Using Decision Trees to Re-evaluate the Impact of Title 1 Programs in the Windham School District

Eugene W. Wang, Ph.D.
Associate Professor
Associate Director,
IMMAP
eugene.wang@ttu.edu
806-834-7738

Rong Chang, M.Ed.
Steven R. Chesnut, M.Ed.
Samuel Meeks, B.S.
Avery D. Miller, M.Ed.
INTRODUCTION

The purpose of this evaluation was to identify the impact of Windham School District (WSD) Title 1 services on academic outcomes, specifically GED attainment, and grade-level gains on the Composite, Reading, Writing, and Math tests of the Test of Adult Basic Education (TABE). “Gains” were computed as the difference between the initial TABE and the highest TABE.

The primary predictors of interest were:

- WSD programming (Academic, Vocational, ESL, SPED, Computer Lab, CIP, and CHANGES), but particularly total number of hours of Title 1 services
- Offender type (ID/State Jail/SAFP), gender, race, age at end of the last school year, and sentence length.

HIGHLIGHTS:

- Title 1 hours were significantly related to grade-level gains on the Test of Adult Basic Education (TABE)
- There was an 8.5:1 return on reading grade-level gains from Title 1 hours up to 172 hours, and 3.7:1 from 172-484 hours.
- There was a 13.1:1 return on math grade-level gains from Title 1 hours up to 128 hours, 10.4:1 from 128-207 hours, and 5.3:1 from 207-373 hours.
- There was a 9.5:1 return on composite grade-level gains from Title 1 hours up to 172 hours, and 4.5:1 from 172-499 hours.
- In most cases, a small amount of WSD programming (Title 1 and other WSD programs) show substantial increases in TABE gains and/or GED receipt.

CAVEATS:

- Outside of Title 1 hours, differences were most evident by offender type and sentence length: state jail offenders had smaller gains but also shorter sentences and thus less opportunity for participation in Title 1 hours.
- As outcome measure, grade level “gains” may not be as useful as highest score.
ABOUT THE SAMPLE

Deidentified data were provided to TTU IMMAP from WSD. The sample included all offenders in School Years (SY) 2011, 2012, 2013, and 2014 WSD accountability data that were 22 years old or less as of 8/31 of each respective year. To be included in the WSD accountability data an offender must have been an academic participant (regular academic, ESL, SPED, and/or Title 1), and had two TABE tests during the school year, or a baseline test from a previous school year and a subsequent test during the SY being reviewed. An offender may or may not have met the accountability criteria for three (3) consecutive school years. Data were included for the year(s) the offender met the accountability criteria.

Below are descriptive data for the entire sample of 17,220 offenders:

- **Outcomes**
  - Received GED During/After cohort
    - Yes = 4,974 (28.9%)
    - No = 12,246 (71.1%)
  - Composite TABE Gains
    - mean = 1.61 (grade levels), range = 0 – 11.5
  - Reading TABE Gains
    - mean = 1.43, range = 0 – 12.8
  - Math TABE Gains
    - mean = 1.68, range = 0 – 11.1
  - Writing TABE Gains
    - mean = 1.92, range = 0 – 12.3

- **Predictors**
  - Offender Type
    - ID (n = 13,063, 75.9%)
    - SJ (n = 3,430, 19.9%)
    - SAFP (n = 727, 4.2%)
  - Gender
    - Male (n = 15,386, 89.3%)
    - Female (n = 1,599, 9.3%)
    - Missing (n = 235, 1.4%)
  - Race
    - White (n = 3,054, 17.7%)
    - Black (n = 6,861, 39.8%)
- Hispanic (n = 7,239, 42.0%)
- Asian (n = 47, 0.3%)
- American Indian (n = 11, 0.1%)
- Other (n = 7, 0.0%)
- Unknown (n = 1, 0.0%)

- Age at end of (last) school year, mean = 20.5, median = 21.0, range = 15 – 22
- Sentence length, median = 1,461, range = 0 – life
- Exposure to academic programs
  - Total Title 1 Hours: n = 3,512; avg. = 171.4, median = 101.0
  - Total Academic Hours: n = 15,532; avg. = 215.4, median = 147.0
  - Total Computer Lab Hours: n = 5,375; avg. = 152.3, median = 105.0
  - Total ESL Hours: n = 88, avg. = 402.4, median = 326.0
  - Total SPED Hours: n = 219; avg. = 344.2, median = 221.0
  - Total Vocational Hours: n = 1,050; avg. = 303.1, median = 258.5
  - Total Cognitive Intervention Program (CIP) Hours: n = 2,535, avg. = 131.7, median = 159.0
  - Total CHANGES Program Hours: n = 4,766; avg. = 136.3, median = 170.0

- Initial TABE Scores
  - Composite: avg. = 6.4, median = 6.0
  - Reading: avg. = 7.3, median = 6.9
  - Writing: avg. = 5.9, median = 5.6
  - Math: avg. = 6.3, median = 5.8

- Highest TABE Scores
  - Composite: avg. = 8.0, median = 8.1
  - Reading: avg. = 8.8, median = 9.0
  - Writing: avg. = 7.8, median = 8.2
  - Math: avg. = 8.0, median = 7.8

- Gains (Grade-Level)
  - Composite: avg. = 1.61, median = 1.00
  - Reading: avg. = 1.43, median = 0.40
  - Writing: avg. = 1.92, median = 0.90
  - Math: avg. = 1.68, median = 1.10
ANALYTIC METHODS

CLASSIFICATION & REGRESSION TREES (“DECISION TREES”)

Classification tree analysis predicts values of an outcome variable from a number of predictor variables and offers a number of advantages over more commonly used statistical techniques (Horner, Fireman, & Wang, 2010). For instance, classification trees are both nonparametric and nonlinear. These features mean that missing data are not a problem and that typical assumptions regarding normality and linear relationship between variables is neither assumed nor necessary. In the present study, for example, a lack of normality and linear relations may mean that hours of services are related to academic outcomes only above a certain number of hours of service. In this initial exploratory stage—with many potential predictor variables and without specific a priori hypotheses about how the predictor variables are related to each other and to academic outcomes—this approach is potentially useful in discovering relations between variables that may have otherwise been missed.

The primary caveat of these types of classification and regression (or “decision”) trees is that the results may not completely generalize to a new sample. For example, with a new sample of offenders (or new predictors), variables may shift up or down in the tree structure. Readers might infer that the variable is more (or less) important, but these shifts may be due to characteristics specific to the sample of offenders being analyzed and the predictors being used.

RETURN ON INVESTMENT:

- To put the Title 1 hours and the TABE grade-level gains on the same metric, transformed the grade-level gains into the presumed instructional time in a public education setting so that we could compute a return on investment.
- We assumed a grade-level gain in a public education setting is analogous to 180 days of instruction with an average of 6 hours of instruction per day (1,080 hours of instruction per grade level gained).
There were five (5) outcomes of interest (receipt of GED, and grade-level gains in TABE Composite, Reading, Writing, and Math) and 13 predictors of interest (8 WSD programs, primarily Title 1 programs; gender, race, age at end of the last school year, inmate type, and sentence length).

Analyses were first conducted looking only at the effect of Title 1 hours on the 5 outcomes of interest, then the “demographic” variables of gender, race, age, inmate type, and sentence length were added as a second set of analyses, and finally the additional WSD programs (Academic, Vocational, ESL, SPED, Computer Lab, CIP, and CHANGES) were added to a final set of analyses.

Title 1 hours showed a strong correlation with grade-level gains on the TABE for all four scores (Composite, Reading, Writing, and Math); Title 1 hours were not significantly correlated with GED attainment.

**Title 1 Hours**

When looking at Title 1 hours alone on the three primary TABE measures of interest (Reading, Math, Composite), Title 1 hours were a significantly positive return on investment. The following table outlines the relationship between Title 1 hours and grade-level gains, along with the return on investment. [Please see the notes below the table.]
<table>
<thead>
<tr>
<th>TABE Content Area</th>
<th>Title 1 Hours</th>
<th>Avg. Grade-Level Gain (from decision tree analyses)</th>
<th>Return on Investment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reading</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>≤ 172</td>
<td>1.346</td>
<td><strong>8.5:1</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>172-484</td>
<td>2.422</td>
<td><strong>3.7:1</strong></td>
</tr>
<tr>
<td></td>
<td>(312 hours)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt; 484</td>
<td>3.219</td>
<td></td>
</tr>
<tr>
<td><strong>Math</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt; 128</td>
<td>1.549</td>
<td><strong>13.1:1</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>128-207</td>
<td>2.312</td>
<td><strong>10.4:1</strong></td>
</tr>
<tr>
<td></td>
<td>(79 hours)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>207-373</td>
<td>3.131</td>
<td><strong>5.3:1</strong></td>
</tr>
<tr>
<td></td>
<td>(166 hours)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt; 373</td>
<td>4.037</td>
<td></td>
</tr>
<tr>
<td><strong>Composite</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>≤ 172</td>
<td>1.506</td>
<td><strong>9.5:1</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>172-499</td>
<td>2.863</td>
<td><strong>4.5:1</strong></td>
</tr>
<tr>
<td></td>
<td>(327 hours)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt; 499</td>
<td>3.859</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The return on investment for second, and subsequent, ranges is the incremental (e.g., additional) return, above and beyond that from the prior range of hours. Because the maximum number of hours is used as the denominator in every hour range, the return on investment figure is the most conservative estimate. Return on investment is not calculated for the highest hour category because it is unbounded at the upper end.
Title 1 + Demographics

When gender, race, age, inmate type, and sentence length were added to Title 1 in the prediction of TABE gains, Title 1 hours were still the best predictor of Math TABE gains, the third best predictor of Reading and Writing and Composite TABE gains. See the decision tree graphs for more specific breakdowns of these relationships.

Title 1 + Demographics + Other WSD Programming

When other WSD programs were added to Title 1 and demographics, these other programs masked the impact of Title 1 on TABE outcomes. For example, Title 1 became the 2nd best predictor of Math (from 1st), the 4th best predictor of Composite (from 3rd), the 5th best predictor of Writing (from 3rd), and was not in the top 5 predictors of Reading (from 3rd). Nonetheless, even in these more complex trees, Title 1’s incremental benefits (e.g., above and beyond other forms of WSD programs) are apparent.

GED TREE HIGHLIGHTS:

- The overall theme of the GED tree is that a minimal amount of WSD programming can lead to substantial increases in the probability of GED attainment.
- A second theme of the GED tree is that there are synergistic combinations of WSD programming that lead to better outcomes.
- For example, minority students (Black, Hispanic, American Indian) with less than 30 Academic hours but between 28 and 70 Computer Lab hours move from a 1.5% rate to a 12.0% rate—an 8-fold increase.
- For minority students with minimal access to Computer Lab hours (< 70) and Academic hours (31-45), CHANGES improves GED attainment from 15.8% to 27.2%.
- Minority students with > 70 Computer Lab hours (but no CIP hours) and between 5-45 Academic hours increase almost 4-fold (10.2% to 41.0%) over those with < 4 Academic hours.
- For minority students, adding Vocational to at least 45 hours of Academics increases GED to 84%.

- Majority students (White and Asian) who are ID/SAFP inmates with minimal (< 42) Academic hours, if they have at least 21 Computer Lab hours there is a 5-fold increase in GED probability (from 6.8% to 35.7%).
- For majority students, giving at least 15 hours of Vocational increases GED probability to 75%.
- For majority state jail students with limited access to Academic hours (< 36), any Computer Lab hours improves GED from 2.9% to 21%.
Caveats:

As an outcome measure, grade-level gains may not be as salient as the highest score obtained. As an outcome measure, grade level “gains” are restricted at the top end of the scale, a phenomenon known as the “ceiling effect.” This is most evident graphically in the following scatterplot, which shows baseline scores on the horizontal (x) axis and highest scores on the vertical (y) axis. “Gains” are defined as vertical displacements from the diagonal. Those who started out with lower scores have more they can gain than those who start with higher scores. The results of the SB213 evaluation predicting post-release wage-earning and recidivism also suggest highest grade-level score is more important than grade-level gains in predicting wage-earning and recidivism.

Another caveat is that the relationship between Title 1 hours and grade-level gains begins to “flatten out” at about 500 total hours; in other words, there is a point of diminishing returns at about 500 hours. This is evident in the four scatterplots showing the relationship between Total Title 1 hours on the horizontal (x) axis and TABE grade-level gains on the vertical (y) axis.

Another caveat is that some of the “predictors” actually restrict the opportunity to receive the “treatment” of interest (Title 1), which confounds interpretation of the results. For example, offender type (ID, State Jail, SAFP) and sentence length were repeatedly important in predicting grade-level gains. Our interpretation of these findings is that sentence length affects the opportunity to participate in Title 1 hours—state jail offenders, particularly, seem to not have as many opportunities to participate in Title 1 hours, because of their short sentences.

Further evaluations should look more closely at these limitations identified.
Scatterplot of Initial Composite by Highest Composite
Reading Grade-Level Gains by Total Title 1 Hours
Math Grade–Level Gains by Total Title 1 Hours
Writing Grade-Level Gains by Total Title 1 Hours
Outcome is TABE Grade-Level Composite Gains
Predictors include Title 1 + Demographics

The overall average was 1.61

[Nodes highlighted in green are those with rates > 1.61]
Outcome is TABE Grade-Level Reading Gains
Predictors include Title 1 + Demographics

The overall average was 1.43
[Nodes highlighted in green are those with rates > 1.43]
Outcome is TABE Grade-Level Math Gains
Predictors include Title 1 + Demographics

The overall average was 1.68
[Nodes highlighted in green are those with rates > 1.68]
Outcome is TABE Grade-Level Writing Gains

Predictors include Title 1 + Demographics

The overall average was 1.92

[Nodes highlighted in green are those with rates > 1.92]
GED Received

Overall 28.9% received a GED.

[Nodes highlighted in green are those with rates > 28.9%]