Author: Anne Walton Garrison

Key Findings

- Among Shelby County Schools (SCS) teachers in their first three years of teaching, those who trained with Memphis Teacher Residency (MTR) outperformed their non-MTR counterparts on every measure analyzed in this study:
  - TVAAS (Tennessee Value-Added Assessment System) growth scores
  - TEM (Teacher Effectiveness Measure) observation scores
  - TEM professionalism scores
  - TEM student-perception scores
- Among SCS teachers with four to seven years of teaching experience, MTR-trained teachers outperformed their non-MTR counterparts on TEM professionalism. However, there was no statistically significant difference between MTR and non-MTR teachers on the three other measures.
- The SCS teachers who mentored MTR residents outperformed the rest of the district on all four measures. As a group, they scored *above average effectiveness* on TVAAS and were rated between *above expectations* and *significantly above expectations* on the three TEM measures, indicating that MTR is choosing high-quality teachers to mentor its residents.

Introduction

Beginning not long after its establishment in 2009–10, Memphis Teacher Residency (MTR) has received an annual evaluation of its program from its primary partner, Shelby County Schools (SCS; formerly Memphis City Schools before the city–county schools merger in 2013–14). MTR and SCS were partners in a state-administered, federally funded grant from 2011–12 to 2013–14, which funded the evaluation for those years. Ever since the grant ended, MTR has hired SCS to continue evaluating the effectiveness of MTR-trained teachers working in SCS. This report presents the evaluation of MTR teacher effectiveness in SCS for 2017–18.

Memphis Teacher Residency

MTR is a teacher recruitment and training program working in collaboration with Union University. Those accepted into the MTR program first complete a residency year, for which they receive free tuition, free housing, and a stipend of $1,000 per month. The residency year consists of a summer of intensive coursework, followed by a school year of apprenticing under an experienced mentor teacher in a high-need Memphis public school (regular or charter) every Monday through Thursday, and attending classes every Friday and most Saturdays.

Upon successful completion of the residency year, each MTR resident is awarded a Master of Urban Education. In exchange, MTR residents commit to teaching for three years in a Memphis urban school (public, charter, or private) in which at least 60% of the students qualify for free or reduced-price lunch. MTR graduates continue to receive formal and informal support and professional development throughout their three-year teaching commitment.

MTR requires graduates who do not fulfill their three-year teaching commitment to reimburse MTR in the amount of $10,000 for each unfulfilled year—for a potential total of $30,000. This provides strong incentive for MTR graduates to enter the teaching profession and remain in it for at least three years.

Methods

This evaluation analyzed four measures to compare the teaching effectiveness of MTR-related teachers with that of their non-MTR-related counterparts: 1) MTR-trained teachers were compared to teachers who did not go through the MTR program, and 2) mentors of MTR residents were compared to the rest of the district.

To assess the performance of the MTR-trained teachers, MTR/non-MTR comparisons were made for two groups of teachers: those with one to three years of SCS teaching experience and those with four to seven years. Three years is a logical cut-point because of: 1) the continued professional support MTR teachers receive for the first three years after graduating from the program, and 2) the financial arrangement mentioned above. The MTR graduates with more than three years of experience had been teaching in the district between four and seven years, hence the 4–7 year category.

To measure the quality of the MTR mentors, the mentors of the 2017–18 MTR residents were compared to the rest of the district, regardless of years of experience. Years of experience was excluded from the analysis for one simple reason: mentor teachers should be of high quality in general, not just in comparison to teachers with similar experience. An ineffective teacher with X years of experience should not be a mentor teacher even if he happens to be a little more effective than the average teacher with similar experience. In other words, the goal is to recruit truly high-quality teachers, not just those who are good for their experience level. About half of the 2017–18 MTR mentors had only three to five years of SCS teaching experience, while roughly a sixth had been teaching for six to nine years, and about a third had been teaching for 10 to 31 years.

Independent samples t-tests were employed to make the comparisons. This method compares the mean score of one group to the mean score of another group and assesses the statistical significance of the difference between the two means. The four measures used for the MTR/non-MTR mean comparisons are discussed below.

Measures

TVAAS Scores

The first measure consists of teachers' scores from the Tennessee Value-Added Assessment System (TVAAS). Teacher-level TVAAS data are designed to show how much *growth* a teacher's students exhibited on achievement tests from one year to the next. A teacher's TVAAS scores are generated by assessing her students' performance on end-of-year state-mandated assessments in light of those students' past performance on such assessments. Students outpacing their past performance will raise the teacher's TVAAS score, while students falling short of their past performance will lower it.

Teachers of tested subjects received a separate 2017–18 TVAAS effectiveness score for each subject or grade level tested, so if a teacher taught eighth-grade mathematics and Algebra I, he received two effectiveness scores. Teachers were also assigned an overall composite score, calculated as a 3-year, 2-year, or 1-year average, depending on the number of years of data available for each teacher. This inconsistency in composite scores is unfortunate, because it is important to be able to look at a teacher's overall TVAAS performance from year to year, without influence from previous years. This is especially important for the first few years of a teacher's career, when we expect the most dramatic improvements to occur.

Since the focus of this evaluation is on MTR teachers' performance specifically in 2017–18, it was necessary to create a single-year composite score for each teacher in the study. This was accomplished by using a weighted average, taking into account the number of students in each tested grade/subject. So if a teacher taught eighth-grade mathematics to 30 students and Algebra I to 120 students, the calculation gives the Algebra I score four times more weight than the eighth-grade mathematics score.

Each teacher's 2017–18 TVAAS composite was computed using the standardized, continuous index variable that underlies the ordinal TVAAS levels of 1 to 5 with which most Tennessee educators are familiar. The index variable consists of an estimate (produced by the TVAAS statistical model) divided by its standard error. The TVAAS effectiveness levels are derived from the values of the index variable as indicated in Table 1.

Table 1. Designation of TVAAS effectiveness levels based on TVAAS index

Level	Label	Underlying index
1	Least effective	index < -2
2	Approaching average effectiveness	-2 ≤ index < -1
3	Average effectiveness	-1 ≤ index < 1
4	Above average effectiveness	1 ≤ index < 2
5	Most effective	2 ≤ index

TEM Component Scores

The other three measures employed in this evaluation are all components of the TEM (Teacher Effectiveness Measure), which is SCS's teacher evaluation system. Every teacher receives a TEM score each year (unlike TVAAS, which applies only to teachers of tested subjects). The TEM comprises multiple measures: observations, professionalism, student perceptions, student achievement, and student growth. The TEM components are each rated on a 1–5 scale, as indicated in Table 2.

Table 2. TEM effectiveness ratings

Rating	Label
1	Significantly below expectations
2	Below expectations
3	Meeting expectations
4	Above expectations
5	Significantly above expectations

The components of the TEM are combined to create an overall TEM score. However, this evaluation separately analyzes three of the individual TEM components, because much of the overall TEM score consists—in varying degrees according to circumstance—of TVAAS performance. Since some teachers are assigned schoolwide TVAAS scores because they do not teach tested subjects, it is much more desirable to analyze separately: 1) TVAAS scores (for teachers of tested subjects), and 2) components of the TEM that are not related to achievement tests (for all teachers). The three non-achievement-related TEM components are discussed below.

Observation scores. Certified TEM observers conduct announced and unannounced observations of all SCS teachers (and other certified staff) throughout each school year. Principals, vice principals, assistant principals, PLC (professional learning community) coaches, content specialists, district coaches, National Board Certified Teachers, and specialty teachers were among the designated positions eligible to become certified TEM observers for 2017–18. New hires were to be observed four times throughout the year: once announced and three times unannounced. Teachers who scored less than a 3 on any one of several TEM components the previous year were to have one announced and two unannounced observations. All other teachers were to have one announced and one unannounced observation. A teacher's final score on the observation component of the TEM consists of the average of all her observation scores for the year.

Professionalism scores. From the 2017–18 TEM Manual (p.12):
> *The Professionalism component is designed to capture a teacher's efforts to enhance their practice through professional learning and growth, use of data, school and community involvement, and leadership....*
>
> *Teachers and school administrators will collect and document evidence of a teachers' [sic] professionalism throughout the course of the school year. The school administrator will then review the evidence and determine a final score during [a conference] at the end of the school year....*
>
> *Administrators and teachers are encouraged to use the Professionalism Rubric Scoring Guide to assist in evidence, artifact collection, and assigning ratings. Teachers should submit 3-5 relevant artifacts per indicator from the current school year.*

Student-perception scores. SCS students are surveyed twice a year using the Tripod survey instrument. According to the 2017–18 TEM Manual (p.11), the Tripod survey:
> *...asks students to assess observable teaching practices in their classroom based on what we refer to as the "Seven Cs":*
> - *Caring about students (encouragement and support)*
>   - *Ex: "My teacher in this class makes me feel that s/he really cares about me."*
> - *Controlling behavior (press for cooperation and peer support)*
>   - *Ex: "Our class stays busy and doesn't waste time."*
> - *Clarifying lessons (success seems feasible)*
>   - *Ex: "My teacher explains difficult things clearly."*
> - *Challenging lesson (press for effort, perseverance, and rigor)*
>   - *Ex: "My teacher wants me to explain my answers – why I think what I think."*

- *Captivating* students (learning seems interesting and relevant)
  - *Ex:* "My teacher makes learning enjoyable."
- *Conferring* with students (students sense their ideas are respected)
  - *Ex:* "My teacher wants us to share our thoughts."
- *Consolidating* knowledge (ideas get connected and integrated)
  - *Ex:* "My teacher takes the time to summarize what we learn each day."

Student Perception Survey data for a teacher will factor into a single score based on a Normal Curve Equivalent (NCE), similar to a percentile rank and create an index score of 1-5, similar to TVAAS scores. The ratings for the Student Perception Survey are calculated based on the comparison of all teachers in each SCS survey level [grade bands K–2, 3–5, 6–8, and 9–12].

Results

The findings of the TVAAS analyses are presented in Table 3, which displays results first in the units of the index variable (i.e., standard errors) and then converted into effectiveness levels as described in Table 1. While the index units show the group comparisons more precisely, the effectiveness levels may be more readily understandable. The TVAAS effectiveness levels of 1 through 5 are well known to most Tennessee educators, so seeing the standard error units translated into the effectiveness levels can help provide substance and context to the results.

As Table 3 illustrates, MTR teachers with 1–3 years of teaching experience outperformed their non-MTR counterparts by 1.33 standard error units. Although both groups' mean TVAAS composite fell within the Level 3 effectiveness category, a difference of one and a third standard error units is nonetheless quite large. Note that the non-MTR teachers were just over a hundredth of a point away from averaging Level 2 (*approaching average effectiveness*), while their MTR counterparts were well within the top half of Level 3 (*average effectiveness*). As for the teachers with 4–7 years of experience, however, the mean difference in scores between MTR and non-MTR teachers was not statistically significant.

Turning now to the MTR mentors, their TVAAS performance far outpaced that of the rest of the district. They averaged 1.72 standard error units above the non-mentors—a very large difference. Moreover, the mentors averaged Level 4 (*above average effectiveness*), while the rest of the district averaged Level 3 (*average effectiveness*).

Table 3. Mean differences† between MTR and non-MTR teachers' 2017–18 TVAAS composites

Years of experience	Teacher group	Mean TVAAS composite‡		Mean difference^		N	Statistical significance
1 to 3	MTR	0.34	Level 3	1.33	0 levels	66	YES
	Non-MTR	-0.99	Level 3			679	(p=.000)
4 to 7	MTR	-0.29	Level 3	-0.72	0 levels	28	NO
	Non-MTR	0.43	Level 3			332	(p=.212)
N/A	Mentor	1.55	Level 4	1.72	1 level	29	YES
	Non-Mentor	-0.17	Level 3			2,788	(p=.002)

† assessed via independent-samples t-test

‡ method for obtaining TVAAS composites described in "TVAAS Scores" section; conversion of TVAAS index into effectiveness levels explained in Table 1; first column: TVAAS index (standard error units); second column: corresponding TVAAS effectiveness level

^ mean differences shown two ways:
 • TVAAS index: the given MTR group's mean TVAAS composite minus that of the non-MTR group
 • TVAAS effectiveness level: the given MTR group's mean effectiveness level minus that of the non-MTR group

Table 4. Mean differences† between MTR and non-MTR teachers' 2017–18 TEM component scores

Years of experience	TEM component	Teacher group	Mean score‡	Mean difference	N	Statistical significance
1 to 3	Observations	MTR	4.13	0.43	79	YES
		Non-MTR	3.70		1,365	(p=.000)
	Professionalism	MTR	4.37	0.48	79	YES
		Non-MTR	3.89		1,365	(p=.000)
	Student surveys	MTR	4.00	0.29	76	YES
		Non-MTR	3.71		1,009	(p=.021)
4 to 7	Observations	MTR	4.32	0.11	23	NO
		Non-MTR	4.21		795	(p=.339)
	Professionalism	MTR	4.64	0.27	23	YES
		Non-MTR	4.37		793	(p=.046)
	Student surveys	MTR	4.18	0.22	22	NO
		Non-MTR	3.96		568	(p=.298)
N/A	Observations	Mentor	4.52	0.38	32	YES
		Non-Mentor	4.14		5,744	(p=.000)
	Professionalism	Mentor	4.87	0.57	32	YES
		Non-Mentor	4.30		5,737	(p=.000)
	Student surveys	Mentor	4.27	0.38	30	YES
		Non-Mentor	3.89		4,060	(p=.006)

† assessed via independent-samples t-test; ‡ on a scale of 1 to 5

The results for the analyses of the three TEM components are presented in Table 4, which shows that MTR teachers with 1–3 years of teaching experience outperformed their non-MTR counterparts on all three measures. Those MTR teachers averaged 0.43 points higher (more than a third of an effectiveness rating) on observations, 0.48 points higher (almost half of an effectiveness rating) on professionalism, and 0.29 points higher (under a third of an effectiveness rating) on student perceptions, as compared to their non-MTR counterparts.

As for the TEM results for the teachers with 4–7 years of experience, only one was statistically significant: on professionalism, the MTR teachers outperformed the non-MTR teachers by 0.27 points (over a quarter of an effectiveness rating). The two groups did not differ significantly on observations or student perceptions.

The MTR mentors' TEM performance exceeded that of the rest of the district on all three measures: by 0.38 points (over a third of an effectiveness rating) on observations, by 0.57 points (over half of an effectiveness rating) on professionalism, and by 0.38 points (over a third of an effectiveness rating) on student perceptions.

Conclusion

The analyses presented here show that in 2017–18, MTR teachers in their first three years of SCS teaching outperformed their non-MTR counterparts by significant margins on four very different measures of teaching effectiveness: student growth on achievement tests, observation of practice, professionalism, and student perceptions. This provides robust evidence that MTR is succeeding at its primary mission: to provide Memphis high-need schools with *effective* teachers.

Moreover, the teachers whom MTR recruited to mentor the 2017–18 residents outperformed the rest of the district on all four measures of teaching effectiveness by substantial margins. Their mean TVAAS composite fell within Level 4, *above average effectiveness*, and their mean ratings on all three TEM measures were approaching the top of the scale, falling between *above expectations* and *significantly above expectations*. Taken together, these findings indicate that MTR is recruiting high-quality teachers to mentor its teachers-in-training. The yearlong apprenticeship under an effective mentor teacher is arguably the most crucial aspect of the MTR training experience. MTR thus appears on track to continue providing SCS with effective teacher recruits.