Examining the Impact of Move to Learn among Third Grade Students in Mississippi Public Schools

Prepared for

The Bower Foundation

Conducted by:

Jerome R. Kolbo, PhD, MSW

Sherry Gilkey, MSW

Lei Zhang, PhD, MBA

Melissa Thompson, PhD, CC-AASP

Daniel Bond, PhD

Observers: Casey Johnson, Tina Davis, Elizabeth Johnson,

Erica Crosby, Heather Donnelly, Angela Alford

March 2017

Overview

This study examined the impact of Move to Learn (MTL) among third grade students in Mississippi public schools during Fall 2016. An initial evaluation of MTL during Fall 2015 explored the relationship between MTL and the duration of Time-on-Task (TOT) and the frequency of Time-on-Task Transitions (TOTT) among students. Changes in the design and procedures in the Fall 2016 study afforded greater opportunity to examine the direct impact of MTL on both TOT and TOTT.

2015 Evaluation

Methods and Findings

In 2015, the evaluation consisted of 900 direct classroom observations over six consecutive weeks in three elementary grade schools, and included 10 third grade teachers and 100 third grade students. During one class period per week, students were observed twice for 60 seconds and teachers twice for 180 seconds. Data were recorded by two trained observers using a computer application created specifically for this study. Data were recorded in the comma-delimited text files and imported into SPSS 22 for analysis. Data were examined as both continuous and categorical data.

The duration of TOT during the implementation of MTL in MTL classrooms was found to be significantly higher than among the same MTL classrooms before implementation and among the other non-MTL classrooms. In the MTL classrooms, the duration of TOT during implementation of MTL was significantly higher than the duration of TOT prior to implementation. The percent of students who were on-task for the entire observation period (60 seconds) increased significantly in the MTL classrooms but not in the non-MTL classrooms

The overall mean for frequency of TOTT (number of transitions between On- and Off-Task during the observation) during the last three weeks in MTL classrooms and the non-MTL classrooms were both found to be significantly different than among the MTL classrooms and the other non-MTL classrooms the first three weeks. In the MTL

classrooms, the mean TOTT frequency was significantly lower during the last three weeks than the first three weeks. However, in non-MTL classrooms, the mean TOT frequency during the last three weeks was also significantly lower than the first three weeks.

Frequency of TOTT was also grouped and examined as categorical data (0 time, 1 time, 2 or more times). The frequency of TOTT differed significantly by group. Specifically, students with no (0) transitions between On- and Off-Task increased from 72.0% the first three weeks to 84.1% during the last three weeks in the MTL classrooms. This was much greater than for the non-MTL classrooms, where the change was from 71.2% to 77.8% between the first three weeks and the last three weeks. In both cases, during the last half of the study, the vast majority of students (84.1% and 77.8%) stayed on task the entire time being observed.

Limitations

The 2015 study advanced the current state of knowledge regarding on-task behavior by examining both duration and frequency, and doing so for longer periods (60 seconds). The larger sample of students (100) and number of observations (900) allowed for various analyses of the data.

Yet, there were a number of limitations in the Fall 2015 study. One of the limitations of the study was that all teachers were selected by, and those who were to implement MTL were chosen by, the school administrators. It is possible that bias was introduced as they were not randomly assigned to implement MTL. Also, observers indicated differences in when the teachers administered MTL, some earlier, and others later in the class. As such there was little way to determine if any of the effects were due to other prior activities, such as arriving at school, coming in from recess, or just finishing lunch.

Another explanation for behavioral improvement among students was the length of the study. Initial data collection took place early in the school year when students were in new classes with new teachers. It's possible that as the study progressed

through the fall, students learned more about the classroom procedures and rules for their new class, thus improving behavior.

While the 2015 study extended the time of observations substantially from previous research (from approximately 5 second observations to 60 second observations), the findings that high percentages of students stayed on-task throughout the observations, suggested that the duration of observations should be expanded. The study did not assess if there were certain times of day, such as morning or afternoon where MTL might have a greater impact. Similarly, the study was unable to determine the how long the impact of MTL lasted, if at all, beyond the 60 seconds.

2016 Evaluation

Purpose

The purpose of the 2016 evaluation was to further examine the relationship between MTL and TOT and TOTT behavior by addressing shortcomings identified in the 2015 evaluation and including additional procedures to better control for possible threats to internal validity, such as:

- TOT and TOTT were observed for much longer periods of time (300 seconds)
- Observations began 20 minutes after class started regardless if MTL was or was not being implemented
- Randomly sampled students were observed only once in each class period
- Observations were conducted twice per week for only six weeks
- The order in which selected students were observed changed each class
- Neither students nor teachers knew which students were being observed
- MTL was implemented after four observations among the randomly selected Intervention Group
- MTL was implemented after eight observations among the randomly selected Delayed Intervention Group
- Observations occurred during early AM, late AM, early PM, late PM

- Extending the observations of the impact over time, recording data over five five-minute observations (for a total of 25 minutes) in each class.

Method

A quasi-experimental design with delayed start was employed over six weeks with observations of 10 randomly selected students in all 12 classrooms. Prior to data collection, observers sat in the classrooms two times in one week in order to minimize the potential bias of their presence during the study and to familiarize themselves with the classroom procedures. During these two times no data was collected.

Among all 12 classrooms, no MTL was used for the first four observations (O). Prior to the fifth observation, six classrooms were randomly selected to begin implementing MTL (X) and to do so throughout the remainder of the study. They became the Intervention Group (R_I). The other six classrooms were not allowed to begin implementing MTL until the ninth observation. They could then use MTL through the remainder of the study. They became the Delayed Intervention Group (R_{DI}).

		Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
R_I	O O	O_1 O_2	O_3 O_4	X_5 X_6	X_7 X_8	X_9 X_{10}	X_{11} X_{12}
R_{DI}	O O	O_1 O_2	O_3 O_4	O_5 O_6	O_7 O_8	X_9 X_{10}	X_{11} X_{12}

In order to minimize the possibility of residual effects of the movement or other activities occurring before classes (arriving at school), lunch or recess, or other movement or activities between classes, the observers delayed recording TOT and TOTT behaviors for 20 minutes.

Subjects and Sampling

Three schools were identified and selected by the MDE OHS (Richland Upper Elementary, Stonebridge Elementary, and North Bay Elementary). The three schools were familiar with MTL. In addition, all three were viewed as having the leadership, willingness, and readiness to participate in the evaluation. The three selected schools

were first contacted summer 2016 by Scott Clements, Director of the OHS, informing them of the purpose and scope of the study and setting up a time to meet in August with the research team. At the meetings with three schools' administrators, the study design, requirements, and Human Subjects protocol were reviewed. Upon agreement to participate in the study, the school administrators were asked to identify four comparable teachers, all teaching at the third grade level, and to schedule another meeting with the four teachers from each of their schools.

At the meetings with the administrators and teachers in late August and early September, the study design, requirements and protocol were reviewed. Each teacher was provided a consent form. All teachers signed the consent forms and where then given packets of consent forms to send home to parents of the students in their selected classrooms. Teachers were instructed to collect all consent forms returned to the school. All consent forms were turned in directly to the Human Resources representative for each school. The representative would generate a student photo for all receiving consent in each classroom. No names or other identifying information was linked to the photos. All of the photos for each classroom were submitted to USM. No other identifying characteristics of either the student sample were recorded (i.e., gender, race, age). Even though more than 10 students received consent in the classrooms, only 10 students were randomly selected from each class. Neither the teachers, nor the students then knew which students were selected to be observed during the study.

Procedures

In July, 2016, the study received Institutional Review Board approval through the Human Subjects Committee at USM. In July, six observers for the study were then selected via an interview process from a pool of master-level student applicants in the discipline of social work. In August, training began with the Project Coordinator providing a detailed explanation of the MTL program, conceptual information of TOT and TOTT behaviors both of students and teachers, along with video clips of authentic teacher-student instruction time.

Operationalized on- and off-task behaviors identified in the empirical literature were reviewed during the training. Over the course of three separate 2-hour training sessions, the observers were able to practice recording time-on-task/time-off-task behaviors of students and teachers with the use of the iOS software designed for the use in the classroom. The observers were then tested on two separate occasions by observing a total of 8 children (4 per test session) in classroom observation videos using the software to test inter-rater reliability. A high level (87%) of inter-rater reliability was established by measuring the string of data collected of each second of a one-minute period observation per child.

Instruments and Data Collection

Data collection was completed using custom software created by USM for the project. The software created in the MTL pilot study was revised by USM iTech to accommodate the change in methodology and data requirements. The application was then installed on six iPads and used in the classroom observations.

Data recorded for this study and included: classroom, time of day, time of day category, observation number, move to learn status, order of observation, student number, accumulated seconds on task, accumulated seconds off task, accumulated seconds of external interruption, accumulated seconds of absent time, transition count, unique comment, and observation string.

Within each classroom, both observers monitored unique groups of 5 randomized students for a total of 10 students per classroom. There were no repeated observations within a single class session or between observers. In other words, the same student would not appear twice in a single list of 10 students. Randomization was used to prevent the list students for observation from being the same between the observers. Students were observed for 300 seconds with a 60 second transition period between students. During the transition period the observer could make optional notes about the observation within the application. A countdown timer alerted the observer of the upcoming observation and the student to be observed was identified using a school

supplied picture with no other identifying information. By default, when the observations began, the software was set to On-Task. If a student would go Off-Task, the observers would have to immediately select the Off-Task icon. Once a student was back On-Task, then the observers immediately selected the On-Task icon. In addition, if an external interruption (e.g., intercom announcement) occurred, observers were to select the External Interruption icon. Once an external interruption had passed, the observers were to select the On- or Off-Task icon, whichever was appropriate at that time. If a student was not present on the day of observation, the observer would select the Absent icon.

For two observations (prior to the study beginning) the observers sat in the classrooms with their iPads, but did not record any behaviors of students. The intent was to increase student and teacher comfort and familiarity with the observers in the classroom and make any final adjustments necessary to the software. The results of each observation were saved in comma separated values (CSV) format in a daily observation file. After completing all observations, the observers simply emailed and manually deleted the file from the iPad. The data were exported as CSV and collected into two electronic locations. The individual observation data sets were merged into a single total observation file. The resulting file contained all the data elements described above for all observations.

Data Analysis

The duration of TOT and frequency of TOTT were measured as both continuous variables and categorical variables. To examine the difference of the means by categories, Analysis of Variance (ANOVA) and Independent –Samples T-Test were used. To further understand the relationship, Chi-Square Test was used to compare the various categorical means within the Intervention and Delayed Intervention groups. All analyses were conducted using SPSS 22. The statistical tests were two-sided, and the mean differences were considered statistically significant if the p-value from the tests was less than 0.05.

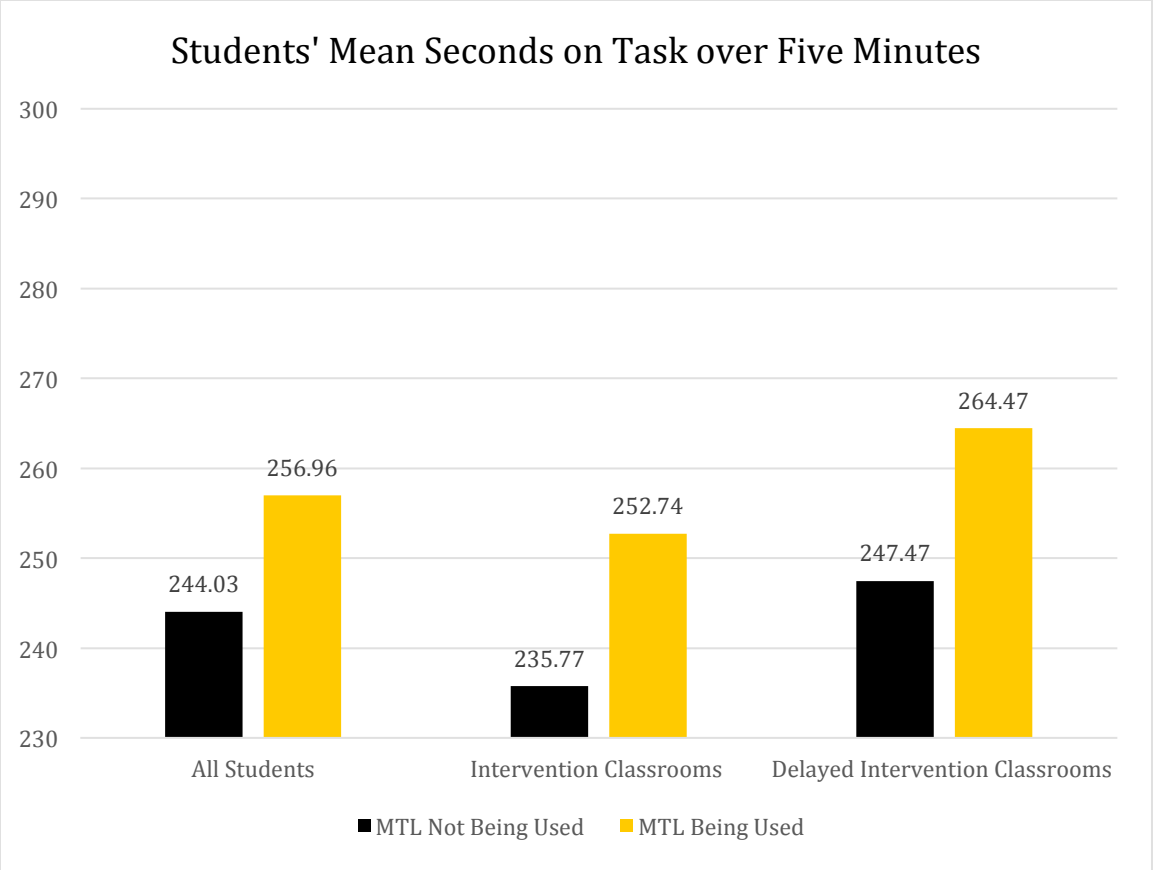
Findings

The evaluation initially consisted of 1,440 observations. After excluding observations with students who were categorized as absent during a majority of their observations, a final total of 1,272 direct classroom observations over 6 consecutive weeks were included in the analysis.

Time On-Task

Mean Seconds On-Task. The overall mean TOT for all students when MTL was being implemented (M=256.96) was significantly higher than the mean for all students when MTL was not being implemented (M=244.03; p=0.002). When examined among the Intervention Group (in which MTL was not implemented until the 5th observation), the mean TOT increased significantly from the first four observations (when MTL was not being implemented) (M=252.74 vs. M 235.77; p=0.013). Among the Delayed Intervention Group (in which MTL was not implemented until the 9th observation), the mean TOT (M=264.47) also increased significantly from the first eight observations (when MTL was not being implemented) (M=247.47; p=0.003) (See Figure 1).

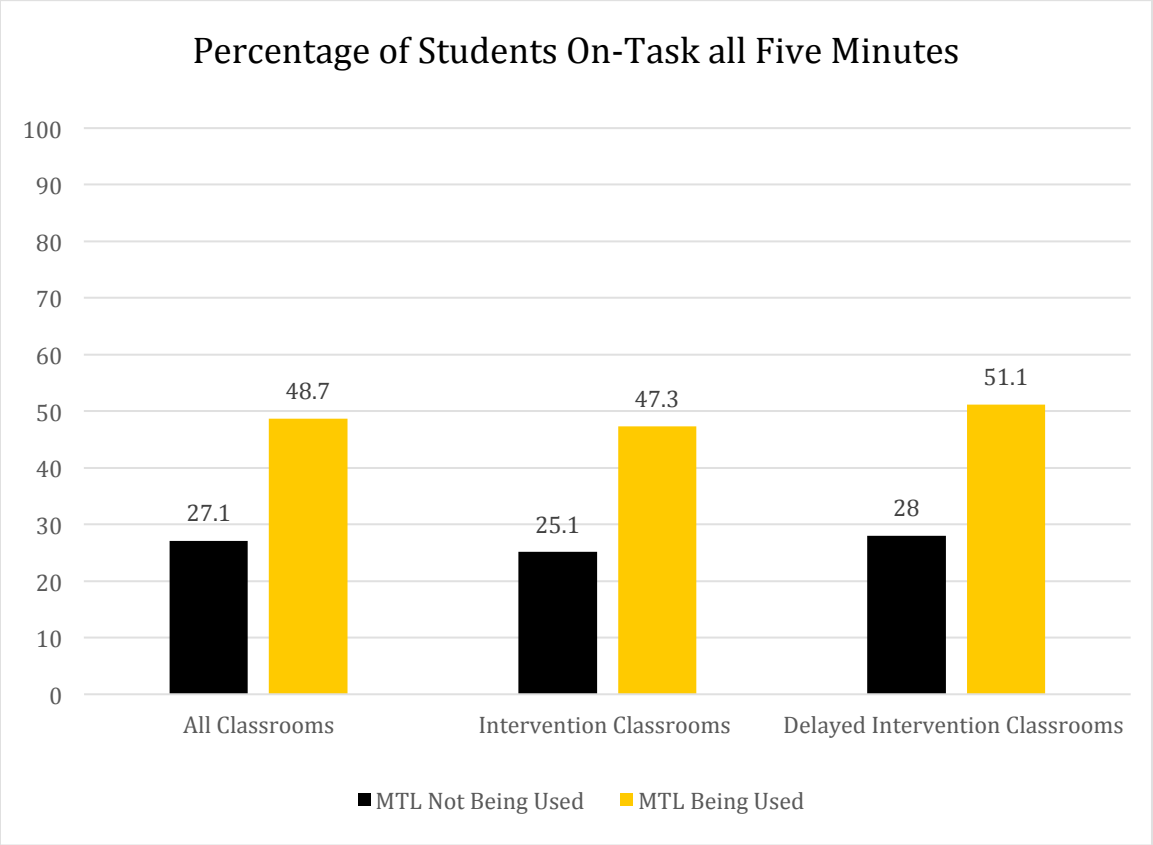
Figure 1: Mean Seconds On-Task



Percentage of Students Remaining On-Task. The percentage of students who remained on-task for the entire 300 seconds when MTL was being implemented (48.7%) was significantly higher than the percentage of students when MTL was not being implemented (27.1%; $p<0.001$). When examined among the Intervention Group (in which MTL was not implemented until the 5th observation), the percentage that remained on-task for the entire 300 seconds increased significantly (47.3%) from the first four observations (when MTL was not being implemented) (25.1%; $p<0.001$).

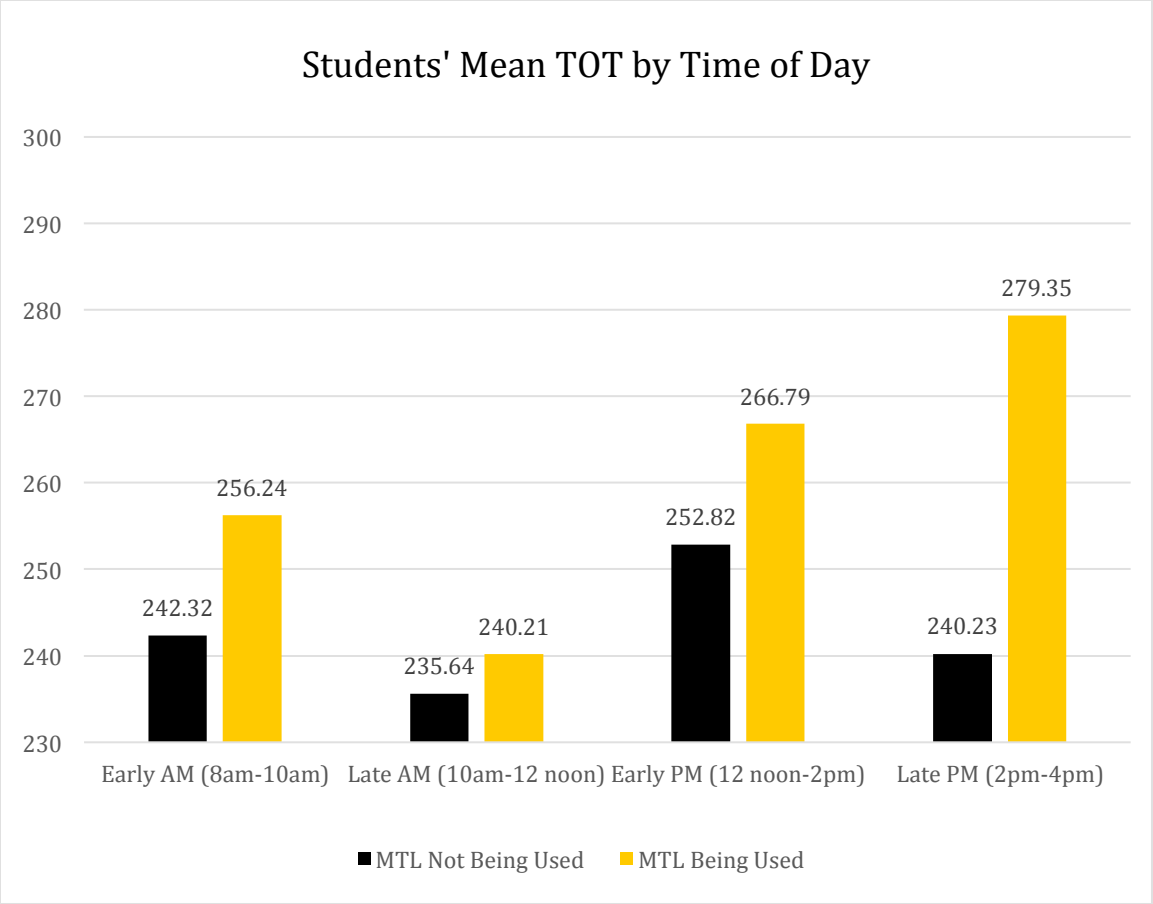
Among the Delayed Intervention Group (in which MTL was not implemented until the 9th observation), the percentage that remained on-task for the entire 300 seconds also increased significantly from the first eight observations (when MTL was not being implemented) (51.1% vs. 28.0%; $p<0.001$) (See Figure 2).

Figure 2: Percentage of Students Remaining On-Task



Mean Seconds On-Task by Time of Day. The overall mean TOT for all students when MTL was being implemented during all four data collected points: Early Morning (8:00am -10:00am)(M=256.24); Late Morning (10:00am – 12:00pm) (M=240.21); Early Afternoon (12:00pm – 2:00pm) (M=266.79); and (Late Afternoon (2:00pm – 4:00pm) (M=279.35) was significantly higher than the mean for all students when MTL was not being implemented during all four data collected points: Early Morning (M=242.32); Late Morning (M=235.64); Early Afternoon (M=252.82); and Late Afternoon (M=240.23) ($p < .001$) (See Figure 3).

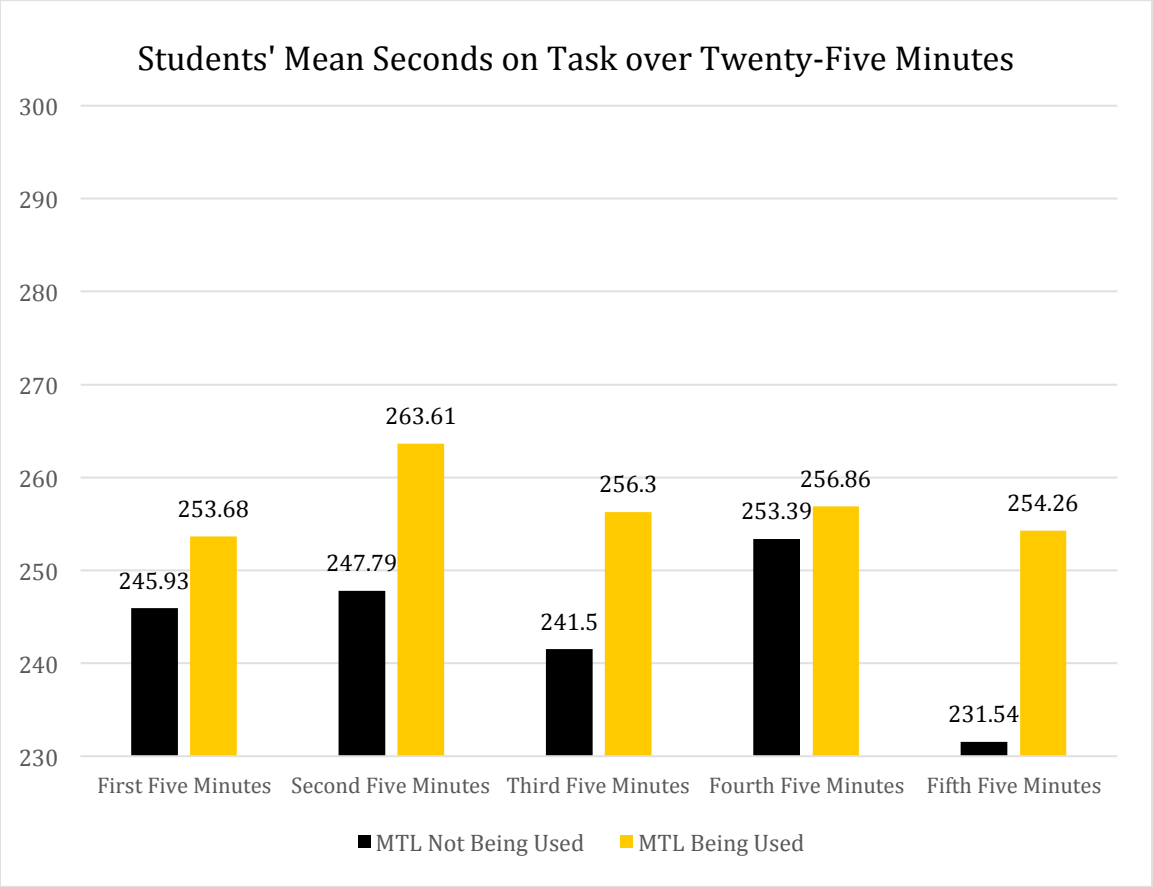
Figure 3: Mean Seconds On-Task by Time of Day



Mean Seconds On-Task over Time. The overall mean TOT for all students when MTL was being implemented did not change significantly over time (five five-minute continuous intervals): 1st 5 minutes (M=253.68), 2nd 5 minutes (M=263.61), 3rd 5 minutes (M=256.30), 4th 5 minutes (M=256.86), and 5th 5 minutes (M=254.26). In other words, TOT did not decline over 25 minutes of observation.

The overall mean TOT for all students when MTL was not being implemented also did not change significantly over time, however was much lower during all five five-minute continuous intervals: 1st 5 minutes (M=245.93), 2nd 5 minutes (M=247.79), 3rd 5 minutes (M=241.50), 4th 5 minutes (M=253.39), and 5th 5 minutes (M=231.54) (See Figure 4).

Figure 4: Mean Seconds On-Task over Time

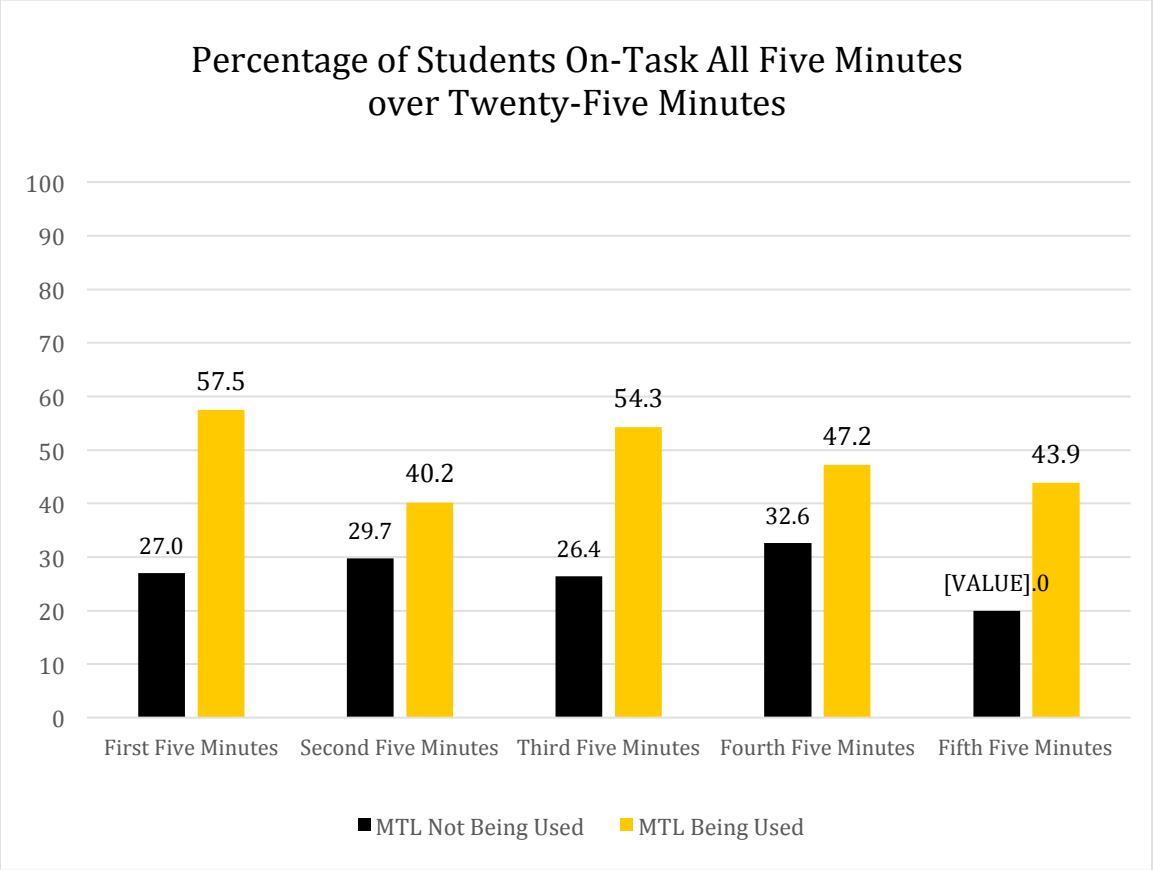


Percentage of Students Remaining On-Task over Time. When MTL was being implemented, the percentage of students who remained on-task for the entire 300 seconds changed significantly over time (five five-minute continuous intervals): 1st 5 minutes (57.5%), 2nd 5 minutes (40.2%), 3rd 5 minutes (54.3%), 4th 5 minutes (47.2%), and 5th 5 minutes (43.9%). In other words the percentage of students on-task for the entire 300 seconds declined over 25 minutes of observation (p=0.003).

However, the percentages remained much higher during all five-minute observations than when MTL was not being implemented: 1st 5 minutes (27.0%), 2nd 5 minutes (29.7%), 3rd 5 minutes (26.4%), 4th 5 minutes (32.6%), and 5th 5 minutes

(20.0%) (See Figure 5).

Figure 5: Percentage of Students Remaining On-Task over Time



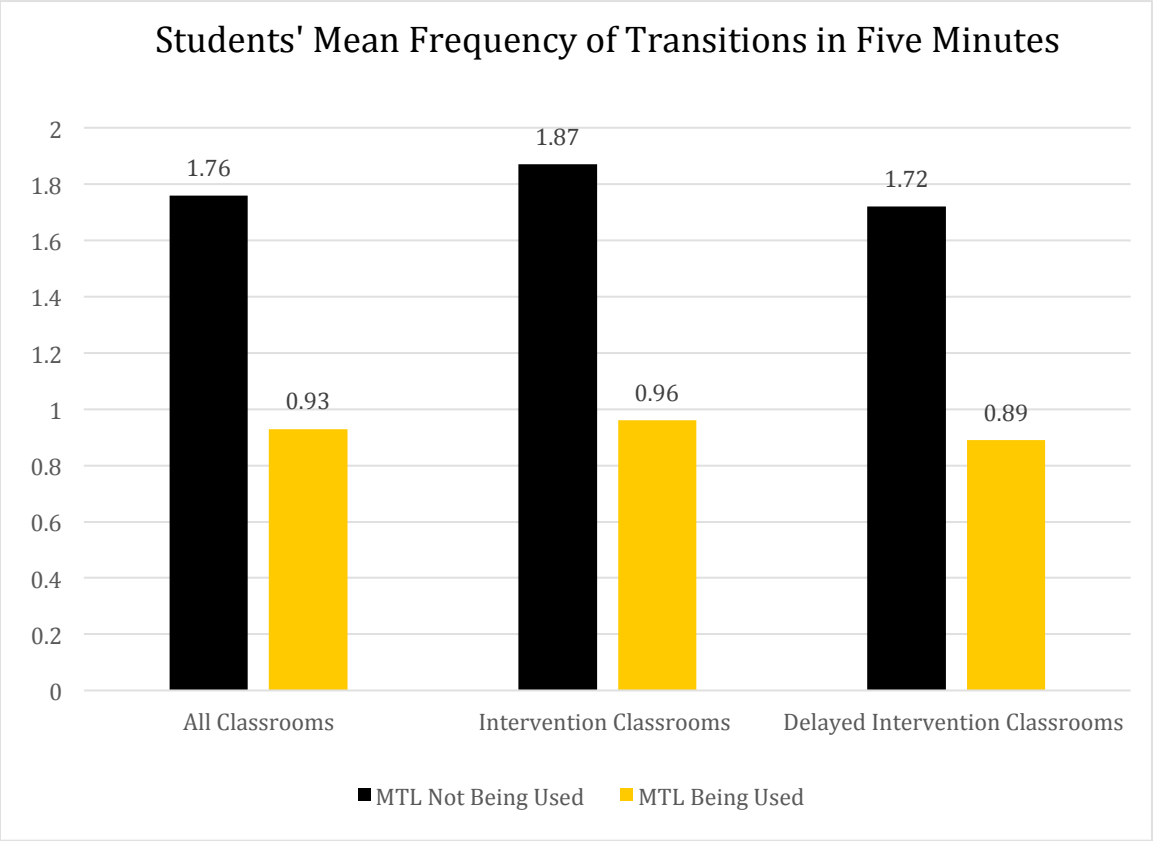
TOTT

Mean Frequency of Transitions. The overall mean TOTT for all students when MTL was being implemented (M=0.93) was significantly lower than the mean for all students when MTL was not being implemented (M=1.76; p<0.001). When examined among the Intervention Group (in which MTL was not implemented until the 5th observation), the mean TOTT decreased significantly from the first four observations (when MTL was not being implemented) (M=1.87 vs. M=0.96; p<0.001).

Among the Delayed Intervention Group (in which MTL was not implemented until

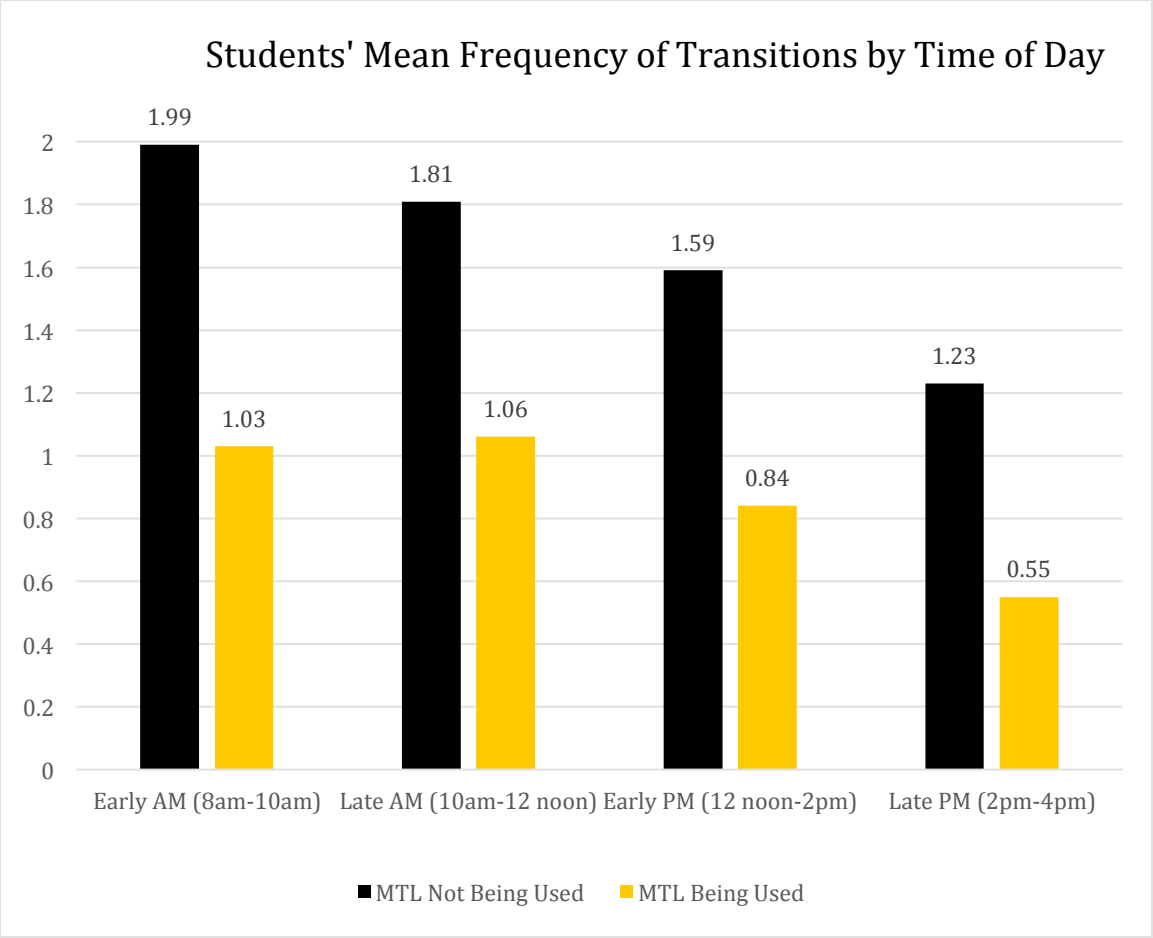
the 9th observation), the mean TOTT also decreased significantly from the first eight observations (when MTL was not being implemented) (M=1.72 vs. M=0.89; p<0.001) (See Figure 6).

Figure 6: Mean Frequency of Transitions



Mean Frequency of Transitions by Time of Day. The overall mean TOTT for all students when MTL was being implemented during all four data collected points: Early Morning (8:00am -10:00am) (M=1.03); Late Morning (10:00am – 12;00pm) (M=1.06); Early Afternoon (12:00pm – 2:00pm) (M=0.84); and Late Afternoon (2:00pm – 4:00pm) (M=0.55) was significantly lower than the mean for all students when MTL was not being implemented during all four data collected points: Early Morning (M=1.99); Late Morning (M=1.81); Early Afternoon (M=1.59); and Late Afternoon (M=1.23) (See Figure 7).

Figure 7: Mean Frequency of Transitions by Time of Day

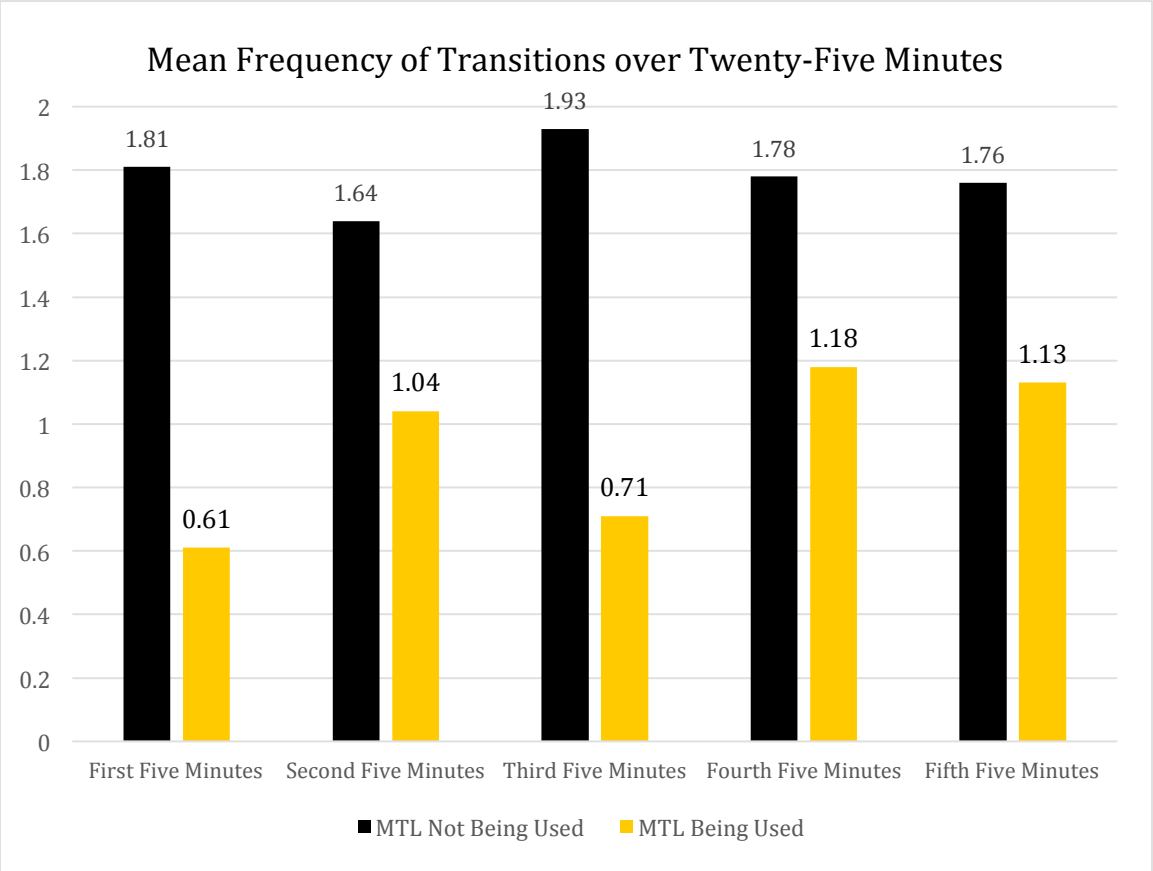


Mean Frequency of Transitions over Time. The overall mean TOTT for all students when MTL was being implemented did change significantly over time (five five-minute continuous intervals1st 5 minutes (M=0.61), 2nd 5 minutes (M=1.04), 3rd 5 minutes (M=0.71), 4th 5 minutes (M=1.18), and 5th 5 minutes (M=1.13). In other words, TOTT increased over 25 minutes of observation (p=0.003).

The overall mean TOTT for all students when MTL was not being implemented

was consistently higher than, but did not change over time (five five-minute continuous intervals1st 5 minutes (M=1.81), 2nd 5 minutes (M=1.64), 3rd 5 minutes (M=1.93), 4th 5 minutes (M=1.66), and 5th 5 minutes (M=1.78) (See Figure 8).

Figure 8: Mean Frequency of Transitions over Time



Limitations and Future Research

While this study advances the current study of MTL and the state of knowledge regarding on-task behaviors, there are a number of questions that arise and a number of ways the evaluation of MTL could possibly be improved in future studies. A couple of the limitations in this study are related specifically to data collection.

First, it is possible that the observers become biased by knowing that MTL is intended to improve TOT and minimize TOTT. This could potentially sway their observation recordings airing on the side of a child being on-task when in question. In the future, observers can be recruited and trained to observe on-task behaviors without necessarily knowing the specific purpose of evaluating the impact of MTL.

Second, while the 2016 study extended the time of observations substantially from previous studies and from the first study using MTL, there were still a high percentage of students who remained on task for the entire duration, which suggests the amount of time may need to be extended. Research focusing on attention of children suggests on average one minute should be allowed for each year of age (i.e. 10 years of age = 10 minutes). Future studies should extend the observation periods to 15 minutes per child in order to determine when the drift to off-task behaviors begin to occur, thus alleviating the recurring ceiling effect in both MTL studies.

We don't know the optimal use (dose) of MTL. This study only examined the implementation once per class period after 20 minutes into the class. It is unknown if the implementation of MTL more times during the day or even a class would further increase TOT or reduce TOTT. Similarly, this study has been conducted over a few weeks in the fall of the year. It is unknown if the effect strengthens or fades as the school year continues.

While classrooms were randomly assigned to implement MTL, it is unknown what impact, if any, the different teaching styles of the teachers had on the implementation of MTL or the on-task behaviors. It is unknown if MTL is more effective when specific videos are paired with specific class topics (e.g., math, science, or language arts).

Another limitation is what we know about the subjects. No identifying information was captured for the students, such as gender, race, age, prior academic performance, learning disabilities, or mental and physical health--all of which could allow for even greater analyses and understanding of the duration of TOT and frequency of TOTT. As such, the data from this study do not indicate whether MTL is more of less effective with

students with different characteristics.

Another limitation falls in our understanding of what aspects of MTL are responsible for the increases and sustained impact of on-task behaviors is warranted. Or, is it even the MTL video or how the teacher uses it that really matters? Related studies using physically active lessons have questioned if it is the physical activity component, being able to participate in something fun that others are excited about, or if the improved results are simply due to a break in academic learning environment.

Finally, it is recommended that future studies of MTL consider the examination of MTL and on-task behaviors on academic and behavioral performance in the classroom.

Acknowledgements

This study was collaborative effort between the Mississippi Department of Education's Office of Healthy Schools, The University of Southern Mississippi, Richland Upper Elementary, North Bay Elementary, and Stonebridge Elementary. It would not have been possible without the leadership of the school administrators and willingness and cooperation of the classroom teachers.