# Longitudinal Analysis of the Correlation of TEAM Observer and Self-Scores

Technical Report

Clint Sattler
Supervisor, Research and Evaluation
Knox County Schools
Department of Research, Evaluation, and Assessment

**January 2018**

**Overview**

The Tennessee Educator Acceleration Model (TEAM) is an educator evaluation process that uses frequent observation and feedback to support educator growth. As part of the TEAM process, educators are rated on a 1 to 5 scale regarding performance in various indicators. The TEAM process requires that educators reflect upon their performance by rating themselves on the same indicators during each evalaution.

At the start of SY1718, the Knox County Schools (KCS) transitioned from collecting TEAM evaluation data in the RANDA Tower platform to the Tennessee Department of Education's TNCompass platform. In the RANDA tower platform, educator self-scores were collected in a form physically separated from observer scores. Although observers could view educator self-scores while they were assigning their own indicator scores, to do so required a deliberate effort. With the transition to TNCompass, educator self-scores are currently visible to observers whenever an observation is being scored (Figure 1).



*Figure 1: Sample Display of Observation and Self-Scores in TNCompass*

KCS principals have suggested that having the educator self-scores displayed during the scoring process may bias observer scores. This analysis attempts to quantify an impacts the TNCompass graphics may have on observer scores. The analysis uses an interrupted time-series approach to determine if the correlation between observer scores and self-scores have shifted with the introduction of TNCompass. However, the findings of this study should not be considered causal. No formal experiment was conducted to determine that any change in correlation was due solely to the data collection platform.

## Methodology

The scope of this study included observer and self-scores collected on the TEAM educator rubric from SY1415 through SY1718. Only first semester data was considered in this analysis because only first semester data was available for SY1718. Additionally, this analysis only used data collected in schools that utilized the TEAM observation process during every year from SY1415 through SY1718. Data collected at 24 locations that used an alternative evaluation system (either TAP or TIGER) was excluded from the analysis. The deployment of alternative systems generally accompanied additional training on evaluation rubrics and more extensive rubric-specific professional development activities which may bias findings.

Observer scores from SY1415 through SY1617 were retrieved from archived spreadsheets on an internal KCS server. Self-scores from these years were accessed from archived databases of RANDA Tower forms. SY1718 observation data (both observer scores and self-scores) were accessed via TNCompass reporting functions.

Fields of data were combined to create unique variables to link RANDA Tower observation scores with self-scores. Teacher license numbers (the license number of teachers being observed) were combined with their observer's name to create the unique linking variable. The month of the observation post conference was included in cases for which the combination of teacher license number and observer name did not result in a unique linking variable. Post conferences may have been held in different months than the teacher observation if the observation occurred at the end of a month. The author manually linked these entries when the evidence (both the dates and the personnel involved with the observation) suggested the records originated from the same observation event. This process was used to link records from SY1314-SY1617.

A more efficient linking variable was available in the SY1718 (TNCompass) data. Teacher license number and observation date were combined to link observer scores to self-scores. This process did not require any manual intervention in order to link the data.

Pearson's product moment correlation coefficients (r) were calculated for each academic year to determine the degree of correlation between individual indicator scores recorded by observers and educators. Basic bootstrapping with 2,000 samples was used to construct 95% confidence intervals. Correlation coefficients and bootstrapped confidence intervals were calculated using R version 3.4.3 and RStudio version 1.0.143. Histograms of observation scores and self-scores were constructed to ensure near-normality of the data. Histograms were constructed using SPSS version 24.

## Results

The mean and standard deviation for observation and self-scores are available in Table 1. The mean for both observer and self-scores have increased each year. The mean observer score became greater than the mean self-score beginning in SY1617.

*Table 1: Annual Means and Standard Deviations for TEAM Indicators*

| Academic Year | N | Observer Scores | | Self-Scores | |
|---|---|---|---|---|---|
| | | Mean | St. Dev | Mean | St. Dev |
| SY1415 | 42083 | 3.57 | 0.79 | 3.61 | 0.75 |
| SY1516 | 41510 | 3.64 | 0.78 | 3.65 | 0.76 |
| SY1617 | 40762 | 3.69 | 0.77 | 3.67 | 0.75 |
| SY1718 | 44967 | 3.78 | 0.77 | 3.75 | 0.77 |

The Pearson product moment correlation was calculated to determine the correlation between indicator observer scores and self-scores. The Pearson product moment correlation coefficient with bootstrapped 95% confidence intervals are available in Figure 2. SY1718 was the only year in the study that did not overlap the confidence interval of other academic years. This suggests a significant change in the level of correlation between observer scores and self-scores occurred during the first semester of SY1718 when compared to previous year.
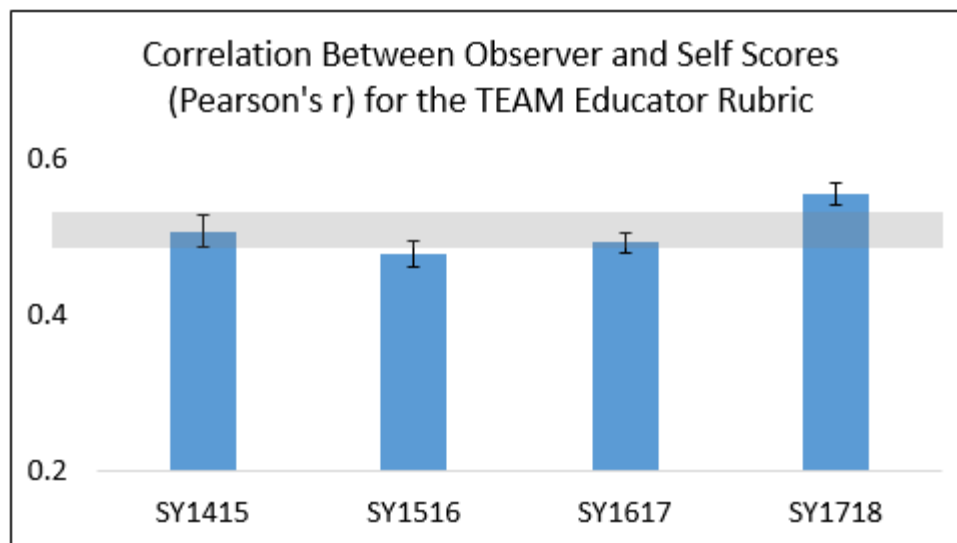


| Parameter | SY1415 | SY1516 | SY1617 | SY1718 |
|---|---|---|---|---|
| r | 0.5079 | 0.4789 | 0.4936 | 0.5551 |
| Upper 95% CI | 0.5284 | 0.4964 | 0.5063 | 0.5684 |
| Lower 95% CI | 0.4882 | 0.4618 | 0.4806 | 0.542 |
| N | 42083 | 41510 | 40762 | 44967 |

*Figure 2: Correlation Coefficients (with 95% CIs) - Observer Scores Vs. Self-Scores*

## Conclusions & Considerations

The change in the TEAM data collection platform presented a unique opportunity to study the possible impacts the platform may be having on observer scores. Qualitative feedback suggested that the "easy" visibility of the self-scores in the TNCompass system may lead to increases in bias in observer scores. There is some evidence that the immediate visibility of self-scores in the TNCompass system may contribute to observer bias. The amount of correlation between observer scores and educator self-scores has significantly increased in SY1718 in comparison to previous years. This increase in the correlation coefficient occurred in the same year in which the data collection platform was changed to TNCompass.

It should also be noted that the current data collection process in TNCompass does not mirror the process in which an observer becomes certified for the TEAM process. No self-scores are visible during the TEAM certification process. TNCompass administrators may wish to align the certification and data collection processes.

**Addendum**

The following figure provides the mean TEAM observer and self-scores for the time periods related to this study.
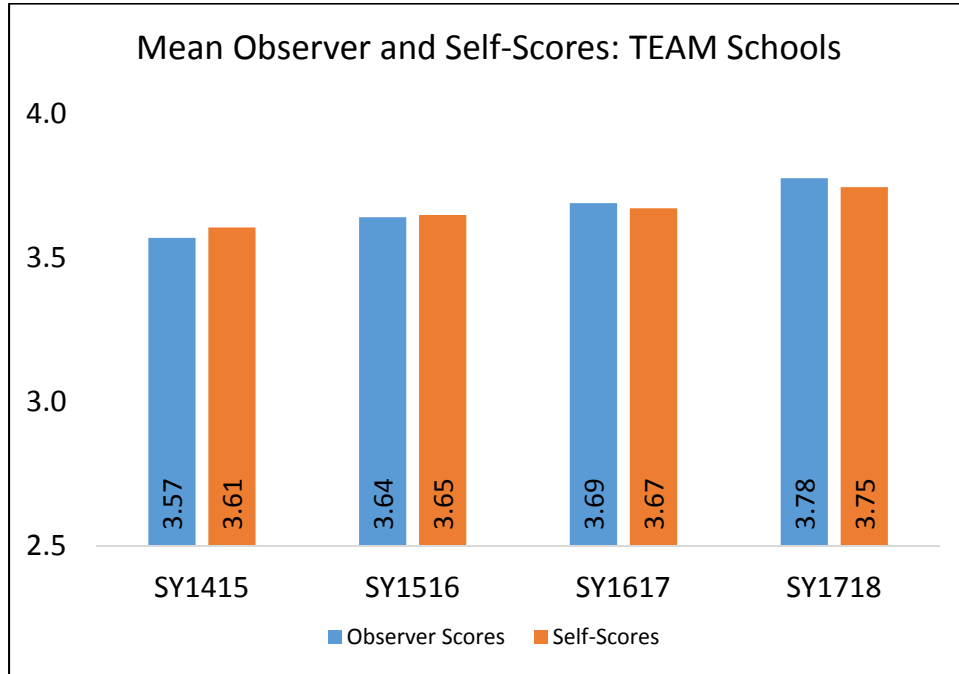


**Mean Observer and Self-Scores: TEAM Schools**

| Academic Year | N | Observer Scores | | Self-Scores | |
|---|---|---|---|---|---|
| | | Mean | St. Dev | Mean | St. Dev |
| SY1415 | 42083 | 3.57 | 0.79 | 3.61 | 0.75 |
| SY1516 | 41510 | 3.64 | 0.78 | 3.65 | 0.76 |
| SY1617 | 40762 | 3.69 | 0.77 | 3.67 | 0.75 |
| SY1718 | 44967 | 3.78 | 0.77 | 3.75 | 0.77 |

*Figure 3: Mean TEAM Scores*