Exploring the Reliability of Measurements of a Teacher Education Program Assessment System: A Pilot Study

Part 1 - Purpose
In the Spring 2014 the Unit’s Accreditation Coordinator and Data Manager decided to conduct an exploratory reliability study of the Teacher Education Assessment System’s by analyzing performance across various accreditation key assessments. At the beginning of this project, it was convened to have a pilot sample whose results would be utilized to guide the planning of a study with a larger sample size.

In March 2014 Datareq generated a database with 439 names and identification codes of candidates who graduated between the years 2012 and 2014. The random selection feature available on Excel was utilized to draw a sample of 50. This sample of 50 participants was moved to a different Excel spreadsheet where names were deleted but identification codes were preserved. For each entry (i.e., participants) data from the following variables were populated:

1. Gender
2. MTLE/Writing Score
3. MTLE/Reading Score
4. MTLE/Math Score
5. MTLE/Pedagogy Score
6. MTLE/Content Score
7. Licensure Area (i.e., Elementary, Secondary, Special Education, Early Childhood)
8. Cumulative Graduate Point Average at Graduation
9. SPED 205 Course Grade
10. SPED 225 Course Grade
11. ED 294 Course Grade
12. ACT Composite Score
13. ACT English Score
14. ACT Math Score
15. ACT Reading Score
16. ACT Science Reasoning Score

Of the 50 entries, 38 were kept due to the degree of available data. The Datareq Office was able to generate most of the aforementioned data. All MTLE scores were collected individually from the Pearson site by a Graduate Research Assistant.

Part 2 - Analyses
Analyses for this pilot study were executed in two phases. Phase I included descriptive statistics reporting for continuous and discrete data. Phase II included inferential statistics reporting for the same continuous and discrete variables.

Phase I
Data analyzed and summarized during Phase I are shown on Tables 1 and 2. Of all variables used in this pilot study the ones with the least data available were the MTLE Basic Skills scores (i.e., Writing, Reading, and Math). Some of our randomly selected candidates were either included in the Pearson database without any available scores for MTLE Basic Skills or candidates were not at all included in the Pearson database. For this reason, whenever results are associated to the MTLE Basic Skills caution in making interpretations should be exercised. All other variables of study (e.g., GPA, ACT composite score) were generated by MSUM Datareq.

Table 1
Descriptive Data – Continuous Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>MTLE Writing</th>
<th>MTLE Reading</th>
<th>MTLE Math</th>
<th>MTLE Pedagogy</th>
<th>MTLE Content</th>
<th>GPA</th>
<th>ACT Composite</th>
<th>ACT English</th>
<th>ACT Math</th>
<th>ACT Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>9</td>
<td>10</td>
<td>7</td>
<td>30</td>
<td>23</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>Mean</td>
<td>234.5</td>
<td>251.6</td>
<td>233.5</td>
<td>260.26</td>
<td>261</td>
<td>3.4</td>
<td>4.4</td>
<td>21.57</td>
<td>21.3</td>
<td>22.56</td>
</tr>
<tr>
<td>Median</td>
<td>234</td>
<td>256</td>
<td>240</td>
<td>259.5</td>
<td>260</td>
<td>3.4</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Mode</td>
<td>NA*</td>
<td>NA</td>
<td>NA</td>
<td>264</td>
<td>248</td>
<td>3.4</td>
<td>19</td>
<td>29</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>19.86</td>
<td>16.58</td>
<td>29.48</td>
<td>15.70</td>
<td>11.96</td>
<td>.32</td>
<td>4.94</td>
<td>4.20</td>
<td>5.17</td>
<td></td>
</tr>
</tbody>
</table>

*NA = Not Applicable

Table 2
Descriptive Data – Discrete Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Gender</th>
<th>License</th>
<th>Area</th>
<th>ED</th>
<th>SPED</th>
<th>ED</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male (21%)</td>
<td>EIE</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Female (79%)</td>
<td>EC</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Second. (11%)</td>
<td>F</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>NI* (20%)</td>
<td>No</td>
<td>Grade*</td>
<td>Grade*</td>
<td>Grade*</td>
<td></td>
</tr>
</tbody>
</table>

*NI = Not Identified

*No Grade = While a grade was given to the student no grade was generated by the system

Results

MTLE

Given that MTLE tests require a minimum score of 240 to pass, on average MSUM candidates encountered challenges in passing the MTLE Writing and Math Basic Skills tests ($Range_w = 193-259$, $Range_m = 175-270$). However, on average, MSUM candidates are capable of passing the MTLE Pedagogy and Content tests without problem ($Range_p = 234-295$, $Range_c = 238-280$).
ACT
For the year 2016-2017, one of the new Council for the Accreditation of Educator Preparation (CAEP) suggested criterion for admission requires the mean ACT scores of the selected pool to be in the top 50 percent. This means a minimum score of 20. Based on this criterion, MSUM candidates are currently fulfilling this requirement with an ACT Composite group mean score of 21.94 (Range = 17-31). For the year 2018-2019, however, the ACT group mean score should be raised to the top 40 percent of the distribution, or a minimum ACT Composite score of 22. By the year 2020, the Unit’s pool of admitted candidates’ will have to have a mean ACT Composite score of 24 (i.e., top 33 percent of the distribution). This progressive increase in ACT scores in the next few years will require careful design of admission strategies.

GPA
CAEP also requires that the pool of selected students has a mean score of 3.00. In this sample of candidates, their mean GPA score at graduation was 3.44. This GPA is reflected in the distribution of course grades presented on Table 2. These courses are critical as they contain MN BOT Standards of Effective Practice and thus constitute key pieces for admission into the Selective Admission and Retention in Teacher Education (SARTE) program. MSUM candidates are, to great extent, performing successfully in these courses.

Phase II
Phase II consisted of inferential statistic reporting for continuous and discrete data. All continuous variables were correlated to determine whether significant associations were found. Analyses of variance were run to determine whether factors such as gender, area of license, or course grades created a difference on overall GPA scores or MTLE scores. Finally, regression analyses were conducted for selected variables to identify the best predictor(s) of MTLE Pedagogy and MTLE Content tests, the last examinations in Teacher Education Program.

Correlations
Several correlations were identified among continuous variables. The most relevant ones, for the purpose of this report, are the variables connected to the ACT Composite score which is one of the admission pieces into SARTE. These variables were:

- MTLE Content, \( r(23) = .65, p < .001 \)
- MTLE Pedagogy, \( r(30) = .61, p < .001 \)
- GPA Graduation, \( r(39) = .59, p < .001 \)

These are all positive correlations, which means that candidates who obtain a high ACT Composite score will most likely obtain high MTLE Content, high MTLE Pedagogy, and high GPA scores. Likewise, candidates who obtain a low ACT Composite score will most likely obtain a low MTLE Content, low MTLE Pedagogy, and low GPA scores. Given that the ACT Composite score is one of the
criterion for admission into SARTE, these data supports the notion of carefully looking into ACT scores for admission.

These correlations are shown on Figure 1.

![Figure 1. Pearson Correlation Coefficients between ACT Composite Scores and other Continuous Variables](image-url)

**Figure 1. Pearson Correlation Coefficients between ACT Composite Scores and other Continuous Variables**

**Analysis of Variance**
ANOVA showed that gender and license area had an effect on MTLE Pedagogy scores. No other scores were found to be affected by these or any other discrete variables (e.g., course grades).

The effect of gender on MTLE Pedagogy scores was significant, $F(1, 28) = 6.201, p < .05$. Post hoc analyses using Bonferroni correction for repeated measures indicated that female candidates' MTLE Pedagogy scores were significantly higher ($m = 263.96$) than their male counterparts ($m = 248.14$).

The effect of license area on MTLE Pedagogy scores was also significant, $F(3, 26) = 3.230, p < .05$. Post hoc analyses using Bonferroni correction for repeated measures indicated that the MTLE Pedagogy scores of candidates in the Early Childhood program ($m = 280.5$) were significantly higher than candidates in the Elementary Inclusive Education program ($m = 256$), Secondary program ($m = 260.75$) or Not Identified program ($m = 264$).

Performance on course grades had no effect on MTLE Pedagogy or MTLE Content scores.

**Regression**
Data showed that the best predictor for both, MTLE Pedagogy and MTLE Content is the ACT Composite score:

ACT Composite scores significantly predicted MTLE Pedagogy scores, $b = 607, t(28) = 4.041, p < .001$. ACT Composite scores also explained a significant proportion of variance in MTLE Pedagogy scores, $R^2 = .368, F(1,28) = 16.330, p < .001$.

ACT Composite scores significantly predicted MTLE Content scores, $b = 645, t(21) = 3.869, p < .001$. ACT Composite scores also explained a significant proportion of variance in MTLE Content scores, $R^2 = .416, F(1,21) = 14.967, p < .001$.

ACT Composite score did not predict scores on MTLE Basic Skills tests.

**Part 3 – Follow-Up Studies**

This was the first attempt at exploring the existing relationships among the Teacher Education Assessment System’s key assessments. The data collected have provided some results that will influence decision making in regards to admission criteria.

Also, based on this pilot study the TEAS committee has decided to conduct a follow-up study having in mind the following modifications:

1. Select a larger sample (e.g., 100 recent graduates)
2. Invest more time in retrieving data from Pearson database. This will require phone calls to the company so that the researchers can ensure the generation of a complete set of MTLE scores, particularly for Basic Skills tests which were significantly limited in this pilot study.
3. Include additional key assessments (e.g., Cooperating Teacher Final Evaluation, Supervisor Survey) in the database and analyses
4. Include analysis of data ethnicity

**Conclusions**

There is significant statistical association between candidates’ performance data at the beginning phase of their program and at the end phase of their program. A follow-up study should explore the links between these two end phases with the middle phase’s Key Assessments. This exploration could help the Unit identify students in need of intervention to ensure successful completion of academic program and licensing requirements.