CHAPTER 10
MEASURING PUPIL PROGRESSION

10.1 Introduction

Juvenile justice educational programs can be evaluated using a variety of different outcome measures. The Juvenile Justice Educational Enhancement Program's (JJEEP's) evaluation of juvenile justice educational programs includes annual quality assurance (QA) reviews, student educational performance within juvenile justice programs, and community reintegration measures. This chapter identifies the research methods that will be used to evaluate programs' pupil progression rates among students while they are incarcerated. For QA findings see Chapter Two and for research methods and findings on community reintegration, see Chapters 11 and 12.

To successfully validate best education practices in juvenile justice education, it is essential that JJEEP is able to determine students' academic gains while they are incarcerated. As described in Chapter seven, juvenile justice students, on average, are educationally deficient on several measures compared with their public school counterparts. To the extent that quality education accelerates students' academic gains, these deficiencies may be ameliorated before students return to the community and their home schools. Once the rate of pupil progression and degree of academic gains are determined for juvenile justice educational programs, several interesting comparisons may be made, including the comparison of program performance based on pupil progression rates and QA scores and a comparison between student academic gains and community reintegration.

Based on prior research, JJEEP's assumption is that quality educational programs produce higher student academic gains, which subsequently result in better community reintegration outcomes. To test this assumption, JJEEP plans to evaluate all long-term residential commitment and day treatment juvenile justice educational programs in Florida by assessing each program's average student gains among incarcerated students. Once this is accomplished, JJEEP will be able to determine the kind of relationship that exists between student academic performance and community reintegration. Results will be correlated with QA scores to determine if high QA performing programs produce higher academic gains per student than do lower QA performing programs.

This chapter is comprised of three subsequent sections. Section 10.2 describes Florida's current system for entry and exit academic assessment testing, and its associated problems. Section 10.3 outlines an alternative research method for measuring pupil progression rates and academic gains during incarceration as an effective alternative to using non-uniform test scores. Section 10.4 provides summary discussion of the research methods described in this chapter.

10.2 Entry and Exit Academic Assessment Measures

Section 1003.51, F.S. requires the Department of Education (DOE), in partnership with the Department of Juvenile Justice (DJJ), district school boards, and providers to develop procedures for the administration of entry and exit academic assessments. Rule 6A-6.05281, FAC further clarifies this requirement to include academic entry and exit assessments that measure student performance in the areas of reading, writing, and math. Additionally, the rule says that all residential commitment and day treatment programs are required to submit the test data on all students through the use of local school district management information systems (MIS) and be included in DOE Survey Five data.

In 2002, DOE also developed and disseminated a technical assistance paper (TAP) titled, *A Guide to Test Instruments for Entry and Exit Assessment in Florida Department of Juvenile Justice Educational Programs*. The TAP provides information about 32 different tests, which have been approved for use as entry and exit assessment measures in juvenile justice educational programs. The TAP also describes the process for reporting student test scores to DOE. The 32 different assessments are reported using different scores, including raw scores, percentile rankings, norm curve equivalents, and scale scores. Test scores are not allowed to be reported to DOE in the form of grade equivalencies. The TAP also includes a section on the proper administration of academic assessment tests.

This current system for assessing students at entry and exit from juvenile justice educational programs and the subsequent reporting of test scores has several problems that make the test data difficult to use for research particularly focused upon program comparisons.

Entry and Exit Assessment Testing Problems

In general, there are three significant problems with the current entry and exit assessment testing system. First, there are an inordinate number of approved assessment tests being used by facilities, and many of the assessment tests use different means to measure academic performance. Second, different types of test scores are used and reported to DOE, which makes reasonable and comprehensive program comparisons and interpretations virtually impossible. Third, the overall administration of the tests differs substantially from program to program.

As discussed previously, DOE in Memorandum 20016, Educational Services for Students in Department of Juvenile Justice Facilities—Student Progress, identified 32 different assessment tests approved for use. These tests vary considerably according to whether or not they are norm-referenced tests (NRTs) or criterion-referenced tests (CRTs). Such variation is problematic in that the tests are designed to measure different standards. While NRTs measure a student's performance in relation to his/her peers of the same age or grade level, the CRTs measure the student's performance relative to an established standard of performance (DOE TAP, 2002).
Further examination of the approved assessment tests reveals that there is considerable variation as to what the tests use to measure skills within the areas of reading, math, and writing/language arts. For example, in measuring reading, some tests look at reading

comprehension, others vocabulary, others fluency, and, still others, word recognition. Within those categories, there is further variation in measurement.

An additional concern is that the scoring varies substantially among the different testing instruments. Generally, while NRTs provide scores on a standard scale, CRT scores are generally pass/fail or based on a level (DOE TAP, 2002). Standard scores and grade equivalencies may be more comparable across different assessment instruments, however, programs are required to report test scores in the form of percentile rankings, scale scores, and norm curve equivalents, which are not comparable across different assessment instruments.[1]

A final administrative problem concerns the administration of entry and exit assessment tests. According to DOE (TAP, 2002), most commercially published assessment tests describe the type of training an individual must have in order to administer the test. According to the TAP, either an educational diagnostician or student services professional is qualified; however, there are many programs that do not have individuals with such training. In such cases, DOE recommends that teachers with special education training be used to administer the tests. Of concern, then, is how these assessments are administered, who is administering them, and the degree of consistency within and across programs for following the administrative guidelines of the assessment test instruments.

As an example of the problems with the testing, JJEEP examined 17 programs' administration of entry and exit assessment testing. The 17 programs used 10 different reading tests, nine math tests, and four writing tests. Additionally, they used at least five different types of scores, including the raw score, percentile rank, normal curve equivalent, standard score, and grade equivalent, making it impossible to reach any meaningful conclusions. Furthermore, JJEEP was not always able to discern what type of score was used. Some programs appeared to only administer a subtest or preliminary locator test instead of the full battery of the test assessment. One program administered a teacher-made test assessment that was not standardized or normed on any population.

Importantly, there were other problems with administering the test at both entry and exit time periods. With the reading tests, eight of the 17 programs (57%) had administered the tests to less than 50% of their students at both entry and exit time periods. Similarly, six of the 17 programs (35%) had administered the math tests to less than 50% of their students at both entry and exit time periods. Finally, only 10 of the 17 programs (59%) administered a writing test, and of those 10, seven (70%) administered the tests to less than 50% of their students at both entry and exit time periods. Some programs were also found to be using a different assessment at exit when compared to the one they used at entry.

Given these problems of numerous testing instruments, inconsistencies in reporting the scores, and varying qualifications of those administering the tests, it is inappropriate to use

[1] There are certain statistical procedures that could be used in some cases to "transform" scores from different assessments and place them on a normal or bell curve, which would make comparisons possible. These transformations, however, can do nothing for the measurement of different competencies in the subject area and should only be used with extreme caution.

the current assessment test data for examining student or educational program performance. In order for these tests to be of use, the current assessment system must be reexamined and modified to be a more consistent, valid measure of the academic gains of students and the performance of juvenile justice educational programs.

10.3 An Alternative Method for Measuring Pupil Progression

Given the previously identified problems with academic assessment tests, JJEEP is employing an alternative strategy for measuring academic gains. Clearly, academic testing is not the only measure of student and educational program performance. Portfolio assessment, which uses several qualitative measures of a student's performance, is often argued to be a more comprehensive and higher quality measure of student performance. Portfolio assessment is expensive, however, and difficult to implement consistently across multiple programs and schools. In consideration of these two factors, in future research efforts JJEEP plans to assess pupil progression and academic gains through the use of DOE student data, which is described in detail below.

The questions JJEEP plans to answer through pre and post juvenile justice education program research include:

1. How much academic progress do students make in a given semester while incarcerated?
2. Do high QA performing programs produce greater academic gains per student than low QA performing programs?
3. And finally, do significant academic gains made during commitment increase community reintegration success?

Because juvenile justice students enter and exit commitment programs at random throughout the year and because most juvenile justice schools operate on traditional school semesters, it is difficult to determine credits earned, advancement to the next grade level, and diplomas earned for students who enter and exit during the middle of school semesters. Therefore, the method that JJEEP will employ is based on the full semester enrollment of juvenile justice students. Specifically, JJEEP will select a pool of juvenile justice students by identifying all students who were enrolled for one complete semester (one pool for each fall and spring semesters) in all long-term residential commitment and day treatment programs throughout the state. This should include approximately 150 juvenile justice schools.

The majority of Florida high school graduation requirements focus on academic courses and passing scores on the Florida Comprehensive Assessment Test (FCAT). Therefore, when assessing credits earned, JJEEP is selecting only core academic and vocational course credits. Specifically, this includes credits in English, math, social studies, science, and vocational course work. The number of credits earned in these core areas will be calculated for all students enrolled in either the fall or spring semester. This information is contained in the DOE transcript file.

The second measure of academic gains is advancement to the next grade level. At the end of the spring semester, DOE data provide a variable that includes students who were promoted to the next grade level, students who were promoted to the next grade level without meeting academic performance requirements, and students who were retained in the same grade level. Using this variable, the number of students who advanced to the next grade level at the end of the spring semester will be calculated for every student enrolled for the entire semester in all long-term residential commitment and day treatment juvenile justice schools.

The final measure of academic gains will include diplomas and certificates earned during commitment. Diplomas earned include standard high school diplomas, district high school diplomas (through the General Educational Development [GED] Exit Option), and GEDs. Certificates earned include any secondary or post-secondary vocational certificate. Age and eligibility for graduation will have to be considered when determining student and juvenile justice school performance based on this measure. For example, students eligible to earn a high school diploma during the year of the study must have earned enough credits to be considered a senior. Students wanting to use the GED Exit Option cannot use it as a means of early release from school, and students wanting to earn a GED must be at least 16 years of age.

After determining individual student pupil progression rates, the data will be aggregated back to the educational program level and pupil progression rates for juvenile justice schools will be determined. This aggregated pupil progression rate for juvenile justice schools will be compared to their QA performance. This type of analysis will serve to validate the QA standards and the best practices that make up those standards.

Overall, then JJEEP will utilize credits earned, advancement to the next grade level, and diplomas and certificates earned as measures of each program's average academic gains and pupil progression rate. Once the pupil progression research has been completed, JJEEP will examine the relationship between educational programs' average academic gains and their students' community reintegration measured in relation to the multiple indicators of recidivism/recommitment, education, employment, and other self-report and guardian survey findings. Moreover, JJEEP will assess the relationship between the programs' average academic gains and their QA scores. With these analyses, JJEEP will be able to determine whether community reintegration, as hypothesized, is associated with academic gains and quality educational programs.

10.4 Summary Discussion

An inherent weakness in the research methods described in this chapter is the possibility of grade inflation or pupil progression that is without academic merit. To control for this, it is important that JJEEP be able to evaluate pupil progression not only by using credits earned, advancement to the next grade level, and graduation rates, but also to have comparable entry and exit assessment test data. Reliable test data, along with these other measures, would provide the State of Florida with the means to more accurately assess actual juvenile justice educational program performance and associated student academic gains.

To validate best practices and generate accurate information associated with positive student outcomes, it is essential that JJEEP conduct pupil progression research over the next two years. To validate the pupil progression results, it is important that DOE, in collaboration with JJEEP, DJJ, school districts, and educational providers, develop and implement a uniform entry and exit assessment testing system that will be able to accurately measure the progress that each student experiences while in a juvenile justice educational program.