outcomes. Two of these prominent theories relate to human capital and signaling. The idea that the GED credential benefits economic outcomes through human capital relies on the notion that preparation for and successful completion of the GED Test allows the candidate to build skills valued in the workplace (e.g., focus, discipline, or time management). However, Tyler, Murnane, and Willett (2000) found that many GED candidates in Florida who initially failed, retested soon after, suggesting that candidates were not building human capital between test sessions. According to the signaling theory, given that the employer knows little additional information about the

applicant, the credential indicates to potential employers that the GED recipient has a set of skills that may be necessary in a job (Boesel, Alsalam, & Smith, 1998).

However, researchers have suggested that other factors may play a more important role in the relationship between high school graduation status (dropout with no credential, GED credential recipient, or high school diploma holder) and economic outcomes. For example, the number of years of schooling a student completed prior to dropping out is likely a significant predictor of economic outcomes (Boesel, Alsalam, & Smith, 1998). Cognitive ability also is thought to be a significant predictor. Boesel, Alsalam and Smith (1998) pointed to research suggesting that GED credential holders have approximately the same cognitive ability as high school graduates, while non-credentialed dropouts tend to be of lower cognitive levels.

In addition to these factors, GED credential holders often represent a different population of students from high school diploma holders. For example, the mean age for those who obtained a GED credential in the United States in 2008 was 25.1 (ACE, 2009a), although the majority of those who obtain a high school diploma do so between the ages of 16 and 19. Those who successfully completed and passed the GED Test in 2008 had spent, on average, approximately eight years out of school before taking the GED Test (ACE, 2009a).

Policies that have the potential to affect candidates' development of human capital through the GED testing process should be examined, particularly in light of test performance and potential effects on underrepresented groups. Our study contributes new evidence of the relationship between GED testing policies and test performance.

Research Questions

The question of how test center and jurisdictional policies affect test performance will be examined by focusing on two research questions:

1. How do policies (at the test centers and in jurisdictions) affect GED Test performance (on the test overall and in the five content areas)?
2. Do policies differentially affect African Americans or Hispanics?

The question of how policies affect test-taking rates (i.e., the number of GED candidates divided by the number in the target population) was considered as a third research question. However, this research question was abandoned for three reasons. First, the estimate of the target population (i.e., the number of adults without a high school diploma who are not in educational programs) by jurisdiction was based on data from the 2000 U.S. census, and it is unlikely that this number will be accurate for 2008. Second, because preliminary descriptive analyses indicate that although test-taking rates remained stable from 2004 to 2007 (perhaps partly because of the constant denominator), many policies did change within jurisdictions over that time period, suggesting that the relevant variables that explain test-taking rates are not being collected. The third reason is because of the relative lack of variability and small values for the observed test-taking rates (which ranged from approximately 0.5 percent to 5.7 percent of the target population).

Methods

Data Source

This study used retrospective data collected by the GED Testing Service. Two datasets were used for this study. The first dataset consisted of candidate and test performance data from 2008, which will yield the most accurate picture of the current state of GED test performance. The second dataset consisted of data on test center and jurisdictional policies. Although 2008 test performance is of interest, data on test center policies for research questions 1 and 2 reflect 2007 policies, as 2007 was the most recent year in which test centers provided information about their policies and characteristics.

Sample and Population

The population of interest for the assessment of policy effects on test performance consists of U.S. candidates who took and completed the GED Test. The sample consisted of U.S. candidates who took and completed the GED Test in 2008. This sample also was restricted to those who took and completed the GED in the same jurisdiction. (This consistency in location occurred for approximately 99.9 percent of U.S. test completers in 2008.) The corresponding sample consisted of 641,245 test completers. Of these people, 119 did not have test center codes, which were required for the analyses that follow, and were therefore dropped from the sample. The remaining 641,126 cases were included in the subsequent analyses. Given the very large sample size, we used a conservative statistical significance criterion, $p < 0.01$, to test effects.

Analyses

The effect of test center and jurisdictional policies on test performance is a question well suited to multilevel modeling, also referred to as *hierarchical linear modeling* (Raudenbush & Bryk, 2002). To assess the effect of policies on test performance (research questions 1 and 2), candidates are considered to be nested within test centers, which in turn are nested within jurisdictions. A primary advantage of conceptualizing research questions 1 and 2 within the framework of a three-level hierarchical linear model (HLM) is that variation in an outcome of interest can be modeled explicitly at the different levels of analysis (e.g., variation in candidate performance as a function of jurisdictional policies). Ignoring the nested nature of data can result in biased tests of parameter estimates (Raudenbush & Bryk, 2002).

Research questions 1 and 2 were addressed by examining standard scores on the GED Test and its five content areas. These outcomes were analyzed using HLM, which was performed using the open source statistical program R 2.9.1 (R Development Core Team, 2009) along with the lme4 package (Bates & Maechler, 2009). An example of the structure of the hierarchical models used is presented in **Figure 1**.

Figure 1. Example of HLM Structure

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk} gender_{ijk} + \beta_{2jk} ethnicity_{ijk} + e_{ijk}$$ } Level 1, Candidates

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01k} practice\,required + b_{0jk}$$

$$\beta_{1jk} = \gamma_{100}$$ } Level 2, Test Centers

$$\beta_{2jk} = \gamma_{200} + \gamma_{210} practice\,required_{jk} + b_{2jk}$$

$$\gamma_{00k} = \gamma_{000} + \gamma_{001} age\,requirement_{k} + g_{00k}$$ } Level 3, Jurisdictions

---

Level 1: Candidate Variables

The predictors used at the student level were gender (male=0, female=1), ethnicity (six dummy coded variables with white candidates as the reference group), primary language (three dummy coded variables with English as the reference group), highest level of education completed (none=1 to grade 12 or higher=12), hours spent preparing for the GED Test, whether the candidate reported taking the Official GED Practice Test (OPT; no=0, yes=1), whether the candidate was testing for the first time or was repeating (first time=0, repeat=1), and the number of years the candidate has been out of school.

Level 2: Test Center Variables

The predictors used at the test center level were as follows:

- The number of full-time staff.
- The number of part-time staff.
- How long the center has been testing (0=did not begin testing yet, 1=less than one year, 2=one to five years, 3=five to 10 years, 4=more than 10 years).
- Whether the center is open during all months of the year (no=0, yes=1).
- How often the center is open (0=less than four times per year, 1=once every two months, 2=once every month, 3=two times per month, 4=one time every week, 5=two to three times per week, 6=five times per week, 7=as requested).
- Whether an OPT is required (no=0, yes=1).
- Whether all content areas of the test are required to be completed in one day (no=0, yes=1).
- Whether candidates have to wait to retest (no=0, yes=1).

Level 3: Jurisdiction Variables

Jurisdiction-level predictors were: the age a person must be to attend school without exceptions, and the number of test centers in that jurisdiction. Other jurisdiction-level variables were available but could not be included because of the reasons detailed in the *Excluded Variables* section. Other variables at the jurisdiction level, such as the last three test center variables, often

varied considerably within states and were therefore more logically modeled at the test center level.

Excluded Variables

The decision to use the variables above and exclude other potential variables was made while taking into account several factors. First, the variables included in the subsequent analyses were of interest and were thought to play a role in test performance. For example, at level 1 it was thought that candidates who spend more hours preparing for the GED Test will tend to have higher scores controlling for other factors.

Many variables were excluded because of the relatively large number of missing values for those variables, or variables with sizable gaps in values. For example, although self reports of grades earned in subject areas prior to leaving school were collected, most candidates did not report this information. At level 3, when we looked at the scale for testing fees, there were not enough incremental points to allow for the assumption of a linear relationship between dollar amount and test performance.

The final factor taken into account while making decisions to exclude variables from the analyses was the extent to which variables correlate highly with one another. In the event that two or more variables correlated highly with one another, one variable was retained so that model error could be reduced and the unique contribution of a variable to the outcome of interest could be estimated.

Model Selection and Interpretation

Research question 1 was addressed by examining parameter estimates at levels 2 and 3 (test center and jurisdiction, respectively). The second research question was addressed by parameter estimates at level 2 (test center) predicting the slope for African Americans (compared with white candidates) and Hispanics (compared with white candidates). For example, referring to Figure 1, if γ_{200} reflects the effect of African-American candidates, compared with white candidates, on the overall GED Test score (and is statistically significant and negative), a statistically significant positive coefficient for γ_{210} would suggest that test centers that required an OPT tend to exhibit less of a difference between African-American candidates and white candidates.

The overall performance of the HLMs was assessed using the r-squared value for the model-predicted values and the observed values. Effects of missing data (Ruben, 1976) were assessed, and findings were cross validated (Browne, 2000). For more information, please refer to the *Missing Data* and *Cross-Validation* sections in the appendix. The Akaike information criterion (AIC), Bayesian information criterion (BIC), and nested chi-square tests were used for model selection.

Prior to fitting conditional HLMs, unconditional random intercept models were fit to each of the content area scores and the mean overall test score. These preliminary analyses allowed for an assessment of the distribution of variance in scores among all levels (jurisdiction, test

center, and candidate) through an unconditional intra-class correlation. For example, if the unconditional intra-class correlation for the Language Arts, Reading scores suggests that 4 percent of the variance in Language Arts, Reading scores is associated with the jurisdiction level, 6 percent is associated with the test center level, and 90 percent is associated with the student level, the most amount of variance that potential jurisdiction-level predictors could account for is 4 percent. A 6 percent maximum would exist for potential test center–level predictors, while candidate-level predictors could account for 90 percent of the variation in the Language Arts, Reading scores at most.

Results

Descriptive Analyses

The first step in this study was to describe characteristics of the sample with regard to variables used in the HLMs. These descriptive statistics are presented in **Tables 1** through **6**. In situations where the sample sizes (N) differ, the difference exists because data were not provided for those variables by either candidates (in Tables 1 through 3) or test centers (Tables 4 and 5).

Table 1. Descriptive Statistics of Mean Overall Test and Content Area Standard Scores

Variable	Mean	SD
Overall GED Test Score	502	72
Mathematics Score	468	77
Language Arts, Writing Score	475	104
Social Studies Score	515	83
Science Score	518	81
Language Arts, Reading Score	537	100

Note: N=641,126. SD=standard deviation. Means and standard deviations are rounded to the nearest integer, and the overall GED Test standard score is computed by taking the mean of the five content area standard scores.

Candidates in the sample tended to score lowest on the Mathematics and Language Arts, Writing content areas (mean=468, mean=475, respectively), and highest on the Language Arts, Reading content area (mean=537).

Table 2. Demographic Characteristics of Sample

Variable	Percentage	N
Gender: Male	57.34	636,483
Gender: Female	42.66	636,483
Ethnicity: White	53.13	592,407
Ethnicity: African American	23.83	592,407

Continued on next page

Table 2 continued

Ethnicity: Hispanic	18.25	592,407
Ethnicity: American Indian	2.22	592,407
Ethnicity: Asian	1.78	592,407
Ethnicity: Native Hawaiian	0.77	592,407
Ethnicity: Other	0.02	592,407
Primary Language: English	93.21	541,951
Primary Language: Spanish	5.87	541,951
Primary Language: French	0.45	541,951
Primary Language: Other	0.47	541,951

The majority of candidates were male (57.3 percent) and white (53.1 percent). The second most common ethnicity was African American (23.8 percent) followed by Hispanic (18.3 percent). For the majority of candidates, English was their primary language (93.2 percent).

Table 3. Descriptive Statistics of Performance-Related Variables

Variable	Mean	SD	N
Highest Grade Completed	10.1	1.5	555,080
Years Out of School	7.4	8.9	488,637

Note: Highest Grade Completed: None=1 to grade 12 or higher=12.

The average highest grade completed by candidates in this sample was approximately 10th grade (mean=10.1, standard deviation=1.5). Candidates spent an average of 7.4 years out of school prior to testing. Of candidates in 2008 who took and completed the GED Test, approximately 73 percent passed the test. Repeat test-takers made up 16.5 percent of the sample.

Table 4. Descriptive Statistics on Test Center Characteristics and Policies

Variable	Mean	SD	N
Number of Full-Time Employees	0.9	1.7	2481
Number of Part-Time Employees	1.8	2.0	2481
How Long the Center Has Been Testing	3.7	0.8	2441

Continued on next page

Table 4 continued

How Often the Center Is Open	4.3	2.0	2430

Note: How Long the Center Has Been Testing: 0=did not begin testing yet, 1=less than one year, 2=one to five years, 3=five to 10 years, 4=longer than 10 years. How Often the Center is Open: 0=less than four times per year, 1=once every two months, 2=once every month, 3=two times per month, 4=one time every week, 5=two to three times per week, 6=five times per week, 7=as requested.

The average GED test center had one full-time employee and two part-time employees. The chief examiner was more likely to serve on a part-time basis (54.5 percent) than full-time (45.5 percent) and had served an average of 6.5 years. The chief examiner personally administered the test approximately 22 times per year.

The test center typically had been open at least 10 years and was open at least once per week, on average. Tests were administered roughly 46 days per year. Most test centers (76.5 percent) were not open on weekends.

Table 5. Descriptive Statistics of Test Center Characteristics and Policies

Variable	Percentage	N
Open All Months	67.6	2481
Official GED Practice Test Required	34.9	2434
GED Test Completion Required in One Day	5.4	2438
Waiting Period Required Before Retesting	52.1	2336

Note: Each variable was coded as “0” for having met the condition and “1” for not meeting the condition; for example, if a center was open all months of the year, it would be coded 1, and if not, it would be coded 0.

Table 6. Descriptive Statistics for Jurisdictions

Variable	Mean	SD	N
Age of Required School Attendance Without Exceptions	16.9	1.0	51
Number of Test Centers	53.7	53.3	51

Note: Age of Required School Attendance Without Exceptions only took on values 16, 17, or 18.

Approximately 67 percent reported taking an OPT prior to taking the GED Test. Also, approximately 68 percent of centers were open during every month of the year (Table 5). Although most students in the sample took an OPT, only approximately 35 percent of test centers required the OPT. Very few test centers required the candidate to complete the GED Test in one day (5.4 percent). However, approximately 52 percent of centers required candidates to wait a certain number of days prior to retesting.

Although not presented in the tables above, most test centers (nearly 85 percent) relied exclusively on test fees to fund the test center. The average test fee for the first administration of the GED Test was $64, with retesting fees averaging $14 per content area retaken.

Inferential Analyses

Unconditional Standard Score Models

Prior to fitting the HLM for the mean overall GED standard score using candidate-, test center–, and jurisdiction-level variables, a three-level fully unconditional random intercept model was fit to the mean overall GED standard score. The fixed effects for this unconditional model are presented in **Table 7**. According to Table 7, a randomly selected 2008 GED Test candidate would have a mean overall GED standard score of nearly 515 in the three-level model with no predictors.

Table 7. Fully Unconditional Three-Level HLM of Mean Overall GED Standard Score

Effect	B	SE B
Intercept	514.7***	2.6

Note: *=p<0.05, **=p<0.01, ***=p<0.001. The B estimate reflects the predicted value for a candidate pulled at random from the sample without taking into account any predictors.

Table 8 reports the amount of variation associated with both the test center level and jurisdiction level (and the sum of these values) for the mean overall GED standard score for the unconditional model (i.e., the three-level model with no predictors). Approximately the same amount of variation was associated with both levels. The total amount of variation in the mean overall GED standard score associated with either test center level or jurisdiction level, as measured by the intra-class correlation (Raudenbush & Bryk, 2002), was approximately 11.5 percent. This suggests that, at most, the test center and jurisdiction characteristics and policies that will be included in the model can account for approximately 11.5 percent of the variance in the mean overall GED standard score. The remaining 88.5 percent of variance is between students.

Table 8. Percentage of Variation of Mean Overall GED Standard Score by Jurisdiction and Test Center

Percentage of Variation for Fully Unconditional HLM			
Variable	Jurisdiction (%)	Test Center (%)	Total (%)
Overall GED Standard Score	5.6	5.9	11.5

Conditional HLM for Mean Overall GED Standard Score

The next step in this study was to assess the effect of various predictors on the test scores. The estimated fixed effects for the final HLM predicting the GED Test score are presented in **Table 9**. To answer research question 2 (do policies differentially affect African Americans and Hispanics?), the conditional random intercept model with only two random effects (test center and jurisdiction) for the GED standard score was compared with a model that allowed the African-American ethnicity effect on the overall GED standard score to vary across test centers. Then, the model that allowed for the African-American effect to vary across test centers was compared with a model that allowed both the African-American effect and the Hispanic effect to vary across test centers. The results of these model comparisons suggested that the model that allowed for the African-American and Hispanic effects to vary across test centers fits the data best. These model comparisons were evaluated using both the chi-square deviance test and the model fit indices, AIC and BIC.

Table 9. Fixed Effects for Mean Overall GED Standard Score

Effect	B	SE B
Intercept	515.0***	2.5
Candidate Predictors		
Gender: Female	-3.7***	0.4
Ethnicity: Other	-32.9*	15.2
Ethnicity: Hispanic	-25.7***	0.7
Ethnicity: American Indian or Alaska Native	-24.7***	1.2
Ethnicity: Asian	-14.7***	1.5
Ethnicity: African American	-44.6***	0.6
Ethnicity: Native Hawaiian or Pacific Islander	-23.8***	2.2

Continued on next page

Table 9 continued

Primary Language: French	-34.1***	4.6
Primary Language: Spanish	-23.2***	1.0
Primary Language: Other	-31.9***	1.8
Highest Grade Completed	5.8***	0.1
Preparation Hours	-0.0***	0.0
Reported Taking Official GED Practice Test	-3.1***	0.4
First-Time Test-Taker or Repeat	-44.7***	0.5
Years Out of School	-0.0	0.0
Test Center Predictors		
Number of Full-Time Employees	-0.2	0.3
Number of Part-Time Employees	0.3	0.2
How Long the Center Has Been Testing	1.0	0.7
Open All Months	1.9	1.1
How Often the Center Is Open	0.1	0.3
Official GED Practice Test Required	-1.0	1.3
Test Completion Required in One Day	-0.3	2.4
Waiting Period Required Before Retesting	-0.5	1.2
Jurisdiction Predictors		
Age of Required School Attendance Without Exceptions	-0.3	2.6
Number of Test Centers	-0.1	0.1
African-American Slope		
Number of Full-Time Employees	0.2	0.3
Number of Part-Time Employees	-0.2	0.3
How Long the Center Has Been Testing	0.1	0.9
Open All Months	1.5	1.6
Official GED Practice Test Required	6.0***	1.9
Test Completion Required in One Day	1.3	2.9

Continued on next page

Table 9 continued

Waiting Period Required Before Retesting	2.6	1.6
Hispanic Slope		
Number of Full-Time Employees	0.3	0.3
Number of Part-Time Employees	-0.2	0.3
How Long the Center Has Been Testing	0.1	1.1
Open All Months	2.6	1.7
Official GED Practice Test Required	-0.4	2.3
Test Completion Required in One Day	3.2	4.2
Waiting Period Required Before Retesting	0.7	1.9

Note: Gender: Male=0, Female=1; Ethnicity was dummy coded with white candidates as the reference group; Reported Taking Official GED Practice Test: No=0, Yes=1; First-Time Candidate or Repeat: First-Time=0, Repeat=1; Open All Months: No=0, Yes=1;, *=<0.05, **=<0.01, ***=<0.001.

Predictors at the student level were group mean–centered (i.e., centered around the group mean within test centers). Predictors at the test center level were also group mean–centered (within jurisdictions), and jurisdiction-level predictors were grand mean–centered (i.e., centered around the grand mean at all levels). Therefore, the estimate of the intercept (for the overall test standard score) in Table 9 (B=515.0) represents the estimated standard score for a typical candidate in a typical test center from a typical jurisdiction (i.e., when the observed value matches the respective mean for that variable, the difference is zero). The interpretation of the fixed effect estimates for candidate, test-center, and jurisdiction characteristics and policies are with respect to the means. Therefore, a candidate who had a highest grade completed of one level higher than the average within his or her test center is predicted to have a standard score 5.8 points higher compared with a candidate who had the average highest level of education within his or her test center. The interpretation for dichotomous predictors, such as ethnicity, gender, and whether a test center was open all months of the year, is as normal. For example, on average, a female candidate predictably would lose 3.7 points from the standard score compared with a male.

The results from the HLM on the mean GED Test score suggest that non-white candidates had lower scores on the GED Test overall. Also, candidates whose primary language is not English had lower overall standard scores than those whose primary language is English. The length of time a candidate had been out of school was unrelated to the score on the test (B=-0.0, $p>0.05$). Although African Americans had lower overall standard scores controlling for other predictors, this gap decreased in test centers that required an OPT (B=6.0, $p<0.001$). This mitigating effect was not found for Hispanic candidates. As a whole, this model (including all predictors) accounted for approximately 15 percent of the total variation in the outcome (mean overall GED standard score). The remaining 85 percent includes candidate characteristics and

other policy-related predictors that we did not measure, either at the test center level or jurisdictional level.

As an example of obtaining a predicted value for a particular student, imagine predicting the standard score for an African-American female who took the GED Test in an average test center, which required an OPT, in an average jurisdiction. To obtain this predicted value, one would simply include the characteristics of the person, test center, and jurisdictions of interest in the model presented in Table 6. For this condition, the equation would look like this:

$$\hat{Y}_{ijk} = \beta_{0\,jk} + (-3.69)gender_{ijk} + (-44.58 + 6.02)AfricanAmerican_{ijk} + (-3.13)tookpracticetest_{ijk} + (-1.01)practicerequired_{jk}$$

$$\hat{Y}_{ijk} = 515 + (-3.69)*1 + (-38.56)*1 + (-3.13)*1 + (-1.01)*1$$

$$\hat{Y}_{ijk} = 468.61$$

Therefore, her predicted overall GED standard score would be nearly 469.

GED Standard Scores by Content Area

Prior to fitting the HLMs for the content area scores using candidate-, test center-, and jurisdiction-level variables, five separate fully unconditional random intercept models were fit to the content area standard scores. The fixed effects for these unconditional models for the content areas are presented in **Table 10**. Each column in Table 10 contains an estimate of the unconditional intercept and the corresponding standard error for a content area. Although the estimated intercepts for each of the content areas are similar to the means reported for U.S. candidates in the *2008 GED Testing Program Statistical Report* (and represent the predicted value for a candidate picked at random, on a given content area), they are not of primary interest here. Instead, this initial step allowed for the assessment of variation in scores (for each content area) associated with test centers and jurisdictions. **Table 11** reports the amount of variation associated with both the test center level and jurisdiction level for each content area in descending order. In general, approximately the same amount of variation was associated with both levels. The total amount of variation in content area scores associated with either the test center level or the jurisdiction level, as measured by the intra-class correlation (Raudenbush & Bryk, 2002) ranged from approximately 6 percent to 10 percent, depending on the content area.

The Language Arts, Writing and Science content areas had the lowest and highest percentage of variation associated with either the test center or jurisdiction, respectively (5.7 percent and 10.5 percent, represented in the "Total" column of Table 11). That is, 5.8 percent (2.8 percent at the jurisdiction level plus 3.0 percent at the test center level) of the variation in the Language Arts, Writing content area standard score was associated with either the test-center or jurisdiction level. Therefore, potential predictors at either the test-center or jurisdiction level can only account for a maximum of 5.7 percent of the variation in the Language Arts, Writing content area score. In contrast, the combination of either test center– or jurisdiction-level predictors have the potential to account for a maximum of 10.5 percent of the variation in the

Science content area score. This suggests that test-center and jurisdiction policies or characteristics are likely to affect the Science content area more than the Language Arts, Writing content area.

Table 10. Fully Unconditional Three-Level HLMs of Standardized Content Area Scores

	Language Arts, Writing	Social Studies	Science	Language Arts, Reading	Mathematics
Effect	B	B	B	B	B
Intercept	479.2*** (2.5)	523.6*** (2.5)	528.9*** (2.7)	546.6*** (2.8)	476.1*** (2.4)

Note: Standard Errors are in ( ). *=p<0.05, **=p<0.01, ***=p<0.001. The B estimates reflect the predicted value for a student pulled at random from the sample without taking into account any predictors.

Table 11. Percentage of Variation of Content Area Standard Scores by Jurisdiction and Test Center (for Fully Unconditional Three-Level HLMs)

	Percentage of Variation		
Subject	Jurisdiction	Test Center	Total
Science	5.3	5.1	10.5
Mathematics	4.8	4.8	9.6
Social Studies	4.3	4.5	8.7
Language Arts, Reading	3.7	3.6	7.3
Language Arts, Writing	2.8	3.0	5.7

Again, because one of the goals of this study was to answer research question 2, each of the conditional random intercept models (one for each content area) without the additional random effects for the African-American and Hispanic slopes were compared with models that allowed for random effects for the African-American and Hispanic slopes (same process as the overall standard score model). The results suggested that the models that allowed these slopes to vary across test centers fit the data best. Therefore subsequent conditional models included random effects for the African-American and Hispanic slopes. For the sake of simplicity, **Table 12** contains only the fixed effect estimates for these models (for each of the five content areas) while **Tables A.1** through **A.3** in the appendix contain both the fixed effect estimates and associated standard errors.

Table 12. Fixed Effects for Content Area Scores from Sample 1

Effect	Language Arts, Writing	Social Studies	Science	Language Arts, Reading	Mathematics
Intercept	484.10***	527.00***	532.80***	551.10***	479.80***
Candidate Predictors					
Gender: Female	25.73***	-16.01***	-24.24***	12.44***	-16.55***
Ethnicity: Other	-0.10	-46.86*	-36.14*	-41.80	-40.12*
Ethnicity: Hispanic Origin or Descent	-19.07***	-26.81***	-34.12***	-26.26***	-23.43**
Ethnicity: American Indian or Alaska Native	-23.05***	-23.69***	-27.61***	-24.29***	-24.93***
Ethnicity: Asian	-6.78**	-20.31***	-26.17***	-28.50***	8.80***
Ethnicity: African American	-31.62***	-48.12***	-56.21***	-43.77***	-44.28***
Ethnicity: Native Hawaiian or Pacific Islander	-7.93*	-29.25***	-33.64***	-32.01***	-16.11***
Primary Language: French	-47.67***	-34.52***	-37.19***	-51.13***	-0.41
Primary Language: Spanish	-18.20***	-29.63***	-20.30***	-45.25***	-2.43*
Primary Language: Other Language	-47.29***	-34.33***	-33.70***	-58.36***	14.26***
Highest Grade Completed	6.58***	6.21***	5.01***	5.30***	6.09***
Preparation Hours	-0.02***	-0.02***	-0.02***	-0.02***	-0.02***
Reported Taking Official GED Practice Test	3.09***	-6.93***	-5.32***	-5.01***	-1.67***
First-Time Test-Taker or Repeat	-33.02***	-47.10***	-47.75***	-57.17***	-38.40***
Years out of School	-0.57***	0.85***	0.07***	0.63***	-1.08***
Test Center Predictors					

Continued on next page

Table 12 continued

Number of Full-Time Employees	0.19	-0.02	-0.32	0.18	-0.65*
Number of Part-Time Employees	0.63*	0.21	0.05	0.16	0.21
How Long the Center Has Been Testing	1.25	0.97	1.50*	.00	1.39*
Open All Months	1.73	1.14	2.09	2.12	3.13**
How Often the Center Is Open	0.20	-0.05	-0.18	0.11	0.03
Official GED Practice Test Required	-0.73	-1.11	-2.12	0.78	-0.98
Test Completion Required in One Day	1.09	-1.59	0.70	-3.08	2.21
Waiting Period Required Before Retesting	0.14	-0.30	-0.82	-0.66	-0.52
Jurisdiction Predictors					
Age of Required School Attendance without Exceptions	1.69	-0.38	-1.65	-0.91	0.27
Number of Test Centers	-0.05	-0.09	-0.11*	-0.10	-0.10*
African-American Slope					
Number of Full-Time Employees	0.42	0.36	-0.20	0.23	0.29*
Number of Part-Time Employees	-0.40	-0.29	-0.30	-0.14	.00
How Long the Center Has Been Testing	-0.42	0.01	0.67	0.17	0.01
Open All Months	4.65*	-0.03	1.25	0.70	0.98
Official GED Practice Test Required	8.84**	7.08**	2.54	8.36**	4.05*
Test Completion Required in One Day	3.46	1.19	1.17	-1.17	3.28
Waiting Period Required Before Retesting	5.24*	2.83	2.79	2.48	0.36
Hispanic Slope					

Continued on next page

Table 12 continued

Number of Full-Time Employees	0.63	0.21	-0.31	0.52	0.28
Number of Part-Time Employees	-0.40	-0.20	-0.25	0.02	-0.09
How Long the Center Has Been Testing	-1.21	1.07	0.40	-0.90	0.50
Open All Months	5.13	2.74	1.74	0.49	2.42
Official GED Practice Test Required	0.16	0.13	-1.02	-0.46	-0.95
Test Completion Required in One Day	-0.01	2.52	4.62	3.53	4.54
Waiting Period Required Before Retesting	0.37	0.70	1.09	0.73	0.21

Note: Gender: Male=0, Female=1; Ethnicity was dummy coded with white candidates as the reference group; Reported Taking Official GED Practice Test: No=0, Yes=1; First-Time Test-Taker or Repeat: First-Time=0, Repeat=1; Open All Months: No=0, Yes=1; *=<0.05, **=<0.01, ***=<0.001.

No jurisdiction-level predictors met our significance criterion ($p<0.01$). A single test center–level predictor, *Open All Months*, was statistically significant. In test centers that are open all year, an average candidate could expect an increase of three points in the Mathematics content area. This variable was not statistically significant for other content areas.

The candidate-level results suggest that females tended to perform better than males in Language Arts, Writing, and Language Arts, Reading (B=25.7, B=12.4, respectively), and performed worse in Social Studies, Science, and Mathematics (B=-16.0, B=-24.2, B=-16.6, respectively), controlling for other predictors. In general, non-white candidates tended to perform worse than white candidates in all content areas (as indicated by the negative coefficients for each of the ethnic categories, which are compared with white candidates, across content areas). However, Asian candidates tended to perform better than white candidates in Mathematics (B=8.8), controlling for other predictors.

The results also suggest that those whose primary language was something other than English tended to perform worse than individuals whose primary language is English. This effect appears to be less detrimental in Mathematics. A *Repeat* candidate status was associated with a lower score in each of the content areas, ranging from 33 to 57 standard score points lower on average. Candidates who completed a higher grade than average could expect an additional five to seven standard score points per additional grade completed.

Although African Americans had lower scores than white candidates on all content areas, this disadvantage was less pronounced in test centers that required an OPT. The effect of whether an OPT was required on the gap between African-American and white candidates varied across

content areas (as indicated by the *Official GED Practice Test Required* coefficients under *Ethnicity: African-American Slope*). An example of this mitigating effect can be seen in the Language Arts, Writing content area: the African-American coefficient under candidate predictors is -31.6, suggesting that this group did not perform as well as white candidates on the Language Arts, Writing content area, controlling for other predictors. However, this effect was lessened when the test center required the OPT (B=8.8). That is, although an African-American candidate is predicted to score 31.6 points lower in the Language Arts, Writing content area, he or she is expected to score lower by only 22.8 points (i.e., -31.6 + 8.8) if he or she tested in a center with an OPT requirement. This mitigating effect of an OPT requirement was not found for the gap between Hispanic and white candidates for any of the content areas.

Once again, as an example of obtaining a predicted value for a particular student, imagine predicting the Language Arts, Writing content area standard score for an African-American female who took the GED Test in an average test center, which required an OPT, in an average jurisdiction. To obtain this predicted value, one would simply include the characteristics of the person for whom they want to obtain predicted values in the following equation:

$$\hat{Y}_{ijk} = \beta_{0\,jk} + (25.73)gender_{ijk} + (-31.62 + 8.84)\,African\,American_{\,ijk} + (3.09)took\,practice\,test_{ijk}$$
$$+ (-0.73)\,practice\,required_{\,jk}$$
$$\hat{Y}_{ijk} = 484.10 + (25.73)*1 + (-31.62 + 8.84)*1 + 3.09*1 - (0.73)*1$$
$$\hat{Y}_{ijk} = 489.41$$

Therefore, this candidate would be predicted to have a Language Arts, Writing standard score of approximately 489.

The coefficients excluded from this function are omitted because their observed values (the values that are multiplied by the coefficients) are zero. As mentioned previously, the value is zero because the candidate for whom we are predicting a score is average on all the other variables (group-mean and grand-mean centering).

For each of the content areas, the models (including all predictors) accounted for a total of approximately 7 percent, 13 percent, 17 percent, 10 percent, and 14 percent of the variation in the Language Arts, Writing, Social Studies, Science, Language Arts, Reading, and Mathematics content area scores, respectively.

In summary, the majority of significant predictors in the models predicting both the overall GED Test standard score and content area standard scores were candidate characteristics such as ethnicity, primary language, highest grade a candidate completed, and repeat test-taker status. Although there does appear to be variability in overall GED standard scores and content area scores associated with the test-center and jurisdiction levels, only one of the predictors at these levels, a test center open all months of a year, helped account for this variation in a single content area.

The most consistent finding associated with test centers was that the gap between African-American and white candidates was smaller in test centers that required the OPT. However, this requirement did nothing to decrease the gap between Hispanic and white candidates. Another interesting finding is that the amount of variation the models explained varies by content area, suggesting that the importance of different variables differs as a function of the content area. The lack of significant predictors at the jurisdiction level suggests that the exploration of different policies and/or characteristics of jurisdictions should be considered.

Cross-Validation

Prior to performing the inferential analyses for research questions 1 and 2, cases from the complete sample were randomly assigned into one of two samples. The parameter estimates for all the HLMs (reported in Tables 9 and 12) were used to compute predicted values for both samples. The amount of variation in the outcome of interest explained by the predicted value was then computed for the overall GED standard score and each content area standard score. These r-square values were very similar between the two different samples and are presented in **Table 13**. In the case where the r-squared values differed most, for the Language Arts, Reading content area, the original sample r-square value was 0.100 while the validation sample r-square was 0.097, a difference of 3 percent. (Table 13 presents the values with two decimal places.) The consistency of this coefficient between samples suggests that the models estimated in the first sample predicted the outcome equally well in the second sample.

Table 13. Variance Explained by Estimated Parameters in HLMs From Sample 1

	Language Arts, Writing	Social Studies	Science	Language Arts, Reading	Mathematics	Overall GED
Sample 1	0.07	0.13	0.17	0.10	0.14	0.15
Sample 2	0.06	0.13	0.17	0.10	0.14	0.14

Note: R-squared values for Sample 1 are higher because Sample 1 data were used for estimating the parameters in the HLMs.

Discussion

Policies of test centers explain a small proportion of variability in GED Test standard scores, after controlling for candidate characteristics. Our results provide evidence that a test center that is open all months of a year can benefit candidate performance in the Mathematics content area of the GED Test. Mathematics generally is considered one of the most challenging content areas for candidates, and for some, it may be the final hurdle to completing, even though candidates are not required to complete content areas in a particular order. Candidates who have increased access to testing in year-round test centers may be more apt to go to the center as soon as they feel confident enough in their Mathematics skills to attempt the GED Test, either in its entirety or perhaps as a last content area. If a test center is not open and a delay results, scores could

drop. We do not want to over-interpret a small effect in a single content area, but we believe the role of access to test centers deserves further study.

This study suggests that candidate characteristics play the biggest role in GED performance both for content area standard scores and the overall GED standard scores. Based on the unconditional models, approximately 88 percent to 94 percent of the variance (depending on the score of interest) is at the candidate level.

The candidate-level results from this study can help enhance preparation practices. The consistently weak relationship between hours spent preparing for the test and actual performance suggests the following: (1) candidates did not report their number of preparation hours accurately before testing; (2) candidates may have spent additional hours preparing after completing a demographic form and before testing; (3) candidates did not prepare for the GED Test in a way that was beneficial to their performance; or (4) scores of candidates who spent numerous hours preparing and still did not improve their GED Test performance balanced out scores of high-performing candidates who prepared very little or not at all. Additional evidence of the relationship between time spent on GED preparation and candidate performance is needed.

Understanding the role that different candidate characteristics play in testing outcomes can help tailor preparation to the needs of individuals. For example, instructional centers may want to offer women preparation materials that emphasize Social Studies, Science, and Mathematics, while offering men preparation materials emphasizing Language Arts, Writing and Language Arts, Reading. It also may be beneficial for candidates who have spent several years out of school to sharpen their Language Arts, Writing and Mathematics skills. For more information on the role of gender in GED preparation, please refer to a recent GED Testing Service study on preparation and performance (McLaughlin, Skaggs, & Patterson, 2009).

The disadvantage African Americans and Hispanics face, in terms of both economic outcomes and high school graduation rates, is concerning. The results of this study suggest that both of these groups are also at a disadvantage when taking the GED Test after controlling for other relevant factors. For example, two candidates, one African American and one white, who are similar in many respects are expected to have different standard scores. Although this disparity should be concerning to policy makers, this preliminary evidence suggests that these effects can be partly mitigated for at least one minority group—African Americans—when a test center requires an OPT.

Although the requirement helps close the achievement gap between African-American and white candidates, the effect of taking the OPT was either slightly positive or slightly negative for candidates, controlling for other factors without referencing ethnicity separately. The magnitude of this effect is surprising. We would expect that taking the OPT would boost standard scores considerably because a candidate would presumably study after obtaining feedback on his or her performance. The lack of a large positive effect associated with taking the OPT may be a flag indicating that the OPT is not being used optimally, and this issue should be explored further. Test centers in particular could collect data on the use of the OPT in their centers and its effects on candidates of diverse ethnic backgrounds.

Although this study suggests that there may not be quick solutions to boost test performance, there are perhaps better ways of tailoring preparation for the GED Test to the individual characteristics of candidates.

Because there are many ways to define repeat test-takers (e.g., asking did the candidate complete the test before retesting? Did the candidate miss minimum score requirements in certain content areas before retesting? Did the candidate retest after completing all content areas, perhaps for an even better score[s]?), interpretation of repeat candidate results is difficult. However, we see the topic of repeat testing, with clear definitions, as an important area for future research.

Our study, using select variables, found that test centers and jurisdictions play less of a role in GED Test performance. The remainder of this variance might be captured by test center and jurisdiction characteristics and policies that were not available in our dataset. Future research must address a key question: How are more policy-related variables associated with test performance? We need to collect and analyze additional quantifiable and distinct variables that reflect important policy issues, such as candidate fees, preparation requirements, testing prerequisites, and retesting requirements. This future work could explain further the role that jurisdictional and test center policies play in the outcomes of GED candidates.

References

American Council on Education. (2008). *2007 GED Testing Program Statistical Report.* Washington, DC: GED Testing Service.

American Council on Education. (2009a). *2008 GED Testing Program Statistical Report*. Washington, DC: GED Testing Service.

American Council on Education. (2009b). *Technical Manual: 2002 Series GED Tests*. Washington, DC: GED Testing Service.

Bates, D., & Maechler, M. (2009). Linear mixed-effects models using S4 classes. http://cran.r-project.org/web/packages/lme4/index.html

Boesel, D., Alsalam, N., & Smith, T. M. (1998). *Educational and labor market performance of GED recipients.* Washington, DC: U.S. Department of Education.

Browne, M. W. (2000). Cross-validation methods. *Journal of Mathematical Psychology*. *44*, 108-132.

Cameron, S. V., & Heckman, J. J. (1993). The nonequivalence of high school equivalents. *Journal of Labor Economics*. *11*(1), pt.1.

McLaughlin, J., Skaggs, G., & Patterson, M. (2009*). Preparation for and performance on the GED Test*. Washington, DC: GED Testing Service.

Murnane, R. J., Willet, J. B., & Tyler, J. H. (2000). Who benefits from obtaining a GED? Evidence from high school and beyond. *The Review of Economics and Statistics*. *82*(1), 23-37.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.

R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria. http://www.R-project.org

Ruben, D. B. (1976). Inference and missing data. *Biometrika*. *63*(3), 581-592.

Smith, T. M. (2003). Who values the GED? An examination of the paradox underlying the demand for the General Educational Development credential. *Teachers College Record*. *105*(3), 375-415.

Song, W., & Hsu, Y. (2008). *Economic and non-economic outcomes of GED credential recipients*. Washington, DC: GED Testing Service.

Tyler, J. H., Murnane, R. J., & Willett, J. B. (2000). *The devil is in the details.* The National Center for the Study of Adult Learning and Literacy, Report #16.

Tyler, J. H. (2001). *So you want a GED? Estimating the impact of the GED on the earnings of dropouts* (Working paper no. 01-34). Providence, RI: Brown University.

Tyler, J. H. (2005). The General Educational Development (GED) credential: History, current research, and directions for policy and practice. *National Center for the Study of Adult Learning and Literacy*, *5*(Chapter 3).

U.S. Department of Education, National Center for Education Statistics, Institute of Education Sciences. (2007). *Status and trends in the education of racial and ethnic minorities*

U.S. Department of Education, National Center for Education Statistics, Institute of Education Sciences. (2009). *The condition of education 2009*. Retrieved December 15, 2009, from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2009081.

U.S. Department of Labor, Bureau of Labor Statistics. (2007a). *Employment outlook: 2006–2016, Occupational employment projections to 2016.* Retrieved December 15, 2009, from www.bls.gov/opub/mlr/2007/11/art5full.pdf.

U.S. Department of Labor, Bureau of Labor Statistics. (2007b). *A Profile of the working poor.* Retrieved December 15, 2009, from www.bls.gov/cps/cpswp2007.pdf.

U.S. Department of Labor, Bureau of Labor Statistics. (2007c). *Charting the U.S. labor market in 2006*. Retrieved July 3, 2009, from www.bls.gov/cps/labor2006/chart2-1.pdf.

U.S. Department of Labor, Bureau of Labor Statistics. (2008). *Education pays*. Retrieved July 3, 2009, from www.bls.gov/emp/emptab7.htm.

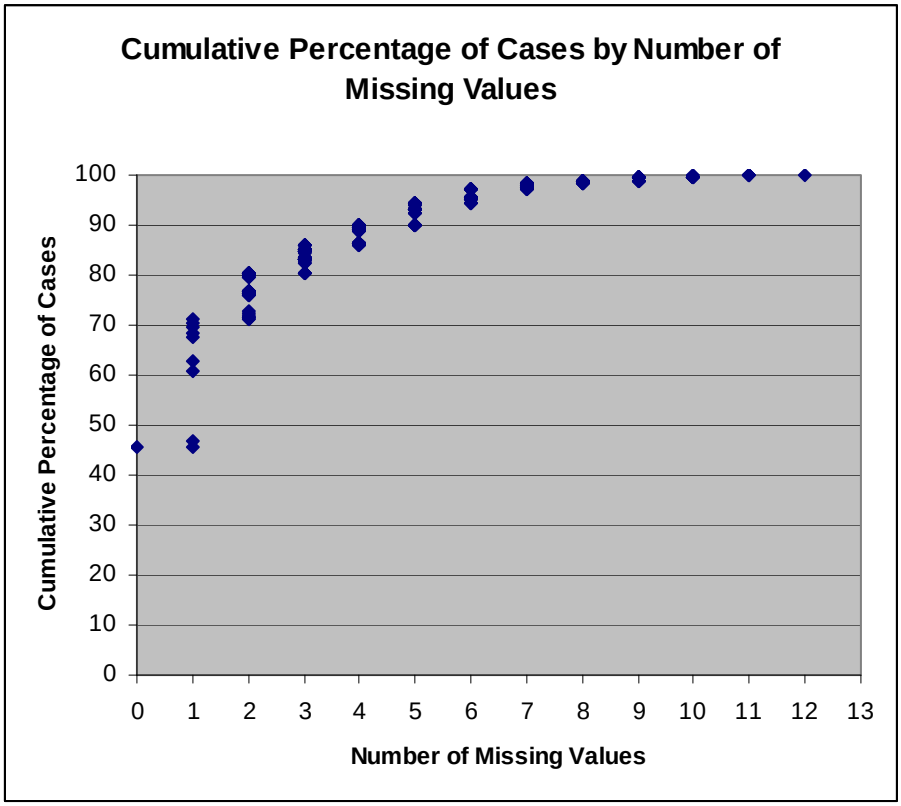
Appendix

Missing Data

The dataset used to answer research questions 1 and 2 (prior to splitting the dataset into two samples for cross-validation) consisted of 641,126 cases (n) with 24 covariates (p) resulting in a data matrix of 15,387,024 (n x p) cells. Of these cells; 640,045 consisted of missing values. Therefore, the dataset consisted of approximately 4.16 percent missing values. The dependent variables for the HLMs for research questions 1 and 2 were the mean overall test score and the scores from each content area (Language Arts, Reading; Language Arts, Writing; Social Studies; Science; and Mathematics). The dependent variables had no missing data. The lack of missing data on the dependent variables is a result of having conditioned the sample on candidates who completed the test in 2008.

For the majority of variables, less than 10 percent of cases contained missing values. However, for the variables of education level, primary language, years out of school, and hours spent preparing for the GED Test, the percentage of cases with missing values were approximately 13 percent, 15 percent, 24 percent, and 36 percent, respectively. Approximately 85.9 percent of cases had three or fewer missing values. The cumulative percentage of cases by the number of missing values is presented below in **Figure 2**.

Figure 2. Cumulative Percentage of Cases by Number of Missing Values



The entire dataset consisted of 361,686 cases or 56 percent of cases with at least one missing value. Cases with at least one missing value were identified, and each of the covariates and the dependent variables of interest were analyzed as a function of whether cases had missing values. The results suggest that although some of the covariates and dependent variables differed by whether cases had missing data, in all of the analyses less than 1 percent of the variation in the covariates and dependent variables was associated with having at least one missing value. This result suggests that cases excluded from subsequent analyses are similar to cases included in subsequent analyses. Subsequent analyses were based on the assumption that data are missing at random (MAR; Ruben, 1976) and the missing data were not imputed prior to analyses.

Cross-Validation

To cross-validate the HLM findings (for research questions 1 and 2), test completers were randomly assigned to one of two datasets (both of size $n = 320{,}563$). The first sample was analyzed, and the parameter estimates as well as the r-squared value for the model based on sample 1 coefficients were compared for both samples. Therefore, parameter estimates based on sample 1 data were used for the sample 2 data. If the r-squared values of the predicted value and the observed values are approximately the same between the two samples, this provides evidence of cross-validation (Browne, 2000).

Table A.1. Conditional Models of Content Area Standard Scores for Language Arts, Writing and Social Studies

	Language Arts, Writing		Social Studies	
Effect	B	SE B	B	SE B
Intercept	484.10***	2.53	527.00***	2.55
Candidate Predictors				
Gender: Female	25.73***	0.53	-16.01***	0.42
Ethnicity: Other	-0.10	22.76	-46.86*	18.19
Ethnicity: Hispanic Origin or Descent	-19.07***	1.07	-26.81***	0.86
Ethnicity: American Indian or Alaska Native	-23.05***	1.75	-23.69***	1.40
Ethnicity: Asian	-6.78**	2.29	-20.31***	1.83
Ethnicity: African American	-31.62***	0.89	-48.12***	0.69

Continued on next page

Table A.1. continued

Ethnicity: Native Hawaiian or Pacific Islander	-7.93*	3.23	-29.25***	2.58
Primary Language: French	-47.67***	6.84	-34.52***	5.46
Primary Language: Spanish	-18.20***	1.48	-29.63***	1.19
Primary Language: Other Language	-47.29***	2.61	-34.33***	2.09
Highest Grade Completed	6.58***	0.21	6.21***	0.17
Preparation Hours	-0.02***	0.00	-0.02***	0.00
Reported Taking Official GED Practice Test	3.09***	0.64	-6.93***	0.51
First-Time Test-Taker or Repeat	-33.02***	0.75	-47.10***	0.60
Years Out of School	-0.57***	0.03	0.85***	0.02
Test Center Predictors				
Number of Full-Time Employees	0.19	0.31	-0.02	0.30
Number of Part-Time Employees	0.63*	0.27	0.21	0.26
How Long the Center Has Been Testing	1.25	0.76	0.97	0.72
Open All Months	1.73	1.29	1.14	1.21
How Often The Center is Open	0.20	0.32	-0.05	0.30
Official GED Practice Test Required	-0.73	1.54	-1.11	1.43
Test Completion Required in One Day	1.09	2.65	-1.59	2.54
Waiting Period Required Before Retesting	0.14	1.34	-0.30	1.26
Jurisdiction Predictors				
Age of Required School Attendance without Exceptions	1.69	2.61	-0.38	2.64
Number of Test Centers	-0.05	0.05	-0.09	0.05
African-American Slope				
Number of Full-Time Employees	0.42	0.42	0.36	0.33

Continued on next page

Table A.1 continued

Number of Part-Time Employees	-0.40	0.35	-0.29	0.27
How Long the Center Has Been Testing	-0.42	1.28	0.01	1.01
Open All Months	4.65*	2.30	-0.03	1.80
Official GED Practice Test Required	8.84**	2.77	7.08**	2.18
Test Completion Required in One Day	3.46	4.14	1.19	3.24
Waiting Period Required Before Retesting	5.24*	2.33	2.83	1.83
Hispanic Slope				
Number of Full-Time Employees	0.63	0.45	0.21	0.37
Number of part-Time Employees	-0.40	0.40	-0.20	0.32
How Long the Center Has Been Testing	-1.21	1.53	1.07	1.23
Open All Months	5.13	2.50	2.74	2.01
Official GED Practice Test Required	0.16	3.30	0.13	2.65
Test Completion Required in One Day	-0.01	6.07	2.52	4.85
Waiting Period Required Before Retesting	0.37	2.66	0.70	2.14

Note: Gender: Male=0, Female=1; Ethnicity was dummy coded with white candidates as the reference group; Reported Taking Official GED Practice Test: No=0, Yes=1; First Time Test-Taker or Repeat: First Time=0, Repeat=1; Open All Months: No=0, Yes=1; *=<0.05, **=<0.01, ***=<0.001.

Table A.2. Conditional Models of Content Area Standard Scores for Science and Language Arts, Reading

	Science		Language Arts, Reading	
Effect	B	SE B	B	SE B
Intercept	532.80***	2.69	551.10***	2.91

Continued on next page

Table A.2 continued

Candidate Predictors				
Gender: Female	-24.24***	0.40	12.44***	0.53
Ethnicity: Other	-36.14*	17.01	-41.80	22.57
Ethnicity: Hispanic Origin or Descent	-34.12***	0.81	-26.26***	1.04
Ethnicity: American Indian or Alaska Native	-27.61***	1.31	-24.29***	1.74
Ethnicity: Asian	-26.17***	1.71	-28.50***	2.27
Ethnicity: African-American Descent	-56.21***	0.64	-43.77***	0.85
Ethnicity: Native Hawaiian or Pacific Islander	-33.64***	2.41	-32.01***	3.20
Primary Language: French	-37.19***	5.11	-51.13***	6.77
Primary Language: Spanish	-20.30***	1.11	-45.25***	1.47
Primary Language: Other Language	-33.70***	1.95	-58.36***	2.59
Highest Grade Completed	5.01***	0.16	5.30***	0.21
Preparation Hours	-0.02***	0.00	-0.02***	0.00
Reported Taking Official GED Practice Test	-5.32***	0.48	-5.01***	0.64
First-Time Test-Taker or Repeat	-47.75***	0.56	-57.17***	0.74
Years Out of School	0.07***	0.02	0.63***	0.03
Test Center Predictors				
Number of Full-Time Employees	-0.32	0.31	0.18	0.33
Number of Part-Time Employees	0.05	0.27	0.16	0.29
How Long the Center Has Been Testing	1.50*	0.72	0.00	0.81
Open All Months	2.09	1.23	2.12	1.37
How Often the Center is Open	-0.18	0.30	0.11	0.34
Official GED Practice Test Required	-2.12	1.44	0.78	1.62
Test Completion Required in One Day	0.70	2.58	-3.08	2.83
Waiting Period Required Before Retesting	-0.82	1.28	-0.66	1.42

Continued on next page

Table A.2 continued

Jurisdiction Predictors				
Age of Required School Attendance without Exceptions	-1.65	2.78	-0.91	3.01
Number of Test Centers	-0.11*	0.05	-0.10	0.05
African-American Slope				
Number of Full-Time Employees	-0.20	0.30	0.23	0.40
Number of Part-Time Employees	-0.30	0.25	-0.14	0.33
How Long the Center Has Been Testing	0.67	0.93	0.17	1.24
Open All Months	1.25	1.66	0.70	2.21
Official GED Practice Test Required	2.54	2.01	8.36**	2.68
Test Completion Required in One Day	1.17	2.98	-1.17	3.97
Waiting Period Required Before Retesting	2.79	1.68	2.48	2.24
Hispanic Slope				
Number of Full-Time Employees	-0.31	0.34	0.52	0.44
Number of Part-Time Employees	-0.25	0.31	0.02	0.39
How Long the Center Has Been Testing	0.40	1.15	-0.90	1.50
Open All Months	1.74	1.89	0.49	2.46
Official GED Practice Test Required	-1.02	2.49	-0.46	3.25
Test Completion Required in One Day	4.62	4.56	3.53	5.93
Waiting Period Required Before Retesting	1.09	2.01	0.73	2.61

Note: Gender: Male=0, Female=1; Ethnicity was dummy coded with white candidates as the reference group; Reported Taking Official GED Practice Test: No=0, Yes=1; First-Time Test-Taker or Repeat: First Time=0, Repeat=1; Open All Months: No=0, Yes=1; *=<0.05, **=<0.01, ***=<0.001.

Table A.3. Conditional Models of Content Area Standard Scores for Mathematics

	Mathematics	
Effect	B	SE B
Intercept	479.80***	2.43
Candidate Predictors		
Gender: Female	-16.55***	0.39
Ethnicity: Other	-40.12*	16.56
Ethnicity: Hispanic Origin or Descent	-23.43**	0.80
Ethnicity: American Indian or Alaska Native	-24.93***	1.28
Ethnicity: Asian	8.80***	1.67
Ethnicity: African American	-44.28***	0.65
Ethnicity: Native Hawaiian or Pacific Islander	-16.11***	2.35
Primary Language: French	-0.41	4.97
Primary Language: Spanish	-2.43*	1.08
Primary Language: Other Language	14.26***	1.90
Highest Grade Completed	6.09***	0.15
Preparation Hours	-0.02***	0.00
Reported Taking Official GED Practice Test	-1.67***	0.47
First-time Test-Taker or Repeat	-38.40***	0.54
Years Out of School	-1.08***	0.02
Test Center Predictors		
Number of Full-Time Employees	-0.65*	0.29
Number of Part-Time Employees	0.21	0.25
How Long the Center Has Been Testing	1.39*	0.68
Open All Months	3.13**	1.15

Continued on next page

Table A.3 continued

How Often the Center is Open	0.03	0.28
Official GED Practice Test Required	-0.98	1.36
Test Completion Required in One Day	2.21	2.42
Waiting Period Required Before Retesting	-0.52	1.20
Jurisdiction Predictors		
Age of Required School Attendance without Exceptions	0.27	2.52
Number of Test Centers	-0.10*	0.05
African-American Slope		
Number of Full-Time Employees	0.29*	0.31
Number of Part-Time Employees	0.00	0.26
How Long Been Testing	0.01	0.94
Open All Months	0.98	1.68
Official GED Practice Test Required	4.05*	2.03
Test Completion Required in One Day	3.28	3.05
Waiting Period Required Before Retesting	0.36	1.71
Hispanic Slope		
Number of Full-Time Employees	0.28	0.34
Number of Part-Time Employees	-0.09	0.30
How Long the Center Has Been Testing	0.50	1.14
Open All Months	2.42	1.87
Official GED Practice Test Required	-0.95	2.45
Test Completion Required in One Day	4.54	4.50
Waiting Period Required Before Retesting	0.21	1.99

Note: Gender: Male=0, Female=1; Ethnicity was dummy coded with white candidates as the reference group; Reported Taking Official GED Practice Test: No=0, Yes=1; First-Time Test-Taker or Repeat: First Time=0, Repeat=1; Open All Months: No=0, Yes=1; *=<0.05, **=<0.01, ***=<0.001.

A Program of the American Council on Education®

One Dupont Circle NW, Suite 250
Washington, DC 20036-1163
(202) 939-9490
Fax: (202) 659-8875
www.GEDtest.org