Preparation for and Performance on the GED Test

GED Testing Service® Research Studies, 2009-2

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

Preparation for and Performance on the GED® Test

Joseph W. McLaughlin
Gary Skaggs
Virginia Polytechnic Institute and State University

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

Table of Contents

Acknowledgments

This research was supported by a graduate research fellowship sponsored by the GED Testing Service, American Council on Education. An earlier version of this paper was presented by the first two authors at the 2007 annual meeting of the American Educational Research Association in Chicago, Illinois. Contact information for the first two authors is: Joseph W. McLaughlin at JWM@exchange.clemson.edu, and Gary Skaggs at gskaggs@vt.edu, or Educational Research and Evaluation Program, School of Education, 319 East Eggleston, Virginia Tech, Blacksburg, VA, 24061. Contact information for the third author is: Margaret Patterson, Director of Research for GED Testing Service, at margaret_patterson@ace.nche.edu. We thank Anne Guison-Dowdy, Wei Song, and Jizhi Zhang of GED Testing Service for their review of and comments on the final paper.

Abstract

GED testing candidates have many options available to them to prepare for the GED Test, including adult education classes, practice tests, and self-study. This study focused on candidates who voluntarily took the GED Test and could choose freely among preparation activities. We examined GED Test preparation activities and created eight mutually exclusive test preparation profile groups: public school adult education with or without a practice test, community college adult education with or without a practice test, individual study with or without a practice test, practice test only, and *none*.

The final sample included 90,032 U.S. candidates who completed the GED Test in 2004 and fell into one of the eight test preparation profile groups. Candidates in the study most often reported studying on their own without a practice test (29 percent) or studying in a public school adult education program without a practice test (28 percent). A smaller proportion studied in a community college adult education program without a practice test (13 percent) or did not prepare at all (14 percent). The public school and community college adult education profile groups had mean GED scores 25 to 34 points higher for candidates who took a practice test. The individual study group means were 20 points higher for candidates who took a practice test.

Among the profile groups, the group members who studied individually with a practice test and the *none* group scored the highest on average. The lowest mean scores were observed in the group who indicated public school adult education preparation without a practice test. All groups' content areas and battery mean scores were higher than the passing requirement (410 and 2,250, respectively). The group with the highest pass rates was the individual study with a practice test group, and the lowest pass rates were recorded for the public school adult education without a practice test group.

Introduction

The purpose of this study was to investigate how candidates prepare for the GED Test battery and how those test preparation activities are related to achievement in the five content areas of writing, social studies, science, reading, and mathematics, as well as on the total test battery. As the stakes for various types of tests have increased in recent years, there has been a concomitant growth in the test preparation industry, which has been most notable and most widely publicized for tests required for entry into educational programs, such as the SAT, GRE, and MCAT. Prior research has shown improvements in test results for test preparation classes and practice tests. However, unlike test preparation activities for entry into educational programs, little is known about how adults prepare for the GED Test or the relationship between GED Test preparation activities and performance on the GED Test. Since 2002, the number of candidates who take the GED Test each year has increased to about 750,000 (ACE, 2006; ACE, 2009). For many, a GED credential is a gateway to improved employment and educational opportunities.

Candidates have a number of available options to prepare for the GED Test, including adult education classes, practice tests, and various self-study programs. In most states, candidates may choose freely among these options, but in a few states, candidates must meet certain prerequisites to take the GED Test, such as a minimum score on an Official GED Practice Test or a specific course background. In this study, we focused on candidates who voluntarily took the GED Test and were able to choose freely among preparation activities, and we examined test preparation activities frequently undertaken by GED candidates.

Literature Review

Previous research on test preparation ranges from concerns about the effectiveness or outcomes associated with completion of the GED Test to concerns about preparation for taking the test itself. Though a substantial portion of the research focuses on student outcomes, this review focuses on research concerning preparation for the examination.

The research is somewhat confounded by the lack of a definitive description of test preparation. Perlman (2004) noted that there is no "single, universally agreed-on definition for test preparation" and that "it is generally used to refer to a broad range of activities which vary in specificity in addressing a particular test" (p. 387). Similarly, Scholes and Lain (1997) defined test preparation as "the utilization of an aid or tool by a test-taker to acquire information and techniques for the purpose of attaining the highest score possible on a test" (p. 1). The authors noted that tools and techniques span across a number of activities, from simple reviews to coaching to intensive instructions in test taking (Scholes & Lain, 1997; Miyasaka, 2000; Briggs, 2001). Scholes and Lain identified activities ranging from familiarization to "drill and practice with feedback, exercises in strategies for various item formats and general test taking, subject-matter review, and/or skill development" (p. 1).

A general term sometimes applied to test preparation techniques is *coaching*. Scholes and Lain (1997) noted that coaching strategies can be used to familiarize students with how tests are administered. Test preparations frequently focus on techniques for effective review and recollection of content areas (Rubenstein, 2004, p. 398). Coaching strategies geared toward familiarity can examine assessment approaches and item formats (Miyasaka, 2000), and can focus on items similar to those found on the type of test being administered (Perlman, 2004), the types of formats appearing on the answer sheet (Perlman, 2004), and on psychological

factors affecting motivation (Miyasaka, 2000). Researchers generally agree that an effective test-taking strategy for any test requires "(a) understanding how the test is scored, (b) having a timing strategy, and (c) having a question selection and priority setting strategy" (Rubenstein, 2004, p. 398).

Rubenstein (2004) recommended that test preparation should, at a minimum, address the following:

- Poor general test-taking strategies (pacing, question selection, and setting priorities).
- Lack of specific problem-solving skills, which relates to the fact that particular questions on a test instrument are not aligned with a student's learned curriculum; that is, assisting the student in understanding how to translate what they know to what the test is asking.
- Lack of practice with the preceding skills and lack of ability to deal with the psychological difficulties attached to a standardized test.
- Physical exhaustion.
- Lack of basic skills that were part of a student's curriculum (Rubenstein, 2004).

With respect to preparation content, Perlman (2004) suggested that content can come from:

- The domain being tested.
- Institution-provided examples.
- Commercial test preparation.
- Parallel forms or old tests.

- The actual test (generally frowned upon and considered inappropriate, unethical, and potentially illegal) (Perlman, 2004).

Issues in Test Preparation

*Short- Versus Long-term Preparation***.** Test preparation can be viewed as long term, which according to ACT research "has continually demonstrated the benefits of taking longer-term college preparatory coursework for increasing ACT scores, *regardless of students' prior achievement in high school.*" ACT states that as long as a student enters these courses ready to learn, any student can benefit from long-term test preparation. ACT also states that high school coursework increases an ACT composite score substantially more than any type of short-term test preparation activity (ACT, 2005).

There are, however, some benefits associated with short-term test preparation, the greatest of which are associated with commercial test preparation courses, test preparation workshops offered by local schools, and test preparation through computer software. Some research, though, has implied that minimal effect has been found through commercial test preparation (Briggs, 2001, cited in ACT, 2005).

Test Validity

One important issue noted by Miyasaka (2000) and Rubenstein (2004) is how test preparation can potentially affect test validity. Miyasaka (2000) stated that the majority of large-scale assessment tests should be designed so that one could make reasonable inferences about the achievement levels of students with respect to content knowledge and/or skills within a given domain. The primary concern is whether a few months of preparation can significantly influence scores. In other words, if preparation does affect test scores, is the purpose of the test

to measure the knowledge of the test-taker (pertaining to the topic of the given test)? If yes, how much knowledge is truly being measured, specifically if test-takers have access to tools that may artificially enhance their scores (Rubenstein, 2004, p. 398)?

ACT has noted that the earning of high scores on the ACT should not be merely a reflection of intrinsic talent or provisional preparation, but should reflect a level of accomplishment resulting from hard work, planning, and a solid commitment (ACT, 2005). But the validity of the test score is compromised when test preparation artificially increases students' test scores without increasing mastery of the content domain, underlying subject-area knowledge, and/or testing skill (Perlman, 2004). Further, any test preparation practice that violates ethical standards would in effect nullify any validity of the students' test results (Miyasaka, 2000). (For more information on test validity, see the Technical Appendix.)

Random Assignment and Self-selection. Another issue central to the study is student self-selection into two groups: those who are *coached* and those who are *uncoached*. Briggs (2004) noted that students in the NELS (National Education Longitudinal Survey) subsample have self-selected into coaching programs, and because they have not been randomly assigned, there is no credible justification to deem coached and uncoached groups of students as comparable. Because coached and uncoached students may diverge along essential characteristics correlated with admissions test performance, any comparison of average score gains that does not control for such divergences probably will be biased. The issues associated with self-reported surveys are well-documented in the literature. Rubenstein (2004) noted, "If the effectiveness of test preparation is to be studied seriously, such a study will have to be carried out on a specific, well-defined form or test preparation, with rigorous standards as to what this sort of preparation constitutes and what it does not" (p. 399).

Briggs (2001) also pointed out another important potential source of bias in estimated coaching effects: Students self-select not only their test preparation, but also whether they will even take the test. In the case of the GED Test, some candidates are required or highly encouraged, depending on the jurisdiction they live in and their status (early release, military, etc.), to take particular test preparations. In addition, Briggs stated that the NELS survey clearly was not designed to evaluate the efficacy of coaching on the SAT, for if it had been, one would have expected more detailed questions about the students' test-taking and test-preparing experiences.

Briggs pointed out that there was little difference in racial background and achievement in terms of distinguishing who is coached from who is not coached. Both coached and uncoached test-takers had similar average PSAT scores, similar enrollment rates in college-preparatory classes during high school, and similar performances on standardized tests in reading and math, administered as part of NELS. On average, neither group appeared more intrinsically motivated, both reported comparable levels of self-esteem (as measured by a cluster of NELS survey items), and both reported producing the same amount of homework per week. The one particular difference between coached and uncoached students was that the coached students were more socioeconomically advantaged and more extrinsically motivated to take the SAT (Briggs, 2004).

Literature Summary

In light of this research, it is apparent that any nonexperimental research on the GED demographic survey data will be primarily observational, as will be any significant correlations between groups. The purpose, therefore, is a search to find out what kinds of test preparation

methods different candidates are using, explore for any particular peculiarities, inspect for apparent trends, and then examine any traits that might expose potential future research questions.

The goal of this study was to provide a profile of what GED test preparation activities candidates used and the relationship of those activities with demographic and educational background characteristics and performance on the GED Test. Specifically, this study addressed the following research questions:

1. How do candidates who use different test preparation activities perform on the GED Test?
2. Is there a relationship between various test preparation activities and GED Test performance, controlling for demographic and educational background variables and reasons for taking the GED Test?

Method

Data

This study used GED Test candidate data from the 2004 test cycle. There are five content areas in the GED Test battery: Language Arts, Writing; Social Studies; Science; Language Arts, Reading; and Mathematics. Each of these five tests is designed to mirror the typical U.S. high school program of study. Standard scores for each test are reported on a 200–800 scale. The GED Testing Service has minimum requirements (410 on each test and a battery total of 2,250), but each jurisdiction may require a higher passing standard.

In addition to GED Test scores, the candidates' responses to the U.S. Demographics Survey were used in this study, which were related to demographic and educational

background, candidates' reasons for taking the GED Test, and the test preparation activities that they used. (For more information on the data file, see the Technical Appendix.)

The U.S. Demographics Survey was completed only once before testing began. Individual GED Test results were merged with the demographic survey information, thereby creating multiple records for each candidate. This study used test records that represented the first time a test was taken by a candidate. Using the five content area scores, the battery average score was calculated. Then, using the content area scores and the GED passing standards (a minimum standard score of 410 per content area and a total standard score of 2,250 across the test battery), fields were assigned for each content area indicating whether the individual met the minimum score requirements. The number of candidates who were recorded as having completed the battery was 350,000.

Because the objective was to study particular test preparation activities—primarily, but not exclusively, practice tests—of candidates who took the GED Test voluntarily (nonobligatory) and chose freely among test preparation activities, the following candidates were excluded from this study:

- Candidates who indicated the following obligatory reasons for taking the GED Test: keep a current job, employer requirement, military entrance, military career, early release, court order, or public assistance requirement.
- Candidates who resided in jurisdictions where passing the Official GED Practice Test was required in 2004: Arkansas, Connecticut, Delaware, District of Columbia, Indiana, Iowa, Kentucky, Maine, North Carolina, and Tennessee.

- Candidates who did not indicate any response to the survey question regarding preparation activities. If a candidate indicated *none* for test preparation, his or her test records were included.

After this round of filtering, the number of candidates who completed the battery in 2004 and met the voluntary testing and free choice of test preparation was 160,000.

Creation of Profile Groups

Next, predominant test preparation activities or combinations of test preparation activities were determined. The survey posed the following question to candidates: How did you prepare for the GED Test? A list of activities followed the question, including public school and community college adult education, home study, library, Internet/computer, and *none*. There were 28 options from which to choose, and candidates were directed to mark all that applied. As a result, there were a large number of combinations of test preparation activities among the 160,000 candidates. After examining these various combinations, eight mutually exclusive test preparation profile groups were created representing the largest configurations of test preparation activities.

The eight mutually exclusive test preparation profile groups were:

- Public school adult education without a practice test.
- Public school adult education with a practice test.
- Community college adult education without a practice test.
- Community college adult education with a practice test.
- Individual study without a practice test.
- Individual study with a practice test.

- Official practice test.
- None. (The candidates indicated they did not undertake any of the listed test preparations.)

The eight groups were mutually exclusive in order to remove interdependence within the data. For example, if an individual indicated preparation via both library and public school adult education, there would be no way to determine which preparation had the stronger effect, nor could it be verified that the effect of public school adult education preparation was a significant factor because the effect also could have been caused primarily, if not solely, by library preparation activities. This interdependence could become more confounding given that candidates who prepared via public school adult education were not the only ones who marked *library preparation*; individuals who marked *community college adult edu*cation and *home study* selected *library preparation*, as well. (For more detail on the profile groups, see the Technical Appendix.)

The final sample included 90,032 candidates who completed the battery in 2004, met the voluntary testing and free choice of test preparation criteria, and fell into one of the eight test preparation profile groups (55 percent of the previously filtered 160,000). The other 45 percent of candidates had selected very low volume test preparation combinations. No other combination of test preparation activities was checked by more than 3 percent of the candidates, so the remainder of the 28 options was excluded from analysis.

Table 1 shows the number and percentage of candidates in each profile group. The largest profile groups included candidates who were involved in public school adult education, community college adult education, and individual study without using practice tests. Three

additional groups were formed from the aforementioned three groups with the addition of practice tests. The final two profile groups were candidates who checked only the practice test as a preparation activity and those who checked *none*.

Table 1
Number and Percentage of Candidates in Test Preparation Profile Groups (N=90,032)

	Number	Percent of the original 160,000	Percent of 90,032
Public School Adult Education without Practice Test	25,384	16	28
Public School Adult Education with Practice Test	3,006	2	3
Community College Adult Education without Practice Test	11,930	7	13
Community College Adult Education with Practice Test	2,150	1	2
Individual Study without Practice Test	26,520	16	29
Individual Study with Practice Test	2,780	2	3
None (Exclusive)	12,229	7	14
Practice Test (Exclusive)	6,033	4	7
Total	90,032	55	100

Propensity Score Analysis

To estimate treatment effects in nonexperimental designs, Rosenbaum and Rubin (1983) proposed a method called *propensity score analysis* to compare two intact groups on a dependent variable. Using this method, logistic regression or discriminant analysis is used to calculate the propensity, or probability, that each participant will be a member of one of the groups (usually a treatment group) from a variety of covariates. Once this model is established, the propensity scores are used to either match or stratify the two groups so that comparisons on the dependent variable can be made between groups that are equal in propensity scores. In this way, the comparisons control for any pre-existing group differences that are measured by the covariates.

In this study, we used propensity score analysis to estimate the effects of taking practice tests and participating in adult education to prepare for the GED Test. We conducted five propensity score analyses to compare the following pairs of profile groups:

- Public school adult education with and without a practice test.
- Community college adult education with and without a practice test.
- Individual study with and without a practice test.
- Public school adult education with individual study.
- Community college adult education with individual study.

(For more detail on propensity score analysis methods and results, see the Technical Appendix.)

Results

Background Characteristics

Table 2 shows demographic variable descriptive statistics for each test preparation profile group: age, gender, ethnicity, income, years since high school, and highest grade completed in high school (in years). Standard deviations are shown in parentheses for age, years since high school, and highest grade completed. Income was summarized at three levels: no income, $1–$3,000, and more than $3,000.

Women were more likely than men to choose adult education. Further, those candidates who indicated undertaking no test preparation activities or preparing only with a practice test were younger and out of high school for fewer years than candidates who indicated adult education and individual study. The practice test only group had the highest percent in the two lowest income brackets. The highest percentage of candidates who indicated an annual income of more than $3,000 were part of the individual study without a practice test group. Candidates who took a practice test tended to be younger than those who did not take a practice test. Candidates who indicated no test preparation had the highest average grade completed. Within the test preparation activities of public school adult education, community college adult

education, and individual study, African-American and Hispanic candidates were more likely to have not taken a practice test, while white candidates were more likely to have taken a practice test.

Table 2
Descriptive Statistics for Test Preparation Profile Groups, by Demographic Variables

Demographic	PSAE without PT	PSAE with PT	CCAE without PT	CCAE with PT	IS without PT	IS with PT	None	PT Only	All Groups
	Mean *(standard deviation)*								
Age	23.65 *(8.63)*	22.73 *(8.18)*	24.61 *(9.31)*	22.79 *(8.20)*	23.29 *(7.51)*	23.00 *(7.94)*	21.38 *(6.29)*	21.43 *(7.13)*	23.14 *(8.04)*
Years Since HS	6.56 *(8.54)*	6.13 *(8.40)*	7.64 *(9.20)*	5.96 *(8.15)*	6.16 *(7.43)*	6.25 *(7.90)*	4.13 *(6.23)*	4.72 *(7.06)*	6.07 *(7.95)*
Highest Grade Completed	10.14 *(1.16)*	10.10 *(1.10)*	10.21 *(1.19)*	10.36 *(1.13)*	10.37 *(1.06)*	10.20 *(1.11)*	10.51 *(0.99)*	10.26 *(1.07)*	10.28 *(1.11)*
	Percentage								
Gender									
Female	53.19	55.50	56.57	56.80	48.89	55.87	40.80	40.10	50.06
Male	46.81	44.50	43.43	43.20	51.11	44.13	59.20	59.90	49.94
Ethnic Group									
Hispanic	22.23	15.48	18.56	9.95	18.31	13.98	15.28	15.94	18.40
American Indian	2.65	2.56	2.85	2.51	2.16	3.17	2.40	2.95	2.50
Asian	1.85	1.31	1.82	2.08	2.69	1.38	1.58	1.25	2.00
Black/African American	19.93	15.28	18.15	15.82	16.43	12.45	12.2	17.32	16.90
Native Hawaiian/ Pacific Islander	0.72	0.40	0.65	0.33	0.79	0.66	0.46	0.57	0.70
White	52.63	64.96	57.96	69.30	59.62	68.37	68.08	61.97	59.40
Annual Income									
$0	28.07	27.83	26.12	26.90	21.19	24.79	22.21	31.28	25.00
$1–$3,000	24.97	27.46	23.19	25.69	23.13	26.02	27.31	30.68	25.03
>$3,000	46.96	44.71	50.69	47.41	55.68	49.19	50.47	38.04	49.97

Note: HS=high school; PSAE without PT=Public School Adult Education without Practice Test; PSAE with PT=Public School Adult Education with Practice Test; CCAE without PT=Community College Adult Education without Practice Test; CCAE with PT=Community College Adult Education with Practice Test; IS without PT=Individual Study without Practice Test; IS with PT=Individual Study with Practice Test; None=No preparation; PT Only=Practice Test Only.

Self-reported Years of High School Study and Grades

The relationship between educational background and test preparation profile groups was examined. Candidates reported the number of years of high school study in five subject areas,

as shown in Table 3, and the grade that best describes the grades received in each area (not shown). The grade scale ranged from 5, "Mostly A," to 1, "Mostly below D."

Table 3
Educational Background, by Test Preparation Profile Group

	PSAE without PT	PSAE with PT	CCAE without PT	CCAE with PT	IS without PT	IS with PT	None	PT Only	All Groups
	Mean *(standard deviation)*								
Years of Study									
English Literature	2.28	2.30	2.24	2.27	2.45	2.36	2.61	2.33	2.38
	(0.98)	*(0.95)*	*(0.98)*	*(0.94)*	*(0.97)*	*(0.96)*	*(0.96)*	*(0.99)*	*(0.98)*
English Composition	1.93	1.97	1.94	1.99	2.07	2.01	2.13	1.95	2.01
	(0.96)	*(0.95)*	*(0.96)*	*(0.96)*	*(1.00)*	*(0.97)*	*(1.02)*	*(0.98)*	*(0.99)*
Social Studies	2.21	2.21	2.18	2.20	2.35	2.25	2.50	2.24	2.30
	(0.96)	*(0.92)*	*(0.95)*	*(0.93)*	*(0.96)*	*(0.95)*	*(0.95)*	*(0.97)*	*(0.96)*
Science	2.17	2.16	2.15	2.16	2.31	2.22	2.45	2.21	2.26
	(0.94)	*(0.90)*	*(0.93)*	*(0.90)*	*(0.93)*	*(0.91)*	*(0.91)*	*(0.95)*	*(0.94)*
Mathematics	2.27	2.26	2.26	2.26	2.43	2.31	2.57	2.30	2.36
	(0.98)	*(0.94)*	*(0.98)*	*(0.94)*	*(0.97)*	*(0.96)*	*(0.93)*	*(0.99)*	*(0.98)*

Note: PSAE without PT=Public School Adult Education without Practice Test; PSAE with PT=Public School Adult Education with Practice Test; CCAE without PT=Community College Adult Education without Practice Test; CCAE with PT=Community College Adult Education with Practice Test; IS without PT=Individual Study without Practice Test; IS with PT=Individual Study with Practice Test; None=No preparation; PT Only=Practice Test Only.

In terms of years of study in all subjects, the candidates who indicated no test preparation had the highest average years of study, while the community college adult education without practice test profile group had the lowest, followed closely by the other adult education profile groups. All groups had very similar grade averages.

Reasons for Taking the GED Test

The reasons reported for taking the GED Test are shown in Table 4. The survey question regarding reasons for taking the GED Test used a *mark all that apply* option; therefore, the percentages in Table 4 do not total 100 percent. The reasons for taking the GED Test were similar across the test preparation profile groups. The most frequently indicated reasons were personal satisfaction, enrollment in postsecondary education, and pursuing a better job. The least frequently cited reasons were job training, pursuing a first job, and skills certification.

Clearly, candidates in the eight profile groups differed on demographic variables, educational background, and reasons for taking the GED Test. As a result, differences between the profile groups for GED Test scores might not be attributable to differences in test preparation activities.

Table 4
Percentage Indicating Reasons for Taking the GED Test, by Test Preparation Profile Group

Reason	PSAE without PT	PSAE with PT	CCAE without PT	CCAE with PT	IS without PT	IS with PT	None	PT Only	All
Enroll in Tech Program	18.69	21.12	19.08	20.88	17.66	20.00	17.03	18.27	18.36
Enter Two-Year College	25.70	32.47	32.12	37.91	28.05	32.23	25.86	26.87	28.06
Enter Four-Year College	17.85	22.72	20.47	23.67	22.35	27.84	24.36	19.26	21.11
Skills Certification	4.53	6.69	5.00	6.47	5.12	8.67	4.55	5.62	5.09
Job Training	5.83	7.75	5.70	7.40	4.92	7.23	4.33	5.98	5.50
Pursue First Job	5.56	8.58	5.55	7.63	4.35	7.34	4.10	7.44	5.34
Pursue Better Job	34.12	44.15	34.75	41.58	34.95	43.88	30.76	29.31	34.48
Role Model for Family	13.30	22.85	13.57	18.88	12.97	24.89	9.79	12.71	13.53
Personal Satisfaction	38.50	57.82	41.33	55.72	44.22	64.71	41.56	46.81	43.40
Other	12.81	14.80	10.54	10.28	13.37	16.33	19.80	15.88	13.95

Note: Tech=technical or trade program; PSAE without PT=Public School Adult Education without Practice Test; PSAE with PT=Public School Adult Education with Practice Test; CCAE without PT=Community College Adult Education without Practice Test; CCAE with PT=Community College Adult Education with Practice Test; IS without PT=Individual Study without Practice Test; IS with PT=Individual Study with Practice Test; None=No preparation; PT Only=Practice Test Only.

GED Test Performance

Table 5 shows mean standard scores for each test preparation profile group for each content area and for the GED Test battery average. Among the five content areas, Mathematics and Language Arts, Writing, mean standard scores were the lowest for all groups. Language Arts, Reading, had the highest mean scores. Among the test preparation profile groups, the individual study with practice test and the *none* group scored the highest on average. The lowest mean scores were observed in the group that indicated public school adult education preparation without practice test. Table 5 indicates that the presence of a practice test along with adult education or individual study resulted in higher average standard scores. All groups'

individual content area means and battery total scores were higher than the passing requirement (410 and 2,250, respectively).

Table 5
GED Test Standard Score Means and Standard Deviations, by Test Preparation Profile Group

Content Area	PSAE without PT	PSAE with PT	CCAE without PT	CCAE with PT	IS without PT	IS with PT	None	PT Only	All Groups
Language Arts, Writing	471.18	496.03	480.96	510.21	493.34	514.40	500.70	506.30	488.46
	(76.53)	*(81.43)*	*(80.79)*	*(88.09)*	*(90.12)*	*(88.99)*	*(92.67)*	*(88.07)*	*(86.14)*
Social Studies	516.25	541.94	528.90	554.29	547.88	564.17	561.87	554.78	539.27
	(81.94)	*(83.75)*	*(86.14)*	*(87.75)*	*(93.37)*	*(86.24)*	*(100.11)*	*(90.91)*	*(91.11)*
Science	506.10	535.06	516.60	542.08	534.03	552.04	553.48	546.29	528.09
	(79.95)	*(79.85)*	*(80.95)*	*(82.07)*	*(87.23)*	*(78.74)*	*(92.83)*	*(84.27)*	*(86.14)*
Language Arts, Reading	539.65	573.99	554.55	583.07	571.23	597.25	583.14	583.46	563.73
	(101.62)	*(105.04)*	*(105.03)*	*(108.18)*	*(110.66)*	*(105.93)*	*(113.78)*	*(108)*	*(108.77)*
Mathematics	459.23	490.23	466.07	499.15	481.44	502.10	511.53	502.00	479.96
	(82.06)	*(85.35)*	*(84.91)*	*(90.67)*	*(96.31)*	*(90.45)*	*(107.09)*	*(94.79)*	*(93.7)*
Battery Average	498.48	527.45	509.42	537.76	525.59	545.99	542.14	538.57	519.90
	(69.27)	*(69.85)*	*(71.81)*	*(74.60)*	*(79.82)*	*(72.39)*	*(85.43)*	*(76.42)*	*(77.64)*

Note: PSAE without PT=Public School Adult Education without Practice Test; PSAE with PT=Public School Adult Education with Practice Test; CCAE without PT=Community College Adult Education without Practice Test; CCAE with PT=Community College Adult Education with Practice Test; IS without PT=Individual Study without Practice Test; IS with PT=Individual Study with Practice Test; None=No preparation; PT Only=Practice Test Only.

The public school and community college adult education profile groups had mean GED Test scores 25 to 34 points higher for those candidates who took a practice test. The individual study group means were approximately 20 points higher for candidates who took a practice test. There was a smaller, sometimes opposite difference between the *none* and practice test only groups; the average score for three of the content areas was lower for candidates who marked practice test only. The only instance without a consistent increase for practice tests was the comparison between those who marked *none* and those who marked *practice test only*, which supports the hypothesis that there are distinct populations of candidates who gravitate toward different test preparation methods.

Table 6 shows GED Test pass rates (i.e., the percentage of candidates who pass among candidates who complete the battery). The passing score for the GED Test battery is a total of 2,250, and a minimum of 410 is required for each content area; therefore, it is possible to meet

the minimum score in every content area and still not pass the battery. A similar pattern in the standard score analyses emerged in the pass rate analyses. The highest pass rates were found in the individual study with practice test group, and the lowest pass rates were found for the public school adult education without a practice test group. Taking a practice test appeared to be related to increased scores and pass rates. Unlike the standard score analyses, a higher percentage of candidates who marked *practice test only* passed than those who marked *none*.

Pass rates for the entire battery increased by 12 to 15 percentage points for adult education and individual study groups with a practice test. In individual content areas, when candidates reported taking a practice test, they were more likely to meet minimum score requirements in Mathematics and Language Arts, Writing Tests; these areas had the overall lowest percentages that met minimum score requirements. Rates in the other three content areas increased by less than six percentage points. Rates differed by three percentage points or less for the *none* and *practice test only* comparison.

Table 6
GED Test Pass Rates (Percentage), by Test Preparation Profile Group

Content Area	PSAE without PT	PSAE with PT	CCAE without PT	CCAE with PT	IS without PT	IS with PT	None	PT Only	All
Language Arts, Writing	82.90	90.69	85.74	91.77	87.22	93.42	88.53	91.43	86.68
Social Studies	93.09	96.47	94.60	97.12	95.58	98.31	95.19	96.78	94.93
Science	90.17	96.01	91.93	95.44	93.53	97.27	94.46	96.34	92.93
Language Arts, Reading	95.58	98.00	96.63	98.05	97.16	99.39	97.06	98.28	96.82
Mathematics	74.58	86.23	76.86	86.98	79.33	88.92	85.13	87.07	79.68
Battery	62.89	77.81	66.54	79.67	71.20	82.95	77.10	79.93	70.41

Note: PSAE without PT=Public School Adult Education without Practice Test; PSAE with PT=Public School Adult Education with Practice Test; CCAE without PT=Community College Adult Education without Practice Test; CCAE with PT=Community College Adult Education with Practice Test; IS without PT=Individual Study without Practice Test; IS with PT=Individual Study with Practice Test; None=No preparation; PT Only=Practice Test Only.

Discussion and Conclusions

Summary of Results

In this study, the test preparation activities of GED candidates who completed the GED Test battery during the 2004 test cycle were examined. Candidates who endorsed test preparation activities related to military installations or correctional facilities were excluded, as were candidates who tested to meet employer or public agency requirements, candidates who were required to take a practice test in order to be able to take the GED Test, and candidates who did not respond when asked about their preparation methods. The remaining candidates were able to choose freely among test preparation activities. The U.S. Demographics Survey asked candidates to indicate which of 28 test preparation activities, including *none*, they undertook. Therefore, the results of this study generalized only to U.S. GED candidates who were able to choose freely among test preparation options.

The most predominant test preparation activities were via public school and community college adult education, individual study, practice tests alone, and no preparation. Based on this predominance, we formed eight test preparation "profile" groups: public school adult education with or without a practice test, community college adult education with or without a practice test, individual study with or without a practice test, practice test only, and *none* (no test preparation).

Our initial intent was to estimate the effects of various test preparation activities. Candidates who reported that they had used none of the listed test preparation activities or who had studied on their own had higher GED Test scores on average and higher pass rates than candidates who reported that they prepared via adult education. However, it became evident that different populations of candidates engaged in different test preparation activities. For

example, women were more likely to enroll in adult education than men, and candidates who entered adult education tended to be slightly older. Adult education candidates were slightly more likely to report low or no income. Adult education candidates were also slightly more likely to have had fewer years of high school study in the five subject areas. As a result, propensity score stratification was used to determine whether GED Test score differences remained between profile groups while controlling for a number of covariates, including demographic and educational background, and reasons for taking the GED Test.

Given these considerations, major findings were:

- Among all the possible combinations of test preparation activities, eight distinct profiles accounted for 55 percent of all candidates: public school adult education with or without a practice test, community college adult education with or without a practice test, individual study with or without a practice test, practice test only, and *none*.
- There was considerable variation in GED Test scores and passing rates among the eight profile groups. The highest scoring groups were individual study with practice test, *none*, and practice test only. The lowest scoring groups were the public school and community college adult education groups without a practice test. However, all groups' mean test scores were higher than the passing requirements (410 minimum standard score on a single test and 2,250 total).
- For all groups of candidates, the most frequently cited reasons for taking the GED Test were personal satisfaction, pursuing a better job, and enrolling in college. The least frequently cited reasons were skills certification, job training, and pursuing a first job.
- Candidates who took a practice test were more likely than their non–practice test counterparts to select more reasons for taking the GED Test, and more likely to select

personal satisfaction, being a role model, and enrolling in a college program as reasons for taking the GED Test.

- Taking a practice test was associated with higher GED Test scores. Among the five content areas, the effect was greatest for Mathematics (see Technical Appendix, particularly Tables A1 through A3). The effect was also greater for candidates enrolled in adult education than for candidates who chose individual study.
- Even controlling for a large number of covariates, GED Test scores were still lower for the adult education groups compared with the individual study group. However, scores for candidates who participated in adult education and took a practice test tended to exceed scores for candidates who pursued individual study but did not take a practice test.

Implications for Practice

For GED candidates who choose to prepare for the GED Test, the variety of available preparation options is encouraging. Whether candidates choose to prepare via adult education, individual study, a practice test, or not at all may depend on a wealth of demographic and background characteristics. A few questions remain: What are the implications of young men's tendencies to not prepare, or to simply take a practice test before registering for the GED Test? What might be associated with women's preferences to study in an adult education program or on their own? Do candidates with little or no income prefer to take a practice test only, while those candidates with additional income opt to study on their own? Adult educators, GED examiners, publishers of preparation materials, and others who want to assist GED candidates should familiarize themselves with these characteristics and compare characteristics of the

candidates they serve at the local or regional level, with particular attention to gender, age, and socioeconomic status.

Another encouraging finding was that standard scores for candidates who made a choice about preparation tended to exceed minimum requirements at even higher levels than for the entire U.S. test-taker sample. Candidates who make this choice may do so in hopes of not simply passing, but doing well on the GED Test. Educators and others who offer test preparation materials and resources must be aware of this finding as they seek to recruit candidates for preparation or offer resources. Counselors, advisers, teachers, and parents can employ this information to encourage potential candidates who are unsure about the advantages of making a choice.

The decision to take a practice test also seems to be critical for candidates who prepare in either type of adult education program or on their own. It is worth noting that taking a practice test alone was not associated with higher average standard scores than taking it in conjunction with individual study or as part of adult education instruction. As an indicator of readiness for the GED Test, a practice test may offer candidates a potential edge as they finish their preparation. Candidates, educators, testing staff, and other stakeholders of GED Test preparation must be aware that this practice test effect could offer particular value for the content area of mathematics, which is typically one of the more difficult subjects for many candidates.

Conclusion

We acknowledge several limitations to this study and offer suggestions for future research. First, the results are limited to one year in the GED Test cycle, yet these results could vary over time. Additional studies of preparation methods across multiple years would yield valuable

information. Second, the background information from candidates is self-reported from a survey, and the accuracy of self-report data is not certain. The demographic survey form could be altered to more accurately measure the experience of candidates—for example, the form could include more specific answer selections about the types of practice tests, library experiences, or computer lab preparation. Third, there may be other test preparation activities that candidates use to prepare for the GED Test that were not listed in the survey. Further research on preparation methods involving computer-based or Internet-based preparation methods would be beneficial. Finally, this study included only GED candidates from 2004 who met the established criteria. The study was descriptive in nature and examined relationships of preparation and performance, not causes. The results of this study generalized only to U.S. GED candidates who were able to choose freely among test preparation options from the eight profile groups described.

Given these limitations, we hope that our results will be useful to professionals engaged in preparing candidates for the GED Test, and also inspire further research to guide how best to prepare future candidates.

References

ACT. (2005). *What kind of test preparation is best?* Iowa City, IA: ACT web flyer retrieved March 14, 2007, from www.act.org.

AERA, APA, & NCME. (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

American Council on Education. (2006). *Who passed the GED Tests? 2004 statistical report.* Washington, DC: GED Testing Service.

American Council on Education. (2009). *2008 GED testing program statistical report.* Washington, DC: GED Testing Service.

Barth, R. P., Gibbons, C., & Guo, S. (2006). Substance abuse treatment and the recurrence of maltreatment among caregivers with children living at home: A propensity score analysis. *Journal of Substance Abuse Treatment*, *30*, 93-104.

Briggs, D. C. (2001). The effect of admissions test preparation: Evidence from NELS:88. *Chance*, *14*(1), 10-18.

Briggs, D. C. (2004). Evaluating SAT coaching: Gains, effects and self-selection. *Rethinking the SAT: Perspectives Based on the November 2001 Conference at the University of California, Santa Barbara*, R. Zwick, ed., Routledge Falmer.

Hahs-Vaughn, D. L., & Onwuegbuzie, A. J. (2006). Estimating and using propensity score analysis with complex samples. *The Journal of Experimental Education*, *75*, 31-65.

Joint Committee on Testing Practices. (1998). *Rights and responsibilities of test takers: Guidelines and expectations.* Washington, DC: American Psychological Association.

Joint Committee on Testing Practices. (2002). *Code of fair testing practices in education.* Retrieved July 28, 2009, from www.theaaceonline.com/codefair.pdf.

Miyasaka, J. R. (2000, April). *A framework for evaluating the validity of test preparation practices*. Presentation in the symposium, Instructionally corrupt test preparation: Can it be detected or deterred? Paper presented at the Annual Meeting of the American Educational Research Association, April 2000, New Orleans, LA.

National Council on Measurement in Education. (1995). *Code of professional responsibilities in educational measurement*. Washington, DC. Available online: www.natd.org/Code_of_Professional_Responsibilities.html.

Perlman, C. L. (2004). Practice tests and study guides: Do they help? Are they ethical? What is ethical test preparation practice? Chapter 27 in Wall, J. E., & Walz, G. R. (eds.). *Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators*, 387-396. Greensboro, NC: CAPS Press.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 76*, 41-55.

Rubenstein, J. (2004). Test preparation: What makes it effective? Chapter 28 in Wall, J. E., & Walz, G. R. (eds.). *Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators*, 387-396. Greensboro, NC: CAPS Press.

Rudner, L. M., & Peyton, J. (2006). Consider propensity scores to compare treatments. *Practical Assessment, Research, & Evaluation*, *11*(9), 1-9.

Scholes, R. J., & Lain, M. M. (1997). *The effect of test preparation on ACT assessment scores*. Presented at the Annual Meeting of the American Education Research Association, March 1997, Chicago, IL.

Technical Appendix

Additional information and detail about the study is presented below, as well as further literature related to test validity, specifics on sampling, profile group development, and propensity score analyses. Additional tables are provided for more technical detail.

Test Validity

The validity of the test score is compromised when test preparation artificially increases students' test scores without increasing mastery of the content domain, underlying subject area knowledge, and/or testing skill (Perlman, 2004). Further, any test preparation practice that violates ethical standards would in effect nullify any validity of the students' test results (Miyasaka, 2000).

Fortunately, a number of publications are identified in the literature to assist in controlling the nature of test preparation to potentially compromise test scores. These publications are designed to clarify practices in testing that are ethical and appropriate. Some of these publications are:

- *The Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999).
- *The Code of Professional Responsibilities in Educational Measurement* (NCME, 1995).
- *The Code of Fair Testing Practices in Education* (JCTP, 2002).
- *The Rights and Responsibilities of Test Takers* (JCTP, 1998; Perlman, 2004).

In a study of NELS:88 data, Briggs (2001) determined that tutors significantly improve both math and verbal performances on the SAT. Commercial courses have large effects on both, too. Other than studying with a book, no other test preparation listed on the NELS had any significant effect. Coaching for the SAT and ACT was slightly significant in both verbal

(SAT), English (ACT), and math (both SAT and ACT) for increasing scores. However, Briggs noted that the average effect of coaching was clearly less than the levels previously suggested by commercial test preparation companies. In his conclusion, Briggs stated that the NELS dataset leaves no evidence that commercial test preparation significantly impacts admissions test performance. He made two important points to consider with his conclusions: (1) In regards to NELS:88, it is important to remember that the data originate from the early 1990s and may not reasonably reflect the modern state of the world; and (2) the benefits of coaching and tutoring may extend beyond potential admission test score improvements by teaching students better study habits and imbuing them with greater discipline and self confidence.

Data

This data file contained only candidate records for which there existed demographic survey information, item-level data, and raw score values that were verifiable based on item-level data. This file represented approximately 92 to 93 percent of the 2004 test cycle data; the states that were affected by the above selection and therefore may have been underrepresented in the analyses include Connecticut, Florida (Mathematics only), Indiana, Ohio, Vermont, California, New Jersey, Arizona, and Illinois.

The GED Testing Service provided the authors a data file with records of candidates whose records indicated that their data could be used for research; they tested within a U.S. credentialing jurisdiction/state; they completed a U.S. Demographics Survey in the English language; they had no testing accommodation; they took the English print version of the GED Test; they took the appropriate test forms for the year; and they had valid standard scores (ranging from 200 to 800).

Many candidates took the GED Test multiple times, and for this study, the first occurrence (date-wise) was used to nominate the *first time taken*. Note that only the 2004 data was used in this study, and it is likely that some candidates took some tests for the first time prior to 2004. The number of *first occurrence* candidates was 549,000. Using the candidate's ID, the five *first occurrence* tests were merged to create a battery profile. Individuals who did not complete the entire battery during the year 2004 were not surveyed for this study.

Profile Groups

The appropriateness of combining test preparation activities was considered, in that public school and community college adult education, as well as home study and self-taught, may have been considered similar test preparation activities. In the end, *home study* and *self-taught* were combined into single profile groups called *individual study*, but public school adult education and community college adult education were kept separate. The primary reasons for this separation included:

- There were few candidates who selected both public school and community college adult education; 28.6 percent of the 160,000 candidates selected public school adult education without selecting community college adult education, 13.5 percent chose community college adult education without selecting public school adult education, and less than 1 percent of the candidates selected both. This lack of overlap indicated that candidates viewed these test preparation activities as two distinct options. In contrast, 13.9 percent of the candidates selected self-taught without selecting home study, 19.1 percent chose home study without selecting self-taught, and 7.9 percent of the 160,000 selected both.

- Conceptually, there is not a clear distinction between *home study* and *self-taught*, yet because of the difference in governing bodies, public school adult education and community college adult education programs appear more distinct.

Finally, if a candidate selected *Official Practice Test* as a preparation activity, it was unknown whether these candidates interpreted the question to mean the Official GED Practice Test developed by GED Testing Service and published by Steck-Vaughn. Other publishers have GED Test preparation materials and practice tests, and there was no feasible way to determine if candidates who filled out the survey distinguished the difference. Therefore, practice testing in this study referred to any GED practice test.

Comparisons of Test Preparation Profile Groups

Although we reported the GED Test results of the profile groups, direct comparisons between profile groups would be suspect because of the lack of experimental design. That is, it is possible that distinct populations of candidates gravitated toward different test preparation methods. For example, when GED Test standard scores were regressed on the profile groups for Language Arts, Writing, those candidates who exclusively marked *none* for preparation had higher predicted scores than those who marked public school adult education, community college adult education, or individual study. This difference does not mean that attending adult education classes or independent study worked against candidate performance. As we show below, these groups differed in important ways, and all groups exhibited average scores above the passing score on the GED Test.

Propensity Score Analysis

To estimate treatment effects in nonexperimental designs, Rosenbaum and Rubin (1983) proposed a method called *propensity score analysis* to compare two intact groups on a dependent variable. This method has two advantages over the more common analysis of covariance (ANCOVA) approach. First, ANCOVA breaks down when there is insufficient overlap between the groups on the covariates. Second, including a large number of covariates reduces the degrees of freedom in hypothesis testing. Since its inception, propensity score analysis has been used in a variety of medical research settings but has received relatively less attention in social science and educational research (Hahs-Vaughn & Onwuegbuzie, 2006; Rudner & Peyton, 2006).

In this study, we used propensity score analysis to estimate the effects of taking practice tests and participating in adult education to prepare for the GED Test. These effects were independent of pre-existing differences between profile groups on a variety of background variables. We conducted five propensity score analyses to compare the following pairs of profile groups:

- Public school adult education with and without a practice test.
- Community college adult education with and without a practice test.
- Individual study with and without a practice test.
- Public school adult education with individual study.
- Community college adult education with individual study.

The first three analyses were intended to estimate the effect of taking a practice test. The last two analyses were intended to estimate the effect of participating in adult education classes. In

these two analyses, the individual study group was chosen as the comparison group. We also may have considered the profile group that checked *none*; however, it is not certain that this group engaged in no test preparation activities, only that they did not use any of the activities listed in the survey question.

In each analysis, logistic regression was used to calculate propensity scores. The covariates in this study included age, income, highest education level, sex, ethnicity, reasons for taking the GED Test, and grades and years of study in literature, composition, social studies, science, and mathematics. To facilitate profile group comparisons, the recommendation of Rosenbaum and Rubin (1983) was followed, in which five propensity score strata were created. The strata contained approximately equal numbers of candidates. Within each stratum, the two profile groups were nearly equal in their propensity scores. According to Rosenbaum and Rubin, the use of five strata accounts for approximately 90 percent of the pre-existing group differences on the covariates. As an alternative strategy, groups were matched on propensity scores, which has been done in several studies (e.g., Barth, Gibbons, & Guo, 2006; Rudner & Peyton, 2006). Matching provides a more precise equalization of groups on propensity scores but also requires a much larger sample size for one of the groups in order to match effectively. This strategy would be a possibility in this study for the three analyses for practice test effects, but would not be feasible for the two adult education comparisons. As a result, stratification was used for all five analyses.

Each propensity score model was validated by a series of analyses, either two-way ANOVA or logistic regression, in which a covariate was the dependent variable and stratum and profile group membership were the independent variables. Then, the significance of the

group and group by stratum interaction effects ware examined. If significant, then additional nonlinear and/or interaction terms were added to the model.

Results

Propensity Score Analysis Results

As mentioned above, five propensity score analyses were carried out. The first three analyses estimated the effect of taking a practice test separately for the public school adult education, community college adult education, and individual study groups. The fourth and fifth analyses underlined the differences between the two types of adult education programs and individual study. To calculate propensity scores, logistic regression was used to predict the probability of a candidate's placement in one of the two profile groups from the following covariates: age, income, highest level of education, sex, ethnicity, grades, and years of study in composition, literature, social studies, science, and mathematics, and reasons for taking the GED Test. We noted that there were high omission rates on the questions regarding grades and years of study in the five content areas, sometimes approaching 50 percent. We compared the distributions of complete cases to those of the cases that had missing values on all other covariates. There were only very small differences between these groups on the covariates. On the other hand, as noted previously, there were differences between the profile groups on educational background. As a result, we included these covariates in the model. However, to make the comparisons between profile groups valid, we used only complete cases throughout so that the effects of the propensity score stratification are based on the same cases.

After creating five equal-sized strata based on propensity scores in each analysis, the equality of covariates within strata were tested using either two-way ANOVA or logistic regression (depending on whether the covariate was quantitative or categorical). The main effect for profile group and interaction effect of profile group by stratum was examined. In the two adult education/individual study analyses, a quadratic term for age was added to improve fit. Out of the 64 hypotheses for the final logistic regression model of each propensity score analysis, there were no more than four significant results for any analysis. None of these significant hypothesis tests resulted in a change to multiple R-square of more than 0.002. As a result, we felt that the propensity score model successfully resulted in strata in which the two profile groups were equal on the collection of covariates.

The results for each of the five propensity score analyses are shown below. For each analysis, the mean difference between the two profile groups in GED standard scores was calculated for each stratum. The total effect is the average difference across strata. This difference is provided in both GED standard score units and effect sizes (in standard deviation units).

Public school adult education with and without a practice test. Table A1 shows the results of estimating the effect of taking a practice test for public school adult education. In this analysis, the two profile groups for public school adult education, one without taking a practice and one with a practice test, are compared on their GED standard scores for each of the five content areas and the battery average. First, the mean and standard deviations of GED standard scores is shown for the two groups. Note that these values differ from those presented earlier in Table 5, which showed the results for all cases. Table A1 shows the results only for complete cases (that is, no missing values on any covariates). As mentioned above, this procedure was

done so that the comparison of effects before and after propensity score stratification would be based on the same cases.

Before propensity score stratification, public school adult education candidates who took a practice test scored on average between 25 and 32 points higher than candidates who did not take a practice test. In terms of effect size, this amounted to a difference between 0.30 and 0.37 standard deviations. After controlling for the covariates through propensity score stratification, the differences were reduced to 17 to 22 points and 0.20 to 0.27 standard deviations. This evidence still suggests a benefit for taking a practice test. The benefit was largest for Mathematics.

Table A1
Estimated Practice Test Effect for Public School Adult Education

	Without Practice Test		With Practice Test		Before Propensity Stratification		After Propensity Stratification	
Content Area	Mean	SD	Mean	SD	Difference	Effect Size	Difference	Effect Size
Language Arts, Writing	479.40	77.79	504.73	82.71	25.34	0.32	17.23	0.22
Social Studies	525.77	83.00	550.68	85.50	24.91	0.30	17.13	0.20
Science	516.28	79.86	542.68	80.45	26.39	0.33	17.32	0.22
Language Arts, Reading	550.87	103.69	582.43	105.52	31.56	0.30	20.72	0.20
Mathematics	468.69	82.54	499.69	86.42	31.00	0.37	22.27	0.27

Note: SD=Standard Deviation.

Community college adult education with and without a practice test. Table A2 shows the results of propensity score stratification that compared the two profile groups for community college adult education, with and without a practice test. These results are similar to those reported above for public school adult education. Before propensity score stratification, candidates who took a practice test averaged 23 to 31 points, or 0.26 to 0.35 standard deviations, higher than candidates who did not take a practice test. After propensity score stratification, the advantage of taking a practice test was reduced to 15 to 21 points, or 0.18 to 0.23 standard deviations. As before, controlling for the covariates reduced the overall

differences for taking a practice test, but there was still a benefit. Again, the largest benefits were for Mathematics.

Table A2
Estimated Practice Test Effect for Community College Adult Education

	Without Practice Test		With Practice Test		Before Propensity Stratification		After Propensity Stratification	
Content Area	Mean	SD	Mean	SD	Difference	Effect Size	Difference	Effect Size
Language Arts, Writing	487.06	80.96	512.90	88.57	25.83	0.31	17.73	0.21
Social Studies	534.56	85.77	557.12	87.72	22.56	0.26	15.34	0.18
Science	522.38	81.01	545.66	83.54	23.28	0.28	15.51	0.19
Language Arts, Reading	560.49	105.44	588.04	108.86	27.55	0.26	19.00	0.18
Mathematics	472.03	86.65	503.34	91.10	31.31	0.35	20.51	0.23

Note: SD=Standard Deviation

Individual study with and without a practice test. The results comparing the two individual study profile groups, with and without a practice test, are shown below in Table A3. As in the above two analyses, candidates who took a practice test had higher GED standard scores than candidates who did not take a practice test. However, the differences were smaller than for the adult education groups. Before propensity score stratification, candidates who took a practice test averaged 14 to 21 points higher, or 0.15 to 0.22 standard deviations. After propensity score stratification, the differences averaged between 10 and 15 points, or 0.11 to 0.16 standard deviations. Compared with the two previous analyses, the practice effect benefit was lower for candidates who prepared on their own than for candidates enrolled in adult education. Again, the largest benefits were for Mathematics.

Table A3
Estimated Practice Test Effect for Individual Study

	Without Practice Test		With Practice Test		Before Propensity Stratification		After Propensity Stratification	
Content Area	Mean	SD	Mean	SD	Difference	Effect Size	Difference	Effect Size
Language Arts, Writing	500.49	89.75	519.83	88.98	19.33	0.22	12.71	0.14
Social Studies	556.23	93.19	569.73	86.07	13.50	0.15	9.78	0.11
Science	541.89	86.31	557.75	78.03	15.86	0.19	12.18	0.14
Language Arts, Reading	580.45	110.09	601.20	104.96	20.75	0.19	13.36	0.12
Mathematics	489.34	96.33	506.83	90.67	17.49	0.18	15.48	0.16

Note: SD=Standard Deviation

Public school adult education compared with individual study. As reported in Table 6, candidates who prepared on their own had higher mean GED Test scores than candidates who enrolled in public school adult education (in which candidates from neither group took a practice test). This finding is somewhat counterintuitive, and, as noted above, these two profile groups also differed considerably on many covariates. Our intent here was to determine if the GED Test score differences persisted when controlling for the covariates. The results are shown in Table A4. Before propensity score stratification, individual study candidates averaged between 21 and 30 points higher than the public school adult education candidates, or 0.23 to 0.34 standard deviations. After propensity score stratification, the differences were 10 to 20 points, or 0.11 to 0.22 standard deviations. In other words, approximately one-third to one-half of the original differences could be attributed to differences on the covariates, but individual study candidates still averaged higher GED Test scores after controlling for the covariates.

Table A4
Comparison of Public School Adult Education and Individual Study

	Individual Study		Public School Adult Education		Before Propensity Stratification		After Propensity Stratification	
Content Area	Mean	SD	Mean	SD	Difference	Effect Size	Difference	Effect Size
Language Arts, Writing	495.15	89.11	474.04	76.71	-21.11	-0.25	-13.63	-0.16
Social Studies	551.07	92.38	520.69	82.12	-30.38	-0.34	-19.99	-0.22
Science	537.52	86.32	510.64	79.84	-26.87	-0.32	-15.58	-0.18
Language Arts, Reading	575.37	109.75	545.18	102.39	-30.19	-0.28	-19.93	-0.18
Mathematics	483.72	95.58	463.14	82.08	-20.58	-0.23	-10.42	-0.11

SD=Standard Deviation

Community college adult education compared with individual study. The public school adult education analysis was repeated with community college adult education candidates and individual study candidates. The results are shown in Table A5. These results were similar to results for public school adult education, but the differences were slightly smaller. Before propensity score stratification, individual study candidates scored on average between 13 and 22 points higher than community college adult education candidates, or 0.18 to 0.24 standard deviations. After propensity score stratification, that difference was reduced to 10 to 16 points, or 0.11 to 0.17 standard deviations. Stratification reduced the difference but again did not reverse it.

Table A5
Comparison of Community College Adult Education and Individual Study

	Individual Study		Community College Adult Education		Before Propensity Stratification		After Propensity Stratification	
Content Area	Mean	SD	Mean	SD	Difference	Effect Size	Difference	Effect Size
Language Arts, Writing	500.49	89.76	487.08	80.96	-13.41	-0.15	-10.10	-0.12
Social Studies	556.23	93.20	534.56	85.76	-21.68	-0.24	-15.49	-0.17
Science	541.90	86.32	522.37	81.00	-19.53	-0.23	-12.66	-0.15
Language Arts, Reading	580.44	110.09	560.48	105.44	-19.96	-0.18	-16.37	-0.15
Mathematics	489.33	96.34	472.02	86.64	-17.31	-0.18	-9.89	-0.11

Note: SD=Standard Deviation

Discussion

Propensity Score Analysis

Propensity score stratification reduced the original observed differences in test scores in all five analyses but did not reverse the direction of the difference. This finding is probably not surprising for the three analyses that examined practice test effects. However, for the two analyses that examined the effects of adult education, these findings may seem disappointing. Unless one believes that adult education leads to poorer test performance than individual study (and we do not believe this to be the case), there are two possible explanations for these results.

First, these propensity score analyses used only the covariates available in the survey. It is possible that these variables alone do not eliminate all important differences between the profile groups. For instance, variables related to geographic location, type of community, and the resources of the local communities were not a part of these data files. Candidates may have selected adult education in general and public school adult education in particular because of the presence or absence of other resources or programs in the community.

There is some circumstantial evidence for this in these analyses. In the three analyses comparing profile groups with and without a practice test, the two groups were much more similar in the covariates than in the two analyses comparing adult education to individual study. Most of the significant results in validating the covariates occurred in the latter two analyses. In other words, there were fewer differences to overcome in the first three analyses.

Second, it is important to realize that the adult education profile groups consisted of candidates who checked only that test preparation activity and none of the others (except for practice test). Therefore, it is possible that adult education candidates who combined adult education with other test preparation activities (e.g., library or online programs) would have

compared more favorably to the individual study group. The numbers of candidates who engaged in these different configurations of test preparation activities were too few to permit more detailed analyses. Clearly, this complex issue is deserving of further research.

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