

**THE DEVIL IS IN THE DETAILS:
Evidence from the GED on the Role of Examination
System Details in Determining Who Passes**

by

**John H. Tyler
Brown University**

**Richard J. Murnane
Harvard University Graduate School of Education and NBER***

**John B. Willett
Harvard University Graduate School of Education**

**NCSALL Reports #16
April 2000**

The National Center for the Study of Adult Learning and Literacy
Harvard Graduate School of Education
101 Nichols House, Appian Way
Cambridge, MA 02138

* National Bureau of Economic Research

We gratefully acknowledge the research support provided by the National Center for the Study of Adult Learning and Literacy (NCSALL). Work supported by NCSALL was supported under the Educational Research and Development Centers Program, Award Number R309B60002, as administered by the Office of Educational Research and Improvement/National Institute on Postsecondary Education, Libraries, and Lifelong Learning, U.S. Department of Education. The contents do not necessarily represent the positions or policies of National Institute on Postsecondary Education, Libraries, and Lifelong Learning, U.S. Department of Education, and you should not assume endorsement by the Federal Government.

Table of Contents

Executive Summary 1

I. Introduction..... 4

II. The GED Credential and Examination System 4

III. Data 10

IV. The Role of Retesting in the GED System 13

A. The Potentially Complicated Route to a GED..... 13

B. Passing, Failing, and Retesting in the GED Program 15

V. Which Tests Prove to be the Highest Hurdles for GED Candidates?.....21

VI. Summary and Implications..... 24

References.....26

Executive Summary

As part of standards-based educational reform efforts, more than 40 states will soon require students to achieve passing scores on standardized exams in order to obtain a high school diploma. Currently, many states struggle with the design of their examination systems, debating such questions as the number of subjects to be tested and the rules regarding opportunities for students to re-take the examinations. There is, however, little systematic information on the impact alternative options would have on the number of students who will meet the standards for a high school diploma. Nor is much known about whether particular design options will have greater impacts on some students, such as students of color, than on other students.

In this paper we use data from a long-standing examination system, the General Educational Development (GED) certificate, to illustrate that the details of examination systems have marked impacts not only on the number of test takers who obtain the desired credential, but also on the racial/ethnic composition of passers. While the examination systems currently in place or being studied in K-12 systems across the nation are relatively new, the GED testing system has been in place since the 1940s. The stability and acceptance of this national examination system make it a unique and important vehicle for studying how the details of an examination system matter in who and how many ultimately pass.

In this study we use detailed data provided by the Florida Department of Education that contain basic demographic and test score information on 189,124 GED candidates who took the GED exams between 1988 and 1998. In these data we have information on the test scores of the first and the most recent tests in all five GED subject areas, allowing us a unique view of the testing history of these GED candidates.

Using these data, several key findings emerge. First we find that initial pass rates on the GED exams vary greatly by race/ethnicity and by age within race/ethnicity. About 77 percent of white GED candidates (of all ages) pass on the first attempt, while only 66 and 46 percent of Hispanic and black (respectively) candidates pass on their first attempt. Among white and Hispanic candidates, the lowest initial pass rates occur among candidates who are age 19-21 years when they take the tests, and the highest pass rates are for candidates who are thirty and older. However, among black GED candidates, the highest pass rates are among the youngest candidates, those who were 16–18 when they took the tests, and the lowest initial pass rates were among those who were age thirty and over. We found no substantial gender differences in initial pass rates.

In the GED system, candidates who fail the GED exams may retake the tests. (Different states have different guidelines regarding the timing of test retakes.) We find that about 65 percent of whites who fail on their initial attempt, retake the exams within three years, while about 60 percent of blacks and Hispanics retake the test. For all groups, younger candidates tend to retake the exams at higher rates than older candidates. Particularly notable is that while black GED candidates had initial pass rates that were 30 percentage points lower than white initial pass rates, there is only about a ten percentage point difference in the pass rates among retakers (71 percent for whites versus 60 percent for blacks). Also, Hispanics who retake the exams have pass rates that equal those of the white retakers. The ultimate pass rates in our data for all GED candidates is 88 percent for whites, 66 percent for blacks, and 80 percent for Hispanics. Again, we find no substantial gender differences in ultimate pass rates.

Our examinations of the time between test attempts show that the modal time between the first and the last attempt to be about two months. (About 20 percent of those who retake do so within two months.) It has been suggested to us by GED officials in Florida that it takes about two months for GED exams to be scored, recorded, and the results to be sent out. This suggests that many GED candidates in Florida who fail on their first attempt may be retaking the test with little additional preparation. Our results are only suggestive, however, and more work with more appropriate data is needed to determine how initial failure on the exams may affect further human capital formation among GED candidates.

We also find that among GED candidates who failed to obtain the credential, the writing test was the lowest score for the highest percentage of males, and the math test was the lowest score for the highest percentage of failing females. Slightly higher percentages of Hispanics had lowest scores on the reading portion of the test, relative to white and black candidates.

It is clear from this work that the ability to retake the GED exams is an important feature for a non-trivial portion of GED candidates. It is also clear that the ability to retake is relatively more important for black and Hispanic candidates than it is for white candidates. Without the ability to retake, and assuming no changes in preparation or the pool of testers, the black–white pass rate gap would be over 30 percentage points instead of 22 points. This suggests that providing individuals the opportunity to retry upon failing is an important provision if one goal of an examination system is diversity in the pool of individuals who “pass.”

The relatively high percentage of candidates who fail the exams and then retake shortly thereafter is a bit troubling. It is at least suggestive that for some

portion of GED candidates, little human capital is being built between attempts. We do note, however, that our findings show that about 80 percent of those who fail on first attempt take longer than two months before they retake the tests, and there may be substantial human capital development for these individuals. The policy recommendations from these particular results are not clear. During the period covered by our study, there was no time constraint on time-to-retest for the second attempt in Florida. Requiring a six-month waiting period between first and second attempt, as some states do, might induce more human capital formation among individuals whose skills prevented them from passing on first attempt. However, such a policy might also be a hardship for individuals who need the GED for timely admission to training, jobs, or post-secondary education.

The age differences on initial pass rates suggest that GED preparation programs are most important for young individuals who have been out of school for two to four years and for black candidates who are age thirty or over. What is not clear from our research are the reasons why individuals who were 19–21 when they initially took the tests consistently passed at lower rates than individuals who were 16–18 when they took the tests. It could be that the older candidate group is composed of individuals who leave school with lower levels of skills than individuals who tend to test sooner after leaving school. Or, it could be that the two groups left school with the same types of skills, but that the skills of the 19–21 year-old group have eroded relative to the skills of the younger group. Understanding these mechanisms is important, because, given the answer, different types of GED preparation programs may be more or less effective in helping students of different ages pass the exams.

Finally, our finding that the math test posed the most difficult barrier for female candidates, while the writing test was the primary barrier to a GED for males, replicates gender differences that have been found in other settings. This finding should reinforce to GED teachers, however, the importance of targeted instruction as they attempt to help individuals obtain the GED credential. It also points out that careful consideration should be given as to how different subject areas might be weighted in any given examination system.

Overall, our research points to the importance that should be given to the details that are associated with high stakes testing systems currently being planned or implemented by many state K-12 systems. Specifically, our examination of the GED examination system in Florida indicates that decisions regarding retesting opportunities and guidelines, as well as the weights that will be given to different subject areas, will have implications for both aggregate pass rates and the composition of those who successfully meet the standard.

**The Devil Is in the Details:
Evidence from the GED on the Role of Examination System Details in
Determining Who Passes**

I. Introduction

As part of standards-based educational reform efforts, more than 40 states will soon require students to achieve passing scores on standardized exams in order to obtain a high school diploma. Currently, many states struggle with the design of their examination systems, debating such questions as the number of subjects to be tested and the rules regarding opportunities for students to re-take the examinations. To date there is little systematic information on the impact of alternative options for the number of students who will meet the standards for a high school diploma. Nor is much known about whether particular design options will have greater impacts on some students, such as students of color, than on other students.

In this paper we use data from a long-standing examination system to illustrate that the details of examination systems have marked impacts not only on the number of test takers who obtain the desired credential, but also on the racial/ethnic composition of passers. In particular, we show that, among school dropouts in Florida who attempted the General Educational Development (GED) examinations, the mathematics exam posed the greatest obstacle to females and the writing exam to males. We also show that the option of retaking the examination was especially important in increasing the proportion of black and Hispanic test-takers who obtained the GED credential.

II. The GED Credential and Examination System

During the late 1930s and early 1940s millions of young Americans left high school to serve in the military. When the veterans returned to this country, they needed a high school completion credential to be eligible for the post-secondary education benefits in the G.I. Bill. Few wanted to return to their high school classrooms. So, with governmental support, the American Council on Education established a system of examinations designed to assess whether returning veterans possessed the skills and knowledge expected of high school graduates. The examinations became known as the General Educational Development exams and those veterans who achieved passing scores were awarded the GED credential.

Beginning in 1952 eligibility to take the GED examinations was broadened, making the alternative credential available to any American lacking a high school diploma who could pass the exams. The 1960s mark the beginning of a period of extraordinary growth in the number of people taking the GED examinations (Cameron and Heckman 1993b). If past patterns hold, about 700,000 individuals who lack a high school diploma will attempt the GED exams this year, with about 500,000 achieving passing scores. Out of the expected half-million individuals who will drop out of school this year (U.S. Department of Education 1997), about one in three will eventually earn a GED (Murnane, Willett, and Tyler 1999). The growth in the use of the GED as an alternative high school credential that began in the 1960s has been accompanied by the influx of millions of dollars of state and federal monies into adult education enterprises that prepare individuals to take the GED exams.

The increased profile of the GED has not escaped the attention of policy analysts and researchers. A 1992 report from the Department of Adult Education at the University of Georgia documents 55 studies examining various outcomes of GED graduates (Johnson and Valentine 1992). In the last decade there have been at least 15 additional studies looking exclusively at the economic benefits associated with GED attainment.¹ All 70 of these studies referenced above, however, are concerned primarily with the outcomes associated with GED attainment. No studies that we have found closely examine how dropouts use the GED system to obtain this alternative credential. This paper attempts to fill this knowledge gap.

Like other credentialing tests such as state bar examinations, dropouts who fail the GED exams may take them again, and many do. How many individuals who fail the GED make additional attempts? Of those who re-examine, what percentage pass? How long do failing candidates wait between attempts? Which subject area tests pose the greatest hurdle to passing? A better understanding of the answers to these questions will inform discussions about the structure of the GED testing program. The answers will also shed light on the design of the examination systems that the majority of states are introducing as part of standards-based educational reforms.

To obtain a GED, individuals must take a seven hour and thirty-five minute battery of five tests. The tests cover mathematics, writing, science, social sciences, and interpretation of literature and the arts.² For expository simplicity we will call

¹ See (Bos 1995; Cameron 1992; Cameron and Heckman 1993a; Cameron and Heckman 1993b; Cao, Stromsdorfer, and Weeks 1996; Cave and Bos 1994; Garet, Jing, and Kutner 1995; Maloney 1992; Murnane, Willett, and Boudett 1995; Murnane, Willett, and Boudett 1999; Murnane, Willett, and Tyler 1999; Passmore 1987; Reder 1994; Sum 1996; Tyler, Murnane, and Willett 1999).

² The literature test is sometimes characterized as primarily a test of reading comprehension.

this last test the “literature” test. In many states GED candidates have the option to take only a portion of the entire battery at any one testing.³ In all states passage is based upon some combination of the minimum test score out of the five tests and the mean score over the five tests. The GED Testing Service sets a minimum passing standard, but individual state education departments are free to set higher standards if they so desire. Table 1 shows the time-pattern of GED passing standards by state from 1980 to 1990⁴ and indicates that most states require a minimum score of at least 35 and a mean score on the five tests of 45.

³ Conversations the authors have had with adult education practitioners reveal, however, that GED candidates in some local settings are not informed of the option to take only a part of the battery, while some states have a de facto policy that all tests are to be taken in one sitting. Florida is such a state.

⁴ There were few changes in state standards between 1990 and January 1, 1997. On the first of January 1997, the GED Testing Service required minimum was raised to the “40 and 45” level of states such as New York, Florida, and California.

Table 1. State GED passing standards by year.

| Standard ² | State by Year ¹ | | | | | | | | | | |
|-----------------------|--|--|--|--|--|--|--|--|--|--|--|
| | 1980 | 1981 | 1982 | 1983 | 1984 | 1985 | 1986 | 1987 | 1988 | 1989 | 1990 |
| 35 or 45 | LA, MT, WV | | | | | | | | | | |
| 40 or 45 | MS, NE, TX | LA, MS, NE, WV, TX | LA, MS, NE, WV, TX | LA, MS, NE, WV, TX | LA, MS, NE, WV, TX | LA, MS, NE, WV, TX | LA, MS, NE, TX | LA, MS, NE, TX | LA, MS, NE, TX | LA, MS, NE, TX | LA, MS, NE, TX |
| 45 avg | SC | SC | SC | SC | SC | SC | SC | SC | SC | SC | SC |
| 40 each | OR | OR | OR | OR | OR | OR | OR | OR | OR | OR | OR |
| 35 & 45 | AL, AK, AZ, AR, CA, CO, CT, DC, GA, HI, ID, IL, IN, IA, KS, KT, ME, MA, MI, MN, MO, NY, NC, OH, OK, PA, RI, TN, VT, VA, WA, WI, WY | AL, AK, AZ, AR, CA, CO, CT, DC, GA, HI, ID, IL, IN, IA, KS, KT, ME, MA, MI, MN, MO, NY, NC, OH, OK, PA, RI, TN, VT, VA, WA, WI, WY | AL, AK, AZ, AR, CA, CO, CT, DC, GA, HI, ID, IL, IN, IA, KS, KT, ME, MA, MI, MN, MO, NY, NC, OH, OK, PA, RI, TN, VT, VA, WA, WI, WY | AL, AK, AZ, AR, CO, CT, DC, GA, HI, ID, IL, IN, IA, KS, KT, ME, MA, MI, MN, MO, NY, NC, OH, PA, RI, TN, VT, VA, WA, WI, WY | AL, AK, AZ, AR, CO, CT, DC, GA, HI, ID, IL, IN, IA, KS, KT, ME, MA, MI, MN, MO, NY, NC, OH, PA, RI, TN, VT, VA, WA, WI, WY | AL, AK, AZ, AR, CO, CT, DC, GA, HI, ID, IL, IN, IA, KS, KT, ME, MA, MI, MN, MO, NY, NC, OH, PA, RI, TN, VT, VA, WA, WI, WY | AL, AK, AZ, AR, CO, CT, DC, GA, HI, ID, IL, IN, IA, KS, KT, ME, MA, MI, MN, MO, NY, NC, OH, PA, RI, TN, VT, VA, WA, WI, WY | AL, AK, AZ, AR, CO, CT, DC, GA, HI, ID, IL, IN, IA, KS, KT, ME, MA, MI, MN, MO, NY, NC, OH, PA, RI, TN, VT, VA, WA, WI, WY | AL, AK, AZ, AR, CO, CT, DC, GA, HI, ID, IL, IN, IA, KS, KT, ME, MA, MI, MN, MO, NY, NC, OH, PA, RI, TN, VT, VA, WA, WI, WY | AL, AK, AZ, AR, CO, CT, DC, GA, HI, ID, IL, IN, IA, KS, KT, ME, MA, MI, MN, MO, NY, NC, OH, PA, RI, TN, VT, VA, WA, WI, WY | AL, AK, AZ, CO, CT, DC, GA, HI, ID, IL, IN, IA, KS, KT, ME, MA, MI, MN, MO, NY, NC, OH, PA, RI, TN, VT, VA, WA, WI, WY |
| 40 or 50 | NM, ND, SD | NM, ND, SD | NM, ND, SD | NM, ND, SD | NM, ND, SD | NM, ND, SD | NM, ND | NM, ND | NM, ND | NM, ND | NM, ND |
| 40 & 45 | DE, FL, MD, UT | DE, FL, MD, NJ, OK, UT | DE, FL, MD, NJ, OK, UT | CA, DE, FL, MD, NJ, OK, UT | CA, DE, FL, MD, NJ, NY, OK, UT | CA, DE, FL, MD, NJ, NY, OK, UT | CA, DE, FL, MD, NY, OK, SD, UT, WV | AR, CA, DE, FL, MD, NY, OK, SD, UT, WV | AR, CA, DE, FL, MD, NY, OK, SD, UT, WV | AR, CA, DE, FL, MD, NY, OK, SD, UT, WV | AR, CA, DE, FL, MD, OK, SD, UT, WA, WV |
| 40 & 50 | | | | | | | | | | | |

More
compli-
cated
passing
standards:
New
Jersey

40 on tests
1-3, 44 on
test 4, 43 on
test 5.

40 on tests 1-
3, 44 on test
4, 43 on test
5.

40 on tests 1-
3, 44 on test
4, 43 on test
5.

42 on test
1, and 40
on tests 2-4,45 on
test 5 & a total
score of 225.
test 5 & a
total
score of
225.

42 on test 1, and 40
on tests 2-4,45 on
test 5 & a total
score of 225.

1. If no states changed standards from the previous year, then the year-column is blank.

2. The first number in each standard represents the required minimum score on any one of the five tests in the battery, while the second number represents the lowest mean score over the five tests that is acceptable.

The probability that a dropout attempting the GED exams obtains the credential is a function of underlying cognitive skills, motivation to succeed at the tests, and of the state in which the attempt is made (as shown by Table 1). But GED attainment is also influenced by the fact that GED candidates who fail to meet passing standards can retake the exams an unlimited number of times, and by the fact that candidates do not have to retake all failing tests at the same time.⁵ Thus, unlike one-shot, one-test exams, the path to a GED is potentially complicated, with candidates making choices about retesting or quitting, as well as which tests to take at each attempt. The various features of the GED system raise two interesting possibilities. First, the opportunity to attempt the GED multiple times may have differential effects by gender, race/ethnicity, and age. Second, it is possible that different subject areas in the GED battery pose a greater hurdle to different gender, racial/ethnic, and age groups.⁶ Using test score data on GED candidates in Florida we are able to address these issues. We show that:

- the ability to retake the GED tests matters, and it matters especially for black and Hispanic dropouts, and
- the tests that prove to be the greatest stumbling blocks to passing the exams vary by gender, and somewhat by race/ethnicity.

In the next section we briefly discuss our data. In Section IV we analyze test taking patterns and in Section V we examine which GED subject areas tend to pose the highest hurdle for candidates. We summarize and discuss the implications of our findings in Section VI.

⁵ For a discussion of the affect that multiple opportunities has on “false positives,” see Millman (1989).

⁶ Baldwin (1992) shows that mathematics and writing tend to be more difficult for GED candidates than the literature, science, and social studies tests. That study, however, did not directly link lower scores in these subjects to the probability of passing the GED, and that study did not look at racial/ethnic and gender differences.

III. Data

The state department of education in Florida provided information on the test scores and demographic characteristics of all people who took the GED examinations in that state during the 1990s. Our data set is composed of individuals from these data who meet specific criteria. The critical criteria were that individuals initially tested no earlier than 1988 nor no later than 1995 and that any individual who retested did so within three years of their initial test date.⁷ Application of these criteria provided a sample of 189,124 individuals.⁸ Table 2 shows the distribution of initial and final test dates in our data.

Table 2. Distribution of initial and final test years.

| Years | Percentage distribution | |
|---------------|-------------------------|------------------|
| | Initial test years | Final test years |
| 1988 | 0.2 | |
| 1989 | 0.6 | |
| 1990 | 1.8 | |
| 1991 | 19.9 | 19.8 |
| 1992 | 19.2 | 19.0 |
| 1993 | 18.6 | 18.6 |
| 1994 | 19.7 | 19.7 |
| 1995 | 20.2 | 19.7 |
| 1996 | | 2.1 |
| 1997 | | 0.8 |
| 1998 | | 0.2 |
| Column totals | 100.0 | 100.0 |

As a result of the way GED data is stored in Florida, our data do not contain the complete test histories of the candidates. Rather, they contain only the initial test scores and the cumulative best scores on each test as of the last test date.

⁷ We limited the data in this manner so that all individuals, regardless of when they initially tested, had the same amount of time to retest. Since the data we received from Florida included individuals who *last tested* in the years 1991 to 1998, we set a lower *initial test year* of 1988 and an *upper initial test year* of 1995.

⁸ To test the sensitivity of our results to our sample definition we formed a data set using only individuals who initially tested in 1991. These individuals had seven years (rather than the more restrictive 3 years) to complete their “GED cycle” in our data, 1991-1998. We obtain essentially the same results using these data that are less restrictively defined.

Thus, in the Florida data we can determine whether individuals tested once or more than once, and we can determine how long was the period between the first and last attempt. However, for those who tested more than one time we cannot determine how many times they tested, nor can we determine the scores from any single test attempt other than the first.

Table 3 shows that our sample contains 189,124 individuals, divided almost equally between females and males. The sample is about 67 percent white, 17 percent black, and 17 percent Hispanic.⁹ The median age in the sample is 21. About 70 percent of the GED candidates met the passing score requirements on their initial attempt. The final GED passing rate, given the three-year attempt “window” that we allow, was 83 percent, giving us the first hint of the potential importance of the retesting feature of the GED. The average mean score over the five tests is about 50 for the entire sample, 52 for those who were eventually awarded a GED, and 40 for those who never managed to pass the exams. The remainder of the paper examines factors that contribute to the ultimate success or failure of dropouts who attempt the GED.

⁹ Individuals who were not classified as white, black, or Hispanic constituted about 2 percent of the data in Florida. We exclude these individuals to simplify the analysis and discussion.

Table 3. Sample descriptive statistics for dropouts who tested in Florida between the years 1988 and 1998.

| | | Florida 1988 cohort |
|--|---------------------------------------|------------------------|
| Number of observations | | 189,124 |
| Percent who are... | female | 48.5 |
| | white | 66.7 |
| | black | 16.5 |
| | Hispanic | 16.8 |
| | pass on their initial attempt | 70.4 |
| | eventually pass the GED exams | 83.0 |
| Mean test scores on the complete test battery... | | |
| | in the full sample | 50.2 (6.9) |
| | Among passers | 52.1 (5.4) |
| | Among failers who completed all tests | 40.2 (5.1) |

IV. The Role of Retesting in the GED System

A. The Potentially Complicated Route to a GED

As stated earlier, the GED program is not a one-shot, one-test system. The GED Testing Service guidelines relating tests, scores, and awarding of a GED have the following unique components.

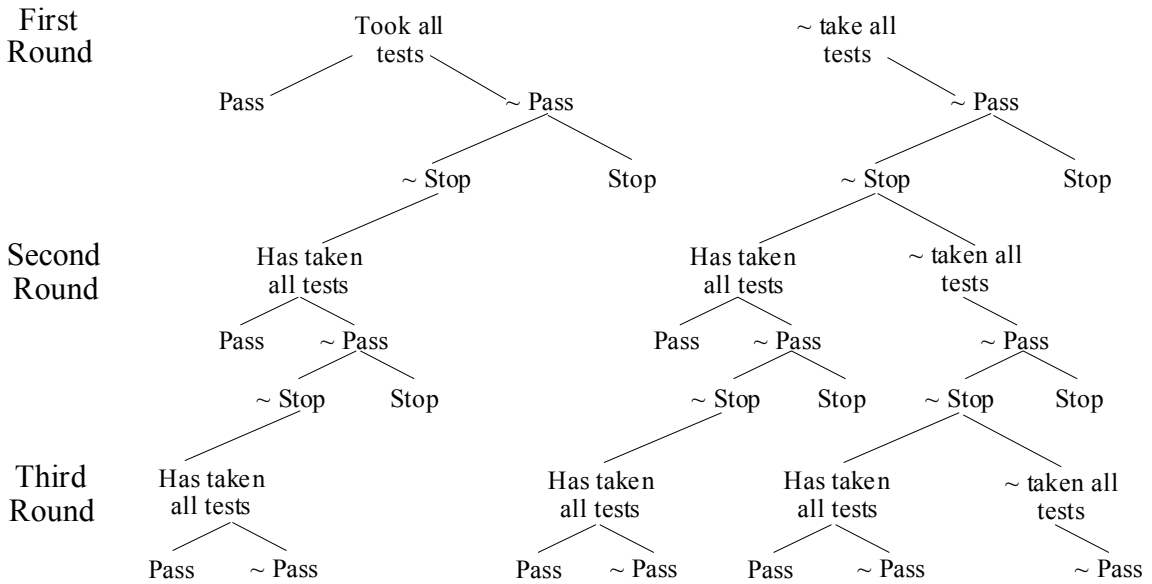
- In most states, individuals can take all five tests of the GED battery at any one attempt, or they can take any chosen subset of the GED battery.
- Individuals who do not achieve scores high enough to pass can retake all or any portion of the GED battery multiple times.¹⁰
- Awarding of the GED certificate is based upon state standards¹¹ that are generally a function of the mean score over the five GED tests and the minimum score out of all of the five tests. The scores used to compute the mean and minimum are based upon the highest scores attained out of all attempts--that is, the accumulated maximum test score in each of the five tests.

As a result of these guidelines, the route to acquiring a GED can be quite complicated for some individuals. Figure 1 shows the possible routes to GED certification for up to three attempts.

¹⁰ While there is considerable variation in state guidelines for retesting, an overarching principle is that individuals may not attempt the GED more than three times within a 12 month period. This is because there are three different forms of the GED exams and GED Testing Service rules stipulate that an individual may not test on the same form within a 12 month period (GED Testing Service 1991, Section 5.3). Florida guidelines stipulate that individuals who fail *a second attempt* have to wait for 6 months before testing again if their overall mean score was less than 43 and if their score on at least three of the tests was not 40 or more (GED Testing Service 1991).

¹¹ These state standards are all at least as high as the GED Testing Service mandated minimum. Until January 1, 1997 that minimum passing standard was a score of at least 40 on each of the five tests in the battery *or* an average score of at least 45 over the five tests (GED Testing Service 1991, Section 5.5). On January 1, 1997, the GED Testing Service minimum requirements were raised so that after that date individuals must have a minimum score of at least 40 *and* a mean score of at least 45.

Figure 1. Patterns of possible GED testing and retesting



An individual's path to a GED begins with the decision on the first attempt of whether to take all or only a portion of the GED exams. In Florida, this decision is sharply curtailed, as individuals are often not offered the opportunity to take only portions of the exam.¹² As illustrated in Figure 1, individuals face decisions at each juncture on the route to a GED. In the first round of testing individuals decide whether or not to take all of the tests in the battery. To pass, individuals have to meet the state score requirements on the GED battery. Since 1988 this battery has been composed of:

¹² 96 percent of the GED candidates in our Florida data took the complete battery on their initial attempt. By contrast, our tabulation of similar data from Texas indicates that about 75 percent of the GED candidates in that state took the complete battery on first attempt. Conversations with officials in Florida indicate that the high percentage of GED candidates who attempt all five tests in the battery at the same time is a result of custom and the direction of GED program providers rather than of official state guidelines (Posco, 1999).

- a writing test containing 55 multiple-choice items and an essay, with a total allotted time of 120 minutes;
- a social studies test containing 64 multiple-choice items--allotted time of 85 minutes;
- a science test containing 66 multiple-choice items--allotted time of 95 minutes;
- a test on interpreting literature and the arts containing 45 multiple-choice items--allotted time of 65 minutes, and;
- a mathematics test containing 56 multiple-choice items--allotted time of 90 minutes.

For those who fail, either because of low scores or because they did not take all tests in the battery, there is the decision of whether to quit at that point (“stop-out”) or to attempt the GED again. The stop-out/retake decision is then faced by non-passers in each subsequent round.

B. Passing, Failing, and Retesting in the GED Program

Given the 7-hour 35-minute length of the exams, individuals might rationally choose to attempt only a portion of the entire battery on the first try as a way of maximizing their scores on any one test. As we have pointed out, however, these decisions are not always open to candidates in Florida. Thus, most individuals in our data are in the left-hand branch of the pass-fail diagram in Figure 1, and as a result the initial GED status of most candidates is a function of their scores on the tests rather than whether or not they attempted all tests in the battery.

The first part of this paper will address three questions that have not been studied before: who tends to pass on the first try at the GED, who tends to retest upon failing the first attempt, and who tends to pass given the inclination and opportunity to retest. Table 4 provides direct evidence on the first question.

Table 4. Percentage of candidates who passed on their initial attempt and ultimately, by racial/ethnic group and age. (* indicates that the difference in the adjacent percentages are significant at the 0.01 level).

| Racial/Ethnic Group | Percentage who passed | | | | | |
|--|-----------------------|--------|--------|--------|--------|------|
| | All ages | Age | | | | |
| | | 16-18 | 19-21 | 22-24 | 25-30 | 30+ |
| Panel A: Initial pass rates | | | | | | |
| (1) Whites | 77.4 | 76.9 * | 75.1 * | 79.7 | 79.3 | 79.1 |
| (2) Blacks | 46.2 | 51.4 * | 47.1 | 48.5 * | 45.8 * | 40.0 |
| (3) Hispanics | 66.1 | 68.7 * | 62.1 | 63.6 * | 66.6 * | 70.7 |
| Panel B: Ultimate pass rates | | | | | | |
| (4) Whites | 88.1 | 89.3 * | 86.8 * | 88.6 | 88.0 | 88.6 |
| (5) Blacks | 66.3 | 71.9 * | 66.8 | 67.1 * | 64.3 * | 57.6 |
| (6) Hispanics | 80.4 | 85.3 * | 78.5 | 77.5 | 78.1 | 81.3 |
| Panel C: Difference in initial and ultimate pass rates | | | | | | |
| Whites Row 4 - Row 1 | 10.7 | 12.4 | 11.7 | 8.9 | 8.7 | 9.0 |
| Blacks Row 5 - Row 2 | 20.1 | 20.4 | 19.7 | 18.6 | 18.5 | 17.6 |
| Hispanics Row 6 - Row 3 | 14.3 | 16.6 | 16.4 | 13.9 | 11.5 | 10.6 |

Table 4 is divided into three panels. Panel A gives figures on *initial* passing rates by race/ethnicity and age. Panel B gives figures on *final* passing rates for the same groups. And for convenience, Panel C shows the differences between initial and final passing rates by racial/ethnic group and age. Within a panel, looking across a row of Table 4 gives within race/ethnicity group information about either the initial (Panel A) or the final passing rates (Panel B) of Florida dropouts by age group. Looking down any column allows for within-age-group pass rate comparisons across race/ethnicity.

The first column in Panel A indicates that there are substantial racial/ethnic differences in initial GED passing rates in Florida.¹³ The initial pass rates of white dropouts of all ages are 30 percentage points higher than the initial pass rates of black dropouts (77.4 percent versus 46.2 percent) and 11 percentage points higher than those of Hispanic dropouts (66.1 percent).

These initial passing percentages computed across all age groups mask interesting differences by age across the racial/ethnic groups. First, young black candidates (age 16-18) pass at a higher rate than any other age group of black GED testers (51.4 percent), while older black candidates (30+) pass at the lowest rates of all candidates, regardless of race/ethnicity (40.0 percent). For white and Hispanic testers, however, the highest initial passing rates are found among older testers. For white GED candidates, initial passing rates are highest beginning at age 22 (differences in the mean passing rates of whites age 16-18 and whites age 30+ are statistically significant at the 0.001 level). For Hispanics, however, the higher passing rates are only among the 30+ age group (differences in the mean passing rates of Hispanics age 16-18 and Hispanics age 30+ are statistically significant at the 0.01 level).

We cannot determine with these data what is generating these age by race/ethnicity patterns in GED pass rates. It could be that young black GED candidates are more skilled than are older black candidates, while the opposite is true for whites and Hispanics. It could also be that the level of test preparation is different by age for blacks versus whites and Hispanics. Whatever the cause, it is important to recognize that GED success rates are not simply functions of race/ethnicity, but that there are also age by race/ethnicity interactions, a fact that GED program providers should consider.

An examination of the entries in Panel B indicate that all racial/ethnic-age groups improved their passing rates as a result of retesting. However, blacks made the most substantial gains. Computed over all age groups, the opportunity to retest allowed blacks to reduce the initial white-black passing rate gap by one-third. The ultimate passing differences between blacks and whites is about 22 points (Panel B), as opposed to the 30 percentage point white-black initial passing rate gap (Panel A). Hispanic testers closed the white-Hispanic initial passing gap from 11 percentage points to 8 percentage points with retesting.

The fact that all groups improved considerably between initial testing and final results after retesting is made clear in Panel C, which simply shows final minus initial passing rates. Comparing the row entries in Panel C reemphasizes

¹³ There are essentially no gender differences in first or ultimate pass rates in the Florida data.

the extent to which black GED candidates benefited from retesting, especially relative to white dropouts.

Table 5 helps to explain some of the racial/ethnic differences in the transition from initial to ultimate pass rates. Given initial pass rates, there are two factors that determine ultimate pass rates: the proportion of first-attempt failers who choose to retest and the proportion of these candidates who ultimately pass. The two panels in Table 5 address these two issues.

The first column in Panel A of Table 5 shows that relative to black and Hispanic dropouts, a higher percentage of the white GED candidates who initially failed the GED retested at least once within three years of their initial attempt. For all racial/ethnic groups it was the youngest GED candidates, those who were 16-18 years of age on their initial test date, who tended to reattempt the GED at the highest rates.

Panel B of Table 5 shows that not only did white candidates tend to retest at higher rates, their pass rates were also higher, conditional upon retesting. However, both black and Hispanic GED candidates substantially closed the passing rate gap with whites as compared to the group passing rate differences on the initial attempt. For example, we saw earlier that the percentage of blacks who passed on the initial attempt was about 30 percentage points lower than the percentage of whites who passed on the first try. The first column of Table 5 shows that among retesters, however, the black-white gap was only 11 percentage points (60 versus 71). Meanwhile, the Hispanic passing rate upon reattempt was nearly identical to that for white retesters.

The passing rates of white and Hispanic candidates who retested (Panel B of Table 5) tended to be higher than the passing rates of black retesters. However, since the initial failing rates were so much higher for black candidates (see Table 4) their retesting and passing rates conditional upon retesting still meant that they made up considerable ground in GED acquisition rates. It is notable that the retesting rates for all groups were highest among the youngest GED candidates, and that the passing rates upon retesting also tended to be higher for the 16-18 year-old age group.

Table 5. (1) Proportion of candidates who retest within three years after a failing first attempt, and (2) the proportion of candidates who pass conditional upon making at least one attempt after a failing first attempt. (* indicates that the difference in the adjacent percentages are significant at the 0.01 level, ~ at the 0.10 level).

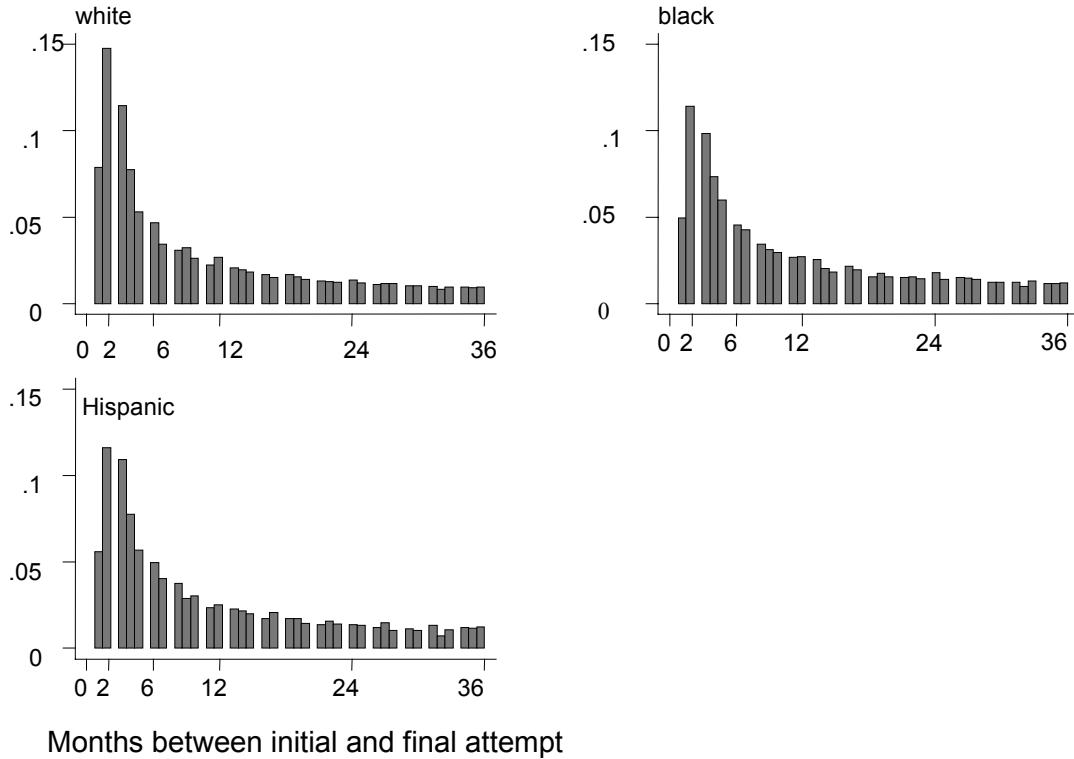
| Racial/Ethnic Group | Percentage of initial failers who retested | | | | | |
|---|--|--------|--------|--------|--------|------|
| | All ages | Age | | | | |
| | | 16-18 | 19-21 | 22-24 | 25-30 | 30+ |
| Panel A: | | | | | | |
| (1) Whites | 65.4 | 71.9 * | 65.3 * | 60.8 | 59.9 | 61.4 |
| (2) Blacks | 58.7 | 65.0 * | 61.5 * | 56.5 | 56.1 | 54.1 |
| (3) Hispanics | 59.9 | 70.2 * | 61.5 * | 54.7 | 51.7 ~ | 54.8 |
| Percentage of retesters who ultimately passed | | | | | | |
| Panel B: | | | | | | |
| (4) Whites | 71.3 | 73.9 * | 71.0 | 71.0 | 69.2 | 68.6 |
| (5) Blacks | 59.9 | 64.5 * | 59.7 ~ | 63.3 * | 60.8 * | 53.9 |
| (6) Hispanics | 69.7 | 75.0 * | 69.7 | 69.4 | 65.7 | 64.8 |

Our data do not allow us to determine how many times on average these retesters attempted the GED. The Florida data only have information on the initial and the final attempt. However, we can determine the length of time between a candidate's first and last attempts. The time between attempts is somewhat conditioned by state regulations. In Florida, individuals may retest

...at the discretion of the local Chief [GED] Examiner. After the second testing they are eligible to retake the test if they have made an overall minimum standard score of 215 [total on the five tests] and have achieved a standard score of at least 40 on three (3) or more of the tests. Any candidate who fails to achieve this standard after the second testing is not eligible to retest for six (6) months. (GED Testing Service 1991)

Figure 2 shows the distribution of time between attempts for white, black, and Hispanic retesters.

Figure 2. Distributions by race/ethnicity of months between initial and final attempt for initial failers who retested at least once within three years.



From Figure 2 we see that the highest percentage of retesters, regardless of race/ethnicity, made their final attempt within two months of their initial attempt.¹⁴ Conversations with officials in Florida suggest that the high percentage of failing candidates who retest within two months is because this is about how long it takes to grade the exams, do the paper work, and inform the candidates of their scores (Posco, 1999). While the mean time between first and last attempt is about one year for all groups, the median time is six months for white, seven months for blacks, and eight months for Hispanic retesters. We found little difference across age groups in the time between test attempts.

¹⁴ These graphs look a little different if they are conditioned on whether or not the candidate failed or passed on their final attempt.

Section V. Which Tests Prove to be the Highest Hurdles for GED Candidates?

Using a sample of GED candidates, an earlier study conducted by the GED Testing Service shows that the math and writing tests tend to be the tests on which GED candidates score the lowest (Baldwin 1992). This suggests that these tests might be the highest hurdles for GED candidates. While suggestive, however, this earlier study does not lead to definitive conclusions regarding which of the five subject areas of the GED exams—math, literature, writing, social studies, and science—prove to be the highest hurdles. Neither does the earlier study provide information on whether different demographic subgroups have relatively more difficulty with different tests in the GED battery. We can use the Florida data to get at these questions.

As a first step in examining which tests are important determinants of who does or does not obtain a GED, we examine the role played by the Florida passing standards itself. In Florida, as is true in most states, the GED passing standard is a function of an individual's mean score computed over the five tests in the GED battery and their minimum score out of all tests in the battery. Over the period spanned by our data, Florida required a mean score of at least 45 and a minimum score of at least 40. Thus, an individual could fail to obtain a GED either because (1) their minimum score was too low, (2) their mean score was too low, (3) both scores were too low, or (4) they never completed all five tests in the battery.

Table 6 shows the distributions across these results for (1) those who failed on their initial attempt and (2) those who failed on their final attempt, conditional upon retesting at least once.

Table 6. Proportion of candidates in Florida who fail as a result of (1) low minimum scores, (2) low mean scores, (3) both low minimum and low mean score, or (4) failure to complete the battery.

| Reason for failing | Percentage distribution | |
|---|-------------------------|-----------------------|
| | 1st failing attempt | Final failing attempt |
| Low min. score | | |
| Mean score was ≥ 45 , but Min. score was < 40 | 14.0 | 7.3 |
| Low mean score | | |
| Min score was ≥ 40 , but Mean score was < 45 | 16.6 | 39.1 |
| Both low min & low mean | | |
| Mean score was < 45 and Min score was < 35 | 56.7 | 50.4 |
| Failed to complete all five tests in the battery | 12.7 | 3.3 |

Most individuals fail because they can neither meet the minimum score nor the mean score requirements, and this is regardless of the first or initial attempt. However, among individuals who can meet one but not the other of these hurdles, there are substantial differences between initial and final failing attempts. Among those who fail on their first attempt, only slightly more do so because they cannot score at least a 45 mean score versus those who fail because they cannot score at least a 40 minimum score (16.6 percent versus 14 percent). However, among those who retry, but still fail, a much higher percentage are unable to meet the mean relative to failure to meet the minimum score requirement (39.1 percent versus 7.3 percent). Note also, that the percentage who fail because they have not finished the battery drops considerably from first to final attempt, as we would expect.

In results not detailed here, we find that the only patterns by race/ethnicity in “reasons for failing” is that relative to blacks and Hispanics, a higher percentage of whites tend to fail because they can meet the mean requirement, but cannot meet the minimum score requirement. Likewise, we find that the only differences by age are that higher percentages of older white and Hispanic testers are able to meet the mean, but fail the minimum score requirement, relative to black testers. There are only small gender differences in reasons for failing.

We turn now to an examination of the individual tests, seeking to determine which are the highest hurdles for candidates who failed to ever achieve a GED, and how the answer might vary by demographic group. Table 7 gives information from the Florida data on the lowest scores of *failing* GED candidates in that state. The first row of the table shows that among all failing GED candidates, the lowest test scores tend to be on either the math or the writing exams (34 and 23 percent respectively).¹⁵ Closer examination of the data reveals substantial gender differences among the failing Florida candidates regarding which tests proved the greatest barrier to passing.

The second row of Table 7 shows that the greatest percentage of males who failed had a lowest score on their writing test (31 percent). The third row indicates that the lowest score for failing females occurred on their math test (44 percent). While not shown in Table 7, we find that for females, these effects tend to get larger with age. For example, 38.7 percent of the females age 16-18 who failed had a lowest score on their math exams. For failing females age 30+, however, that figure is 47.9. Fully 50 percent of the white females in the 30+ age group who failed their GED had a lowest score on their math exams. We did not find similar age effects among unsuccessful male candidates.

¹⁵ Our examination of the data indicate that either math or writing or both were the lowest test scores for the great majority of individuals whose lowest scores occurred on two or more tests (the last column in Table 7).

Table 7. Lowest score distributions for candidates who failed to attain a GED, by demographic group.

| Group | Test that was the lowest score | | | | | |
|-----------|--------------------------------|------------|---------|----------------|---------|------------------------------------|
| | Math | Literature | Writing | Social Studies | Science | Multiple tests with same low score |
| All | 33.5 | 16.4 | 22.8 | 8.8 | 6.1 | 12.5 |
| Males | 23.5 | 19.6 | 31.3 | 8.9 | 5.1 | 11.7 |
| Females | 43.6 | 13.1 | 14.2 | 8.7 | 7.1 | 13.3 |
| Whites | 35.4 | 14.6 | 24.6 | 8.7 | 4.7 | 12.1 |
| Blacks | 34.6 | 16.3 | 20.0 | 9.1 | 7.1 | 13.0 |
| Hispanics | 28.5 | 19.2 | 23.4 | 8.7 | 7.6 | 12.7 |
| Other | 24.9 | 24.4 | 24.7 | 8.1 | 6.4 | 11.5 |

The bottom part of the table shows that there are no racial/ethnic patterns as pronounced as the gender patterns in lowest test score. We do note that a higher percentage of Hispanics tend to have lowest scores on either the literature or the writing exams, relative to the other race/ethnicity groups.

VI. Summary and Implications

The evidence from the GED examination system in Florida indicates that the details of examination systems matter. In particular, a mathematics examination tends to pose particular obstacles for females and a writing examination for males. To the extent that adult education practitioners do not currently take these differences into account, GED pass rates could possibly be improved if more gender specific instruction were provided in adult education classes. The lesson for state-based examination systems is that, depending on the design of the examination system (e.g., how different subject areas are weighted in computing the “passing score”) potential gender differences by subject may affect success rates by gender.

The GED testing experience also documents the critical role of the re-testing option. Without this option, not only would the number of test takers who received the credential in Florida have been 13 percentage points lower (70 versus

83 percent overall pass rates), the gap between the success rate of white test takers and that of test takers of color would have been considerably larger.

We close by posing two potential concerns associated with any credentialing system that incorporates retesting. First, some attention should be given to the role of “false positive” credentialing. As Millman (1989) has shown, in an environment that allows retesting, some portion of individuals whose “true” skills lie below a given passing threshold will be able to capitalize on chance to achieve a passing score. That is, through repeated attempts, some individuals who should not pass will pass solely because of “random” high scores on repeated attempts. Since the GED system allows essentially unlimited retesting, this may be of some concern. Second, even if we thought that false positive certification was a minor problem in an examination system, it may be that upon repeated testing, individuals become “test wise.” The scores of some retesters improve because they have learned better test taking techniques, not because they have increased their skills. Of course this argument assumes that individuals’ “true” skills were measured on the first test. It might be that individuals score lower than their true skills on the first test because of “test anxiety” or unfamiliarity with the GED testing format. In this case, we may actually be more accurately measuring true skills with subsequent tests

More so than many other developed countries, our education system has traditionally been one of second chances. In this spirit, the opportunity to retest seems a natural component of a “high stakes” examination system. As a part of this second chance philosophy, we want to believe that individuals succeed on retests because they have increased their skills between attempts. To the extent that higher subsequent scores are not the result of increased skills, but rather a function of Millman’s “false positive” hypothesis or of adaptation to the test, retesting in a high stakes credentialing system may not always achieve desired ends. However, until further research can determine whether either of these potentially undesirable outcomes is a substantial problem associated with the GED, we have shown that the opportunity to retest is an important “second chance” component of the GED system. Given this result, it is likely that retesting, where allowed, would play an equally important role in other examination systems.

References

- Baldwin, Janet. *GED Test Performance of Adult Examinees and High School Seniors*. Washington, D.C.: GED Testing Service of the American Council on Education, 1992, 5.
- Bos, Johannes M. "The Labor Market Value of Remedial Education: Evidence from Time Series Data on an Experimental Program for School Dropouts." Unpublished Ph.D. dissertation, The Robert F. Wagner School of Public Service, New York University, 1995.
- Cameron, Steve V. "Assessing High School Certification for Women Who Dropout." Working Paper, Department of Economics, University of Chicago, 1992.
- Cameron, Steve V., and James J. Heckman. "Determinants of Young Male School and Training Choices." In *Skill Formation by the Private Sector: International Comparisons*, ed. Lisa M. Lynch. Chicago: University of Chicago Press, 1993a.
- Cameron, Stephen V., and James J. Heckman. "The Nonequivalence of High School Equivalents." *Journal of Labor Economics* 11, no. 1 (1993b): 1-47.
- Cao, Jian, Ernst W. Stromsdorfer, and Gregory Weeks. "The Human Capital Effect of the GED on Low Income Women." *Journal of Human Resources* 31, no. 1 (1996): 206-228.
- Cave, George, and Johannes Bos. *The Value of a GED in a Choice-Based Experimental Sample*. New York: Manpower Demonstration Research Corporation, 1994. Paper presented at the 106th Annual Meeting of the American Economic Association, Boston.
- Garet, Michael S., Zhongren Jing, and Mark Kutner. *The Labor Market Effects of Completing the GED: Asking the Right Questions*. Washington, D.C.: American Institutes for Research, 1995.
- GED Testing Service. "GED Examiner's Manual." Washington, D.C.: American Council on Education, 1991.
- Johnson, Mark, and Thomas Valentine. *Outcomes of GED Graduation: An Annotated Bibliography of Research Reports*. Athens, Georgia: University of Georgia, 1992.

Maloney, Tim. "Estimating the Returns to a Secondary Education for Female Dropouts." Working Paper No. 100, Department of Economics, University of Auckland, 1992.

Millman, Jason. "If at First You Don't Succeed: Setting Passing Scores When More than One Attempt Is Permitted." *Educational Researcher* 18, no. 6 (1989): 5-9.

Murnane, Richard J., John B. Willett, and Katherine P. Boudett. "Do High School Dropouts Benefit from Obtaining a GED?" *Educational Evaluation and Policy Analysis* 17, no. 2 (1995): 133-147.

Murnane, Richard J., John B. Willett, and Kathryn Parker Boudett. "Do Male Dropouts Benefit from Obtaining a GED, Post-Secondary Education and Training?" *Evaluation Review* 23, no. 5 (October 1999): 475-502.

Murnane, Richard J., John B. Willett, and John H. Tyler. "Who Benefits from a GED? Evidence from *High School and Beyond*." *The Review of Economics and Statistics* 82, Issue 1 (February 2000): 23-37.

Passmore, David L. *Employment of Young GED Recipients*. Washington, D.C.: American Council on Education, 1987, 14.

Posco, Bob, Adult Education Section of the Florida State Department of Education, Personal Communication. , 1999.

Reder, Steve. *The Nature and Impact of GED Training in Oregon*. Portland, Oregon: Northwest Regional Education Laboratory, 1994.

Sum, Andrew. *Literacy in the Labor Force: Results from the National Adult Literacy Survey*. Washington, D.C.: National Center for the Educational Statistics, 1996.

Tyler, John H., Richard J. Murnane, and John B. Willett. "Estimating the Labor Market Signaling Value of the GED." *The Quarterly Journal of Economics* 115, no.2.(1999): 431-468.

U.S. Department of Education. *Dropout Rates in the United States: 1996*. Washington, DC: National Center for Education Statistics, 1997. NCES 98-250, by Marilyn M. McMillen and Phillip Kaufman.