
Lindsey D. Basileo, PhD
Michael D. Toth

April 19, 2015
Table of Contents
INTRODUCTION ................................................................................................................................. 3
BACKGROUND ................................................................................................................................. 4
OVERVIEW OF THE TNTP REPORT ............................................................................................... 4
FINDING 1 & RECOMMENDATION 4 ............................................................................................... 5
FINDING 3 & RECOMMENDATIONS 1 & 3 ..................................................................................... 8
  Number of Elements Scored per Teacher ................................................................................... 8
  Number of Observations ............................................................................................................. 10
  Frequency of Elements Scored .................................................................................................. 11
    Inaccurate Conclusion Drawn From Data Findings ................................................................. 12
    Misinterpretation of the Marzano Framework ......................................................................... 13
CONCLUSION .................................................................................................................................. 20
List of Tables
Table 1. Number of Lessons TNTP Observed in Elementary, Middle, and High Schools .............6
Table 2. Total Number of Schools in PCPS 2014-15........................................................................7
Table 3. TNTP Survey Sample Frames and Response Rates...........................................................7
Table 4. TNTP Report: Percent of PCPS Teachers Scored per Element......................................11
Table 5. National Frequencies of Scored Elements Compared to PCPS.....................................13
Table 6. TNTP Report: Missed Opportunities in Marzano Criteria...............................................15
Table 7. Providing Rigorous Learning Goals and Performance Scales...........................................16
Table 8. Identifying Critical Content ...............................................................................................17
Table 9. Scale for Identifying Critical Content ...............................................................................17
Table 10. Tracking Student Progress ..............................................................................................18
Table 11. Reviewing Content .........................................................................................................19

List of Figures
Figure 1. TNTP’s Figure: Observed Elements Scored per Teacher .................................................9
INTRODUCTION

On April 8, 2015, the *Tampa Bay Times* published an article titled, “What does the TNTP report say about Pasco County schools?” The article outlined summary findings from a more detailed report submitted to Pasco County Public Schools (PCPS) by The New Teacher Project (TNTP) titled, “A Roadmap to Rigor: How Pasco County Schools Can Prepare Students for College and Career.”

Learning Sciences International (LSI) implements the Marzano Teacher Evaluation Model (MTEM) referred to in the TNTP report. The LSI research team has found several critical flaws and major oversights in the TNTP report and the supplementary PowerPoint. Additionally, there was a gross misinterpretation of the MTEM which we will address in this response. These inaccuracies were based on misleading findings and inaccurate policy implications. Consequently, the recommendations made by TNTP were not supported by the data findings presented in their report.

The unnamed authors of the TNTP report provide a highly imbalanced view of their data findings. The TNTP report failed to provide context to the state of Florida where current legislation requires termination track for teachers if they receive evaluation scores below “effective.” There is enormous pressure on principals to retain teachers in this current environment. While we agree that PCPS needs to work toward proper implementation of the standards, the TNTP report paints a generally negative portrait of PCPS. This is despite the fact that the report’s authors claim that their purpose was to “identify areas of strength as well as opportunities for improvement” (p. 3). The only promising practices outlined by TNTP are detailed in the third paragraph of the introduction. The tone and content of the report is overwhelmingly critical of PCPS, including the liberal use of negative callout quotes. Further, TNTP included only negative data findings (or data they perceived as negative). They failed to fully report all relevant details of the study design. Critical parts of their analysis were incomplete and missing. Had a balanced assessment of district practice and implementation been the true objective of the report, more positive or exemplary practices would certainly have been discussed. In many respects, PCPS is likely keeping pace with peer districts in Florida, as the new academic standards were only recently adopted and are being tested this year.

The most problematic issues in this report were related to the paper’s high-stakes policy recommendations. The data findings and inexpert data analysis provided in the report were not sufficient to support such high-stakes claims. The report is riddled with factual errors and misleading misstatements. Moreover, TNTP does not disclose that they are a vendor of competing services and would benefit from the report recommendations. While this conflict of interest does not necessarily disqualify TNTP from conducting objective research, a lack of scientific integrity is evident.

---

4“The majority of TNTP’s revenue comes from its work with clients on a fee-for-service basis.” [http://tntp.org/about-tntp/our-funding-model](http://tntp.org/about-tntp/our-funding-model) TNTP also works to “advance evaluations” in districts. [http://tntp.org/what-we-do/grow-all-teachers](http://tntp.org/what-we-do/grow-all-teachers)
throughout the report. Most of these issues, which we address in detail below, could have been clarified if all study findings and methodologies had been disclosed. TNTP’s inaccurate statements related to the MTEM could have been prevented had the authors conducted a thorough review of the research.

We feel that a thorough review of the data and findings is necessary to flush out any errors and overstatements made by TNTP. Of course, we also encourage practitioners to read the TNTP report in its entirety to draw their own conclusions.

BACKGROUND
In order to disclose any possible conflicts of interest, Dr. Robert J. Marzano is the Executive Director of Learning Sciences Marzano Center and Michael D. Toth is the CEO of LSI. Dr. Lindsey D. Basileo is a Senior Research Analyst at LSI with more than 10 years of experience conducting process and outcome evaluations to assess the effectiveness of interventions. The research department at LSI is committed to conducting high quality and objective research dedicated to producing evidence-based results, ultimately facilitating sound policy and best practices in education. Learning Sciences partners with Dr. Marzano and districts throughout the nation to implement the MTEM with fidelity.

Learning Sciences is the creator of the iObservation™ technology platform. iObservation™ is an instructional and leadership improvement system widely used in K-12 schools throughout the United States and abroad. iObservation™ was developed to support teachers and administrators to observe and document specific pedagogical and leadership strategies; to provide and receive targeted, timely feedback; and to enhance the observational process with a built-in library of professional development resources. LSI supports schools and educational initiatives in 11 countries, three provinces in Canada, and 40 states in the United States. We currently serve more than 420 school districts; 5,500 school buildings; 311,000 teachers; and 4 million students.

PCPS is also a customer of Learning Sciences and uses the iObservation™ platform. Under the previous superintendent, PCPS adopted the MTEM as part of the national Race to the Top Initiative and proceeded to implement the MTEM without evaluator training on the model from a certified source. It was only under new leadership that PCPS began to seek guidance from LSI to implement the MTEM with fidelity. In May 2013, PCPS observers began receiving a series of trainings on Domain 1 of the model. They have continued these trainings in the 2014-15 school year, making progress toward a deeper understanding of the model.

OVERVIEW OF THE TNTP REPORT
TNTP conducted an analysis which included the use of surveys, leader interviews, focus groups, and observations. Observation data were collected to assess the extent to which teachers, school leaders, and students incorporated standards-based instruction in PCPS (p. 3). The overarching aim of the site visits was to identify opportunities for improvement as PCPS worked to implement the new Florida standards. The main finding of the TNTP analysis was that “teachers and school leaders are not aligned on a common definition of instructional excellence and, thus, are not fully prepared to
implement the new Florida Standards (p. 3).” TNTP reports the following four major findings from their analysis:

1. Students are not consistently engaging with rigorous content.
2. Professional learning communities have the potential to help teachers improve their understanding of the Florida Standards and raise the rigor of their instruction.
3. The teacher evaluation system makes it difficult for school leaders to provide meaningful feedback that will raise the rigor of instruction in Pasco classrooms.
4. Pasco County Schools has not prioritized the hiring and retention of effective teachers, which is critical to sustaining its progress toward improved instruction.

TNTP also outlined “A Roadmap to Rigor: Solutions and Timeline,” which offers recommendations to achieve rigor. TNTP made the following recommendations:

1. Develop a unifying definition for instructional excellence in Pasco that shapes the district's overall strategy.
2. Develop a goals-based model for hiring and retaining great teaching talent, focused on opening the 2016-17 school year with no teacher vacancies to ensure sustained academic improvements.
3. Revisit the design and implementation of the teacher evaluation and development system to ensure that it is accurate and useful for supporting teacher development and for making high-stakes human capital decisions.
4. Revise central curricular resources to fully align with intent and demands of the standards.
5. Build a deep base of knowledge about the demands of the new Florida Standards, beginning with PLC facilitators, coaches, school leaders and select central office staff.
6. Improve PLC structures and content to ensure that they become opportunities for teachers to raise the rigor of the instruction and student work in their classrooms.

The aim of LSI’s response is not to address all errors in the TNTP report. However, we will outline several of their misleading findings and conclusions, discussing both TNTP’s findings and subsequent recommendations. First, we will address Finding 1 and Recommendation 4, followed by a discussion of Finding 3 and Recommendations 1 and 3. We will also include a discussion around the misinterpretation of the MTEM. In conclusion, our response will discuss the need for scientific integrity in educational research.

**FINDING 1 & RECOMMENDATION 4**

- Finding 1: Students are not consistently engaging with rigorous content.
- Recommendation 4: Revise central curricular resources to fully align with intent and demands of the standards.

The recommendation made by TNTP to revise central curricular resources is based on a small sample of observation data collected (about 2% of PCPS teachers). Providing recommendations based on a non-representative sample is scientifically problematic when making district-wide, resource-intensive recommendations.
Although we agree with TNTP’s general statement that districts should align curricular resources to academic standards, we disagree with the scientific method used to reach that conclusion. Finding 1 and Recommendation 4 were based on site visits at only six schools. The school types and exact number of teachers included in the analysis were not reported. However, TNTP did state that there were 100 lessons observed over four days, from January 12-15, 2015 (p. 4). Neither the schools chosen to be included in the study, nor the teachers observed, were randomly selected. TNTP states that the schools were selected “in partnership with the district leaders to represent demographic and performance variation across the district.” Furthermore, they state that the teachers observed were selected by the principal as “representative of the variation in instructional practices within the school” (p. 28). While the authors seem to imply that a representative sample was selected, they do not report any descriptive information about the schools or teachers. They also do not provide any further information or data in the report to support this recommendation.

Table 1. Number of Lessons TNTP Observed in Elementary, Middle, and High Schools

<table>
<thead>
<tr>
<th># of Lessons Observed</th>
<th>ES</th>
<th>MS</th>
<th>HS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELA</td>
<td>33</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Math</td>
<td>28</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>Science &amp; Social Studies</td>
<td>39</td>
<td>3</td>
<td>?</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>25</strong></td>
<td>≥29</td>
</tr>
</tbody>
</table>

Table 1 reports the number of lessons observed in the TNTP analysis—a total of 100. The majority of observations conducted were in Science and Social Studies (39), followed by ELA (33) and Math (28). When disaggregating the information by school type, there were very few lessons observed at the high school(s)—only 5 lessons were observed in ELA and Math each. It was not reported how many Science and Social Studies lessons were observed at the high school(s). While TNTP notes the small number of lessons observed at the high school(s), the authors only note this as a reason to combine middle and high school samples into one “secondary” category (p. 5, footnote 5). They do not state any limitations in regard to the generalizability of findings due to the small sample size. In addition, the number of lessons observed from middle and elementary schools was also very small. There is still a large discrepancy between the TNTP sample size and the number of teachers in PCPS, even when assuming that each lesson observed was taught by a different teacher. In 2014-15, there was a total of 5,006 teachers and 308 administrators in the district.⁵ Thus, the study sample includes less than 2% of teachers at PCPS, and less than 7% percent (6 out of 87) of schools in the district.

Table 2. Total Number of Schools in PCPS 2014-15

<table>
<thead>
<tr>
<th>School Type</th>
<th>Number of Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary</td>
<td>44</td>
</tr>
<tr>
<td>Middle</td>
<td>15</td>
</tr>
<tr>
<td>High</td>
<td>13</td>
</tr>
<tr>
<td>Educational Centers</td>
<td>4</td>
</tr>
<tr>
<td>Charter (K-8)</td>
<td>10</td>
</tr>
<tr>
<td>Virtual (K-12)</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>87</strong></td>
</tr>
</tbody>
</table>

Table 2 reports the number and types of schools in the 2014-15 school year in Pasco County. There were 44 elementary schools, 15 middle schools, and 13 high schools. TNTP visited only six of these schools over the four days in their analysis. Estimating that there were four elementary schools included in the TNTP analysis, one middle school, and one high school, that would equate to observations conducted in 9% of elementary schools, 6% of middle schools, and 8% of high schools across the district.

The main issue we address here is that TNTP does not disclose or describe how the study sample used compares to the district as a whole. Researchers typically provide such information to indicate how study findings can be generalized to the larger district. It is common scientific practice to disclose any possible sample bias inherent in findings because larger samples typically yield less sampling error. Apart from failing to disclose either the distribution of school type or the number of teachers included in the analysis, the data collected comprise a very small fraction of the full district. To justify high-stakes recommendations like the ones proposed by TNTP, recommendations which are meant to be applicable district-wide, full justification of findings and including all relevant data are in order. Furthermore, if TNTP is making recommendations based on findings not described in their report, those findings should also have been published.

Lastly, it should be noted that while response rates were high for the three surveys that were administered, there is no discussion as to why some schools and administrators were left out of survey administration. Table 3 of this report illustrates TNTP survey sample frames and response rates. Note that in the 2014-15 school year, PCPS was comprised of a total of 5,006 instructional employees and 308 administrators in 87 schools.\(^6\) TNTP should have disclosed why some schools, teachers, and administrators were not asked to participate in the surveys. Again, for the sake of full disclosure, such inconsistencies should be addressed. Moreover, while these large-scale surveys were conducted, only small sets of findings from the surveys were reported.

Table 3. TNTP Survey Sample Frames and Response Rates

<table>
<thead>
<tr>
<th>Survey</th>
<th>TNTP Number Completed</th>
<th>TNTP Respondent Sample Frame</th>
<th>TNTP School Sample Frame</th>
<th>TNTP Reported Response Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insight Culture Survey</td>
<td>3,287</td>
<td>4,446</td>
<td>81</td>
<td>74%</td>
</tr>
<tr>
<td>(Teacher Survey)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School Leader Survey</td>
<td>166</td>
<td>214</td>
<td>78</td>
<td>78%</td>
</tr>
<tr>
<td>Coaching Survey</td>
<td>133</td>
<td>163</td>
<td>74</td>
<td>82%</td>
</tr>
</tbody>
</table>

\(^6\) District School Board of Pasco County Fact Sheet. (2014-15).
FINDING 3 & RECOMMENDATIONS 1 & 3

- Finding 3: The teacher evaluation system makes it difficult for school leaders to provide meaningful feedback that will raise the rigor of instruction in Pasco classrooms.
- Recommendation 1: Develop a unifying definition for instructional excellence in Pasco that shapes the district’s overall strategy.
- Recommendation 3: Revisit the design and implementation of the teacher evaluation and development system to ensure that it is accurate and useful for supporting teacher development and for making high-stakes human capital decisions.

While the exact wording of Recommendation 3 does not explicitly state that the evaluation framework should be changed, a closer read of the report overtly does. TNTP authors make the following recommendation:

```
Develop a plan for transitioning to a better-aligned framework by the 2016-17 school year. The high-level analysis of the Marzano framework included in this report (and expanded on in Appendix IV) shows evidence that Pasco’s observation rubric does not identify the most crucial aspects of instructional practice in light of the new standards, including content alignment. Based on this analysis, the Marzano framework is not likely to be the best tool for setting clear expectations and for driving teacher observations in Pasco. Transitioning to a new observation framework is a complex process that often includes rapid decision making, piloting and testing of a new tool, intensive training, significant change management and communications, and labor relations. In order to change responsibly from Marzano to a better aligned tool by the 2016-17 school year, Pasco should immediately begin developing a transition plan. (p. 23)
```

TNTP’s recommendation to abandon the evaluation model was based on two descriptive findings. First, TNTP presented the number of scored elements per teacher. Second, TNTP outlined the frequency of elements scored. Both are fundamentally issues of implementation of the MTEM, as we discuss below.

**Number of Elements Scored per Teacher**

TNTP represents the same data inconsistently in their final report (pp. 12-13, footnote 9) and in Figure 1 taken from their PowerPoint (p. 24). The authors provide two different descriptions of where the data was obtained. The written report states,

```
These data come from the 2013-14 classroom observation ratings based on the Marzano framework that were used, in part, to calculate Pasco teachers' summative evaluation
```
ratings. This dataset includes 3,979 teachers, with over 50,000 total elements scored across the 41 elements in Marzano’s Domain 1. Conversations with central office staff confirmed that a small number of teachers had more than one observation recorded in the data, which we were unable to disaggregate. Those teachers with more than 41 elements rated were likely those with more than one observation. (pp. 12-13)

In TNTP’s PowerPoint, the figure reproduced below stated that the information was obtained “from Pasco’s 2013-14 evaluation data. Teachers with more than 41 are teachers who received more than 1 observation and were scored on a high number of elements.” (p. 24)

As noted earlier in this response, TNTP observed 100 lessons across six schools. Their study sample did not include 3,979 teachers in that analysis. Clearly the data cited above did not come from observed schools. Rather, TNTP obtained this data from an export from iObservation™ without reporting this in their methodology section as a secondary data source used in their analysis.7

TNTP ought to have reported that the dataset illustrated above includes both evaluative and non-evaluative scores. PCPS is running dual systems for feedback to teachers, one used for evaluation and the other for growth. LSI has been working with PCPS to correct these issues in their implementation of the MTEM. For a full accounting of the number and frequency of strategies used

---

7It should be noted that LSI creates extensive observation reports for districts (PCPS included) and presents an observation analysis of district and school-level data. The reports include which elements were scored more often, the score distribution, the number of observations received by teachers, the scored element count, the average teacher score, the element distribution by design question, a score distribution for each building and observer, and the frequency with which feedback in the form of comments was given to teachers with high and low scores. The LSI reports include recommendations for districts to improve their data and evaluation systems.
in PCPS classrooms and the feedback given to teachers, non-evaluative data are just as critical as evaluative scores. To fully understand how observers are scoring teachers using the MTEM in PCPS, it is crucial to know what data are being reported and why. Typically, districts report only evaluative scores.

More importantly, TNTP’s interpretation of the data is inaccurately represented. TNTP states, “Only 6 elements are scored for 50% or more of teachers.” However, Figure 1 shows that the majority of teachers are scored on 11 to 15 elements, not six. As seen in the figure, the bar representing 11 to 15 elements is larger than the bar representing six to 10 elements. Moreover, while they state six elements in the PowerPoint, they state 10 elements in their report (p. 34). Another error is reported in the horizontal axis of the figure, where it shows the number of elements scored as 1-5, 6-10, 11-15, 16-20, 21-25 and then 16-30 (rather than 26-30). These many inconsistencies add up to a marked carelessness in data analysis and multiple errors throughout the report.8

TNTP uses the data cited above to draw the following conclusion, a conclusion which is in fact based on a faulty understanding of the research base and practical application of the MTEM, “Because school leaders choose which of the 41 elements to rate, the observation tool is inconsistently implemented across the district.” (p. 24)

The data illustrated above does not provide any justification that the model is being inconsistently implemented across the district. To prove such “inconsistency,” a much deeper analysis is called for, one which would include disaggregating the data by school type and by investigating scoring patterns of observers. Moreover, strategies that teachers use in elementary, middle, and high school may vary simply because the learning needs of age groups change; frequency of observed strategies could further vary based on the needs of student populations in individual schools.9 Strategies teachers use rightly vary by the type of lesson being taught: lessons in which students interact with new knowledge; practice and deepen their understanding of new knowledge; or apply knowledge by generating and testing hypotheses.10 Moreover, observers should not expect every lesson to include the same strategies.11 Such an expectation would, in fact, constitute a fundamental error.

Number of Observations
TNTP’s second justification for abandoning the MTEM is based on the fact that the majority of teachers in PCPS received only one observation. Our analysis of the evaluative and non-evaluative observation data show that 59% of teachers in PCPS received one observation, 30%

---

8See also a reference to “Appendix C” in footnote 11 page 13. No “Appendix C” is included in TNTP’s report. Also, see discussion around PCPS facing challenges in keeping its strongest teachers. The attrition rates of teachers working in schools in the bottom 25% of “instructional culture” are still lower than the national average they cite (20% < 25%) (pp. 18-19).


11Marzano & Toth, 2013.
of teachers received two, 8% of teachers received three, and 3% of teachers received four or more observations. While 59% of teachers only received one observation, this finding shows that the MTEM is not being implemented using best practices in PCPS as more observations are recommended—not that the framework is ineffective.

In *Teacher Evaluation that Makes a Difference*, Marzano and Toth (2013) discuss the reason teachers should be observed multiple times. First, multiple observations decrease sampling error in measurement (p. 61). But even with as many as 10 observations per teacher (which is rarely practical), there is an approximately 40% chance that all three types of lessons (introducing new content, practicing and deepening, and applying knowledge) would be observed. To combat issues in sampling error, Marzano and Toth offer five recommendations to minimize the amount of sampling error.

- Start with teacher self-evaluation.
- Use announced observations for specific types of lesson segments.
- Use video recordings of the three planned observations.
- Use data from brief walkthroughs to augment other observation scores.
- Ask teachers to provide video evidence or artifacts for specific strategies (pp. 61-64).

### Frequency of Elements Scored

The next set of data that TNTP references in support of replacing the MTEM comes from the frequency of scores that pertain to elements in the model that relate to the Florida standards. The following excerpts are taken from Appendix IV of the TNTP report:

In its 2014 teacher evaluation protocol, Marzano suggests that teachers and observers emphasize seven specific elements from the Content design questions that most closely align to the Common Core. These seven elements are identified in the first column of the table below. (p. 34)

<table>
<thead>
<tr>
<th>Marzano element</th>
<th>% of Pasco teachers scored</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifying critical information (element 6)</td>
<td>56%</td>
</tr>
<tr>
<td>Elaborating on new information (element 11)</td>
<td>24%</td>
</tr>
<tr>
<td>Recording and representing knowledge (element 12)</td>
<td>43%</td>
</tr>
<tr>
<td>Examining similarities and differences (element 17)</td>
<td>21%</td>
</tr>
<tr>
<td>Examining errors in reasoning (element 18)</td>
<td>14%</td>
</tr>
<tr>
<td>Revising knowledge (element 20)</td>
<td>12%</td>
</tr>
<tr>
<td>Engaging students in cognitively complex tasks (element 22)</td>
<td>7%</td>
</tr>
</tbody>
</table>
TNTP continues:

We found that Pasco observers typically only rate a teacher on about 10 elements, which is not surprising given that most observations last less than 30 minutes. If observers are narrowing their scope out of necessity, it makes sense that they should do so consistently across the district, and that they focus on the highest leverage components to ensure students are mastering standards-aligned content that will prepare them for college and career. (p. 34)

There are several issues noted with TNTP’s interpretation. First, this finding has more to do with implementation of the model than with intrinsic deficiencies in the model as related to standards-based learning. Further, a comparison of PCPS data to a national dataset reveals that PCPS is scoring elements that are related to standards-aligned strategies at a greater frequency than a national sample of organizations using the same model. This will be discussed in more detail in the next section.

Inaccurate Conclusion Drawn From Data Findings

TNTP used the same data export from iObservation™ as noted in “Number of Elements Scored Per Teacher” section above to create the frequencies shown in Table 4. Using that same dataset with a different download date, LSI replicated the TNTP analysis and came to the same data findings reported in Table 4. However, although these findings are accurate, TNTP misinterprets what the data actually mean. The data that TNTP presents actually shows promise for PCPS. Considering that PCPS averaged only one observation score per teacher, the frequency of these essential strategies for achieving rigor indicate positive movement toward standards-based learning. If PCPS required an increase in the number of observations per teacher, administrators could most likely provide teachers with enhanced feedback in these elements.

In our own analysis of more than 2.1 million data points collected from more than 260 organizations reporting data in iObservation™, we found that, across the board, districts are struggling to provide feedback on the strategies associated with more rigorous learning.12 Table 5 shows the national distribution of scores for the 2012-13 school year in addition to the frequencies which PCPS scored these elements in 2013-14. The percentages are calculated by taking the number of teachers scored for each strategy divided by the total number of scored elements in Design Questions 2, 3, and 4. These design questions are where rigorous content strategies are concentrated in the MTEM that lead to more student-generated, complex thinking.

One of the major findings from Marzano and Toth (2014)13 was that instruction focused on achieving rigor is rare and not frequently observed in classrooms. The lack of such instruction amounts to a crisis if we expect students to meet standards. Even if classrooms are randomly sampled, one would not expect to see rigor 100% of the time. When comparing how PCPS

---

13 Marzano & Toth (2014).
compares to national frequencies of elements scored, it seems that PCPS teachers are actually using these strategies more often than teachers in our national sample. Those strategies in bold, below, are those aligned to Florida standards. When looking at strategies aligned to standards, only two elements are scored less frequently than the national sample, Identifying Critical Information and Elaborating on New Information. While this is an overall positive finding, it does not discount the fact that PCPS should continue to focus on rigor in the classroom with student evidence. LSI supports the shift to rigorous learning called for by the new Florida standards.

Table 5. National Frequencies of Scored Elements Compared to PCPS

<table>
<thead>
<tr>
<th></th>
<th>National Frequencies 2012-13</th>
<th>PCPS Frequencies of Evaluative Scores Only 2013-14</th>
<th>PCPS Frequencies of Evaluative &amp; Non-Evaluative Scores 2013-14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifying Critical Information</td>
<td>93,560 12.5%</td>
<td>2,108 10.3%</td>
<td>2,706 10.6%</td>
</tr>
<tr>
<td>Practicing Skills, Strategies, and Processes</td>
<td>89,847 12.0%</td>
<td>1,773 8.7%</td>
<td>2,318 9.1%</td>
</tr>
<tr>
<td>Chunking Content into &quot;Digestible Bites&quot;</td>
<td>88,655 11.8%</td>
<td>1,793 8.8%</td>
<td>2,258 8.9%</td>
</tr>
<tr>
<td>Reviewing Content</td>
<td>80,054 10.7%</td>
<td>2,183 10.7%</td>
<td>2,818 11.1%</td>
</tr>
<tr>
<td>Organizing Students to Interact With New Knowledge</td>
<td>53,907 7.2%</td>
<td>1,630 8.0%</td>
<td>2,029 8.0%</td>
</tr>
<tr>
<td>Previewing New Content</td>
<td>53,491 7.1%</td>
<td>1,485 7.3%</td>
<td>1,864 7.3%</td>
</tr>
<tr>
<td>Organizing Students to Practice and Deepen Knowledge</td>
<td>51,110 6.8%</td>
<td>1,677 8.2%</td>
<td>2,071 8.1%</td>
</tr>
<tr>
<td>Recording and Representing Knowledge</td>
<td>48,154 6.4%</td>
<td>1,568 7.7%</td>
<td>2,023 7.9%</td>
</tr>
<tr>
<td>Processing New Information</td>
<td>38,391 5.1%</td>
<td>1,519 7.4%</td>
<td>1,899 7.5%</td>
</tr>
<tr>
<td>Elaborating on New Information</td>
<td>36,166 4.8%</td>
<td>884 4.3%</td>
<td>1,089 4.3%</td>
</tr>
<tr>
<td>Providing Resources and Guidance</td>
<td>27,879 3.7%</td>
<td>270 1.3%</td>
<td>317 1.2%</td>
</tr>
<tr>
<td>Examining Similarities and Differences</td>
<td>20,043 2.7%</td>
<td>789 3.9%</td>
<td>933 3.7%</td>
</tr>
<tr>
<td>Reflecting on Learning</td>
<td>19,559 2.6%</td>
<td>971 4.8%</td>
<td>1,129 4.4%</td>
</tr>
<tr>
<td>Examining Errors in Reasoning</td>
<td>14,549 1.9%</td>
<td>534 2.6%</td>
<td>621 2.4%</td>
</tr>
<tr>
<td>Using Homework</td>
<td>9,184 1.2%</td>
<td>288 1.4%</td>
<td>353 1.4%</td>
</tr>
<tr>
<td>Engaging Students in Cognitively Complex Tasks</td>
<td>9,087 1.2%</td>
<td>253 1.2%</td>
<td>282 1.1%</td>
</tr>
<tr>
<td>Revising Knowledge</td>
<td>7,957 1.1%</td>
<td>450 2.2%</td>
<td>530 2.1%</td>
</tr>
<tr>
<td>Organizing Students for Cognitively Complex Tasks</td>
<td>6,679 0.9%</td>
<td>221 1.1%</td>
<td>241 0.9%</td>
</tr>
</tbody>
</table>

In conclusion, the two data elements presented by TNTP simply point to implementation issues that may be addressed by PCPS administration as they move toward a deeper implementation of the model. The next section will further demonstrate misconceptions TNTP presented in the report around the implementation of the MTEM.

**Misinterpretation of the Marzano Framework**

There were several areas where TNTP misinterpreted the MTEM. First, in order to address Recommendation 1, it should be noted that PCPS already has a definition of instructional excellence. The MTEM is grounded in a model of instruction and common language. Our reply in this section will detail the importance of the use of research-based instructional strategies within a robust framework that requires student evidences at the “Applying” and “Innovating” levels. It also may be necessary for PCPS to more intentionally use the model of instruction to shape common
language in the district. Confusion around the definition of instructional excellence is typically seen when districts are not implementing the model with fidelity. Thus, there is not a need for a “unifying definition” but rather a need for more frequent use of the common language. Second, in order to address TNTP’s basis for making Recommendation 3, the misinterpretation of Identifying Critical Information and Reviewing Content and TNTP’s apparent misunderstanding of the use of Providing Rigorous Learning Goals and Scales will be discussed. The omission of Providing Rigorous Learning Goals and Scales is a critical error when discussing Critical Content because these two strategies, when utilized correctly, form the basis of the standards alignment in the framework.

The TNTP interpretation of the Marzano Framework is cited below. The following excerpts are taken from Appendix IV of the TNTP report.

To better expand on the analysis of Pasco’s use of the Marzano framework, this appendix addresses additional challenges faced when using this tool for the purpose of teacher evaluation and development. This diagnostic found that the Marzano framework does not adequately assess the quality or grade-level appropriateness of taught content or whether students have achieved the desired learning. Furthermore, Pasco’s implementation of Marzano has made it difficult for school leaders to focus on elements that are most closely aligned with the standards. The Marzano Framework prioritizes the assessment of teacher actions almost to the exclusion of assessing the rigor of content and impact on student learning. A teacher may have mastered a robust toolkit of teaching strategies, but that does not guarantee that students are learning grade-level, standards-aligned content that will prepare them for college and career. Nearly all of the Marzano Framework focuses on what the teacher does to prepare or deliver a lesson and ignores two fundamental questions that observers should ask when assessing a lesson: Did the lesson focus on content that advances students toward grade-level standards? In addition, does student work and effort demonstrate that they are on track to achieve the standard or learning goal?

Within Domain 1, the elements in Design Questions 2-4 are intended to assess the content of the lesson, but they tend to do so ineffectively. In Pasco, teachers are most frequently observed on two elements within the content focused design questions: Element 6 - Identifying Critical Information (56 percent) and Element 14 – Reviewing Content (58 percent). The table below illustrates how these two elements miss the mark when it comes to ensuring that students are learning standards-aligned content.
Like the vast majority of Domain 1 elements, those focusing on content zero-in on teacher actions and instructional techniques without demanding evidence of student learning. Although Marzano has made some revisions to accommodate Common Core State Standards, Pasco’s implementation remains a barrier to using the tool for teacher development. The Marzano Framework was originally designed as a lesson-planning framework many years ago. In the interim, the framework has been adapted to become a teacher evaluation tool and is now used by many states and districts. Although the framework has not been fundamentally reworked to support standards-aligned instruction, its creators have released guidance on how to adapt the framework for the new standards. (pp. 33-34)

The TNTP report incorrectly identifies the Marzano Framework as “originally designed as a lesson-planning framework.” In fact, only one of the 10 design questions from the original The Art and Science of Teaching (2007) focuses on planning.14 The Art and Science of Teaching represented a confluence from a number of previous works on classroom instruction and management, including Classroom Instruction That Works15 and a Handbook for Classroom Management That Works,16 which focused on understanding the characteristics of effective teachers within schools. In the next section, we will review the elements that were misinterpreted in the TNTP report to further show how the strategies can be used as a basis of standards alignment within the Marzano framework.

### Identifying Critical Content

In Table 6 of this report (Table 5 in TNTP report), TNTP authors identify the following question as a missed opportunity for the element Identifying Critical Content, “Was the lesson focused on critical content as articulated by the standards?” (p. 33). First, the question suggests TNTP does not understand how Identifying Critical Content is directly related to the element Providing Rigorous Learning

---

**Goals and Performance Scales.** This element requires instruction to be aligned with a standards-based scale, with a progression of learning targets moving from less complex to more complex. As students demonstrate mastery of the learning targets through evidence of their work, they move up the standards-based scale as they track their progress to attaining those standards.

Table 7. Providing Rigorous Learning Goals and Performance Scales

<table>
<thead>
<tr>
<th>1. Providing Rigorous Learning Goals and Performance Scales (Rubrics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The teacher provides rigorous learning goals and/or targets, both of which are embedded in a performance scale that includes application of knowledge.</td>
</tr>
</tbody>
</table>

**Example Teacher Evidence**
- Teacher has a learning goal and/or target posted for student reference
- The learning goal or target clearly identifies knowledge or processes aligned to the rigor of required standards
- Teacher makes reference to the learning goal or target throughout the lesson
- Teacher has a scale that builds a progression of knowledge from simple to complex
- Teacher relates classroom activities to the scale throughout the lesson
- Teacher has goals or targets at the appropriate level of rigor

**Example Student Evidence**
- Students can explain the learning goal or target for the lesson
- Students can explain how their current activities relate to the learning goal or target
- Students can explain the levels of performance, from simple to complex, in the scale
- Student artifacts demonstrate students know the learning goal or target
- Student artifacts demonstrate students can identify a progression of knowledge

Table 7 above shows a table taken from the 2014 Marzano Center Teacher Observation Protocol. Each element within the protocol includes a focus statement that has a description for teacher and student evidences that should be observed. In the example evidences shown, there is a clear progression of knowledge aligned to the rigor of the standards in the scale and the lesson should focus on this progression. Students should be able to articulate how the lesson relates to these targets in the scale. This facilitates the actions in the element Tracking Student Progress, in which students track their progress on the scale to demonstrate mastery of the standard (addressed in Table 9). Therefore, it is evident that the element Providing Rigorous Learning Goals and Performance Scales not only aligns to standards but that the required student evidence is also aligned to standards.

Secondly, the question posed by TNTP also appears to be a misunderstanding of Identifying Critical Content. The MTEM’s definition of Identifying Critical Content does, in fact, include alignment to standards.

---

Table 8. Identifying Critical Content

<table>
<thead>
<tr>
<th>6. Identifying Critical Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>The teacher continuously identifies accurate critical content during a lesson or part of a lesson that portrays a clear progression of information that leads to deeper understanding of the content.</td>
</tr>
</tbody>
</table>

Example Teacher Evidence
- Teacher highlights critical content that portrays a clear progression of information related to standards or goals
- Teacher identifies differences between the critical and non-critical content
- Teacher continuously calls students’ attention to accurate critical content
- Teacher integrates cross-curricular connections to critical content

Example Student Evidence
- Students can describe the level of importance of the critical content addressed in class
- Students can identify the critical content addressed in class
- Students can explain the difference between critical and non-critical content
- Formative data show students attend to the critical content (e.g., questioning, artifacts)
- Students can explain the progression of critical content

Table 8 shows that the focus statement for the element *Identifying Critical Content* is that, “The teacher continuously identifies accurate critical content during a lesson or part of a lesson that portrays a clear progression of information that leads to deeper understanding of content.” Furthermore, Senn and colleagues (2014) expound upon what *Identifying Critical Content* is and how the definition of it actually aligns with the standards, stating that, “Critical content is defined as the knowledge and skills necessary for students to demonstrate standards.”

A thorough understanding of the definition of *Identifying Critical Content* clarifies that the focus is on the standards. The protocol additionally cites teacher evidence in the first bullet in Table 8, “Teacher highlights critical content that portrays a clear progression of information related to standards or goals.” Moreover, evidence of student learning is at the very core of the MTEM and is embedded in the scale.

Table 9. Scale for Identifying Critical Content

<table>
<thead>
<tr>
<th>Identifying critical content</th>
<th>Not Using</th>
<th>Beginning</th>
<th>Developing</th>
<th>Applying</th>
<th>Innovating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategy was called for but not exhibited.</td>
<td>Uses strategy incorrectly or with parts missing.</td>
<td>Signals to students critical versus non-critical content and portrays a clear progression of information, but the majority of students are either not monitored for or not displaying the desired effect of the strategy.</td>
<td>Signals to students critical versus non-critical content and portrays a clear progression of information and monitors for evidence of the extent to which the majority of students are attending to critical versus non-critical content.</td>
<td>Adapts and creates new strategies for unique student needs and situations in order for the desired effect to be evident in all students.</td>
<td></td>
</tr>
</tbody>
</table>

As noted in Table 9, the elements that focus on content instruction require evidence of student learning at the “Applying” and “Innovating” levels of implementation. Teachers should also be aware of the extent to which the majority or all students understand and have a deeper understanding of the content. Because implementing effective instruction is a skill, the MTEM scale is designed to assess stages of skill development. Focusing on the stages of implementation is

---

necessary in order to provide specific feedback to teachers, especially when the desired student learning outcome is not obtained.

**Reviewing Content**

In Table 6 of this report, TNTP authors identify a “missed opportunity” for the element *Reviewing Content* and ask the questions, “Was the content accurately aligned to learning goals and was it accurately sequenced?” and “Do students appropriately apply previously learned skills to help them master new ones?” The first question posed by TNTP actually pertains to a different element in the model, *Tracking Student Progress*, which was not referenced by TNTP. It is *Tracking Student Progress*, not *Reviewing Content*, that focuses on whether both teacher and student are able to track student progress toward learning goals and targets. Learning goals and targets (if scaffolded correctly) relate directly back to standards.

<table>
<thead>
<tr>
<th>2. Tracking Student Progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>The teacher facilitates tracking of student progress on one or more learning goals and/or targets using a formative approach to assessment.</td>
</tr>
</tbody>
</table>

**Example Teacher Evidence**

- Teacher helps students track their individual progress on the learning goal or target
- Teacher uses formal and informal means to assign scores to students on the scale or rubric depicting student status on the learning goal
- Teacher uses formative data to chart progress of individual and entire class progress on the learning goal

**Example Student Evidence**

- Students can describe their status relative to the learning goal using the scale or rubric
- Students systematically update their status on the learning goal
- Students take some responsibility for providing evidence in reference to their progress on the scale
- Artifacts and data support that students are making progress toward a learning goal

Table 10 shows that the focus statement for *Tracking Student Progress* is that, “The teacher facilitates tracking of student progress on one or more learning goals and/or targets using a formative approach to assessment” (Marzano 2014 Protocol, Appendix B, p. 26). To reach the “Applying” level of *Tracking Student Progress*, the teacher should facilitate tracking of student progress toward learning goals and targets using a formative approach to assessment and by monitoring the extent to which students understand their level of performance. On the other hand, *Reviewing Content* speaks to the cumulative nature of the content.
**Table 11. Reviewing Content**

<table>
<thead>
<tr>
<th>The teacher engages students in a brief review of content that highlights the cumulative nature of the content.</th>
</tr>
</thead>
</table>

**Example Teacher Evidence**
- Teacher begins the lesson with a brief review of content
- Teacher systematically emphasizes the cumulative nature of the content
- Teacher uses specific strategies to help students identify basic relationships between ideas and consciously analyze how one idea relates to another
  - Summary
  - Problem that must be solved using previous information
  - Questions that require a review of content
  - Demonstration
  - Brief practice test or exercise
  - Warm-up activity

**Example Student Evidence**
- Students identify basic relationships between current and prior ideas and consciously analyze how one idea relates to another
- Students can articulate the cumulative nature of the content
- Student responses to class activities indicate that they recall previous content
  - Artifacts
  - Pretests
  - Warm-up activities

Table 11 shows that the focus statement for Reviewing Content is that, “The teacher engages students in a brief review of content that highlights the cumulative nature of the content” (2014 Protocol, Appendix B, p. 15). As for the “missed opportunity” to address whether students appropriately apply previously learned skills to help master new ones, the MTEM addresses this in another element. Depending on how the previously learned skill is being used in the lesson, it would be more appropriate to use either Previewing New Content (2014 Protocol, Appendix B, p. 9) or Helping Students Practice Skills, Strategies, and Processes (2014 Protocol, Appendix B, p. 20).

In summary, it is not reasonable to expect one or two elements of an evaluation model to focus on all aspects of teaching. There are multiple factors that can impact the probability of raising student achievement with the use of research-based instructional strategies. As already noted, not all strategies are appropriate for all types of lessons. Moreover, the MTEM takes into account that certain strategies are more effective within different types of lessons. The issue should not be whether Identifying Critical Content and Reviewing Content address student demonstration of progress toward mastery of the standard. Rather the issue should be whether there is a place in the MTEM for an evaluator to assess these important student actions, and the answer to that question is “Yes.”

Teaching is complex. There are multiple aspects to an effective lesson and effective student-teacher interactions. Therefore, an effective teacher evaluation tool should include all of these aspects so that evaluators may assess and provide feedback on specific areas of growth for the teacher. A robust evaluation model recognizes the complexity of teaching by allowing teachers to utilize different types of lessons and strategies for different purposes.
The data and arguments used by TNTP to support changing the evaluation framework in PCPS actually showed evidence that the MTEM is not being implemented with fidelity. It was rather reckless for TNTP to reach this conclusion, when they themselves recognize the importance of proper implementation of evaluation frameworks:

Again and again over the past few years, we have seen states and districts design better evaluation systems on paper that produce the same bad results in practice: wildly inflated ratings and little useful feedback to teachers. By contrast, districts that have devoted more time and political capital to implementation have begun to produce more meaningful evaluation ratings.\textsuperscript{19}

In districts where the MTEM was implemented with fidelity, such as in a pilot in Pinellas County, Florida, professional development for teachers and leaders with additional coaching sessions led to increased student achievement using a statistically matched control group.\textsuperscript{20} We found there were significant and moderate correlations with teacher observation ratings and value-added metrics (both unit and Florida State VAMs). The magnitude of significant validity coefficients ranged from .239 to .532, and these were slightly higher than those found in similar studies.\textsuperscript{21} TNTP’s suggestion to replace the Marzano framework should be accompanied by evidence of at least equal validity coefficients where state VAMs are the referenced criterion. Moreover, using Hierarchical Linear Modeling we found that students who attended treatment schools had significantly increased growth scores (.37 to .39 standard deviations above prediction) compared to students at control schools, accounting for individual and school characteristics. Relative to its utility in terms of enhancing teacher practice and student learning, the MTEM is one of the few teacher evaluation models that demonstrates that student achievement is enhanced by training teachers in its use.

Again, any recommendation of replacement models should be accompanied by similar evidence that student achievement is positively influenced by their use.

CONCLUSION

Scientific integrity should be the number one priority for educational researchers. In high-stakes situations where research studies can impact the lives of students, teachers, and administrators, studies like these could potentially have an impact nationwide. In environments where students are over-tested and teachers are underpaid, acting without scientific integrity can prove more harmful than good. It is negligent and unethical to only partially report data related to important findings and to fail to thoroughly outline methodologies and data sources.


\textsuperscript{20}Basileo, L., Toth, M., & Kennedy, E. (in progress). \textit{Final report: Pinellas County Public Schools, 2013-14 Results from a multiple measures pilot.}


Recommendations based on incomplete reports can cause harm in such high-stakes situations. Educational researchers should hold themselves to the highest standards—making sure that data reported are accurate, without flaws, and able to be replicated. Scientific researchers are under an obligation to provide readers with sufficient information to make informed decisions. These are ethical obligations scientists share with their colleagues in the scientific community, even when findings may contradict the ultimate aim of the organization. Researchers should operate under circumstances where it is their duty to report negative findings, even when doing so does not benefit the organization responsible for implementing the study.

Authors of research studies are most familiar with their study’s technical limitations. As such, all study limitations should be fully disclosed. Ultimately, we all have the same goal—to provide students with the best education possible. Operating with these ethical obligations in mind will only help us better achieve this aim.