TutorSmart works:
An external validation of a community-based tutoring model

Scott Andrews[1]
David Ledrick[2]
Tiffany Adamski[3]

October 2025

1. Associate Director of Strategic Improvement, TutorSmart. *Corresponding Author:* sandrews@tutorsmart.us
2. Board Member, TutorSmart. djledrick@gmail.com
3. Executive Director, TutorSmart. tadamski@tutorsmart.us

The authors would like to acknowledge Nancy Buderer's work on the statistical analysis and assistance in reporting the results. Ms. Buderer holds a master's degree in Biostatistics from the University of Michigan and has over 33 years of experience.

ABSTRACT

Introduction: TutorSmart is a community-based nonprofit organization that partners with schools to implement a data-driven, high-impact tutoring model focused on closing gaps in student achievement ("gap closure"). In the 2023-24 school year, TutorSmart partnered with Toledo Public Schools to support Out-of-School-Time literacy interventions for 3rd-graders. In this paper, we compared the scores on the standardized state tests of participants in the TutorSmart program to those of Business as Usual students as a means of external validation for TutorSmart's effectiveness.

Methods: This was a descriptive study that used literacy scores provided by Toledo Public Schools from the i-Ready Test and Ohio's State Test for all 3rd-graders on a Reading Improvement and Monitoring Plan (RIMP) at schools where the TutorSmart model was implemented. Students who participated in TutorSmart were compared to classmates who did not. The TutorSmart group was further evaluated based on the students' level of participation. Mean test scores between groups were compared for each exam session using two-tailed Student's t-tests.

Results: At schools where TutorSmart ran programming, 224 students were on RIMPs; fifty-four (54) were enrolled in TutorSmart, leaving one hundred seventy (170) as the comparison group. The two groups started with similar i-Ready (mean 451 vs. 446, p=0.51) and OST (mean 662 vs. 653, p=0.10) scores. By Spring 2024, the TutorSmart group had significantly higher i-Ready performance (mean score 493 vs 469, p=0.02; mean change 40 vs. 25, p=0.005). A subgroup analysis showed that students who attained either 30 hours of tutoring or a 66% attendance rate (42 students) demonstrated significant gains [mean change TPS 25 (29) v. TutorSmart 41 (28) p=.003], which were further accentuated if both (21 students) criteria were met [49 (26) p=.001]. The removal rate from a RIMP (attaining 700 on the Ohio's State Test) did not reach statistical significance for any subgroup, though the TutorSmart students (16/54 [30%]) fared better than the TPS comparison group (43/170 [25%]). This rate was further improved in the subgroups that met either one (15/42 [35%]) or both (8/21 [38%]) goals for attendance and dosage.

Conclusion: The TutorSmart program model was able to demonstrate significant gains on the i-Ready diagnostic test, and trended towards improved OST scores and RIMP removal rates. Student performance was enhanced in participants who met program requirements.

INTRODUCTION/BACKGROUND:

TutorSmart was established in 2017 to provide targeted academic support for underperforming students in the Greater Toledo area by partnering with local school districts[1]. It is a facilitating organization that provides support to classroom teachers (tutors) as they deliver research-based intervention in small-group settings before or after school. The instructional model is described in more detail in the methods section. It pays strict attention to: session frequency, group size optimization, curriculum alignment, pedagogical delivery, tutor consistency, and systematic progress monitoring. This structured approach aligns with district partners' academic needs and State proficiency requirements. In the 2023-24 academic year, TutorSmart served more than 350 students, with capacity expansion planned for subsequent terms.

TutorSmart focuses on students who have fallen behind expectations and, therefore, measures progress by estimating the number of months a student is able to catch up while participating in TutorSmart. There's certainly precedent for using equivalent months of school to measure effect size[2],[3]. As a metric, gap closure directly aligns with TutorSmart's mission and informs the use of scarce resources. These internal measures, however, would benefit from external validation against an independent, authoritative standard. Furthermore, as TutorSmart is a community-run nonprofit organization that depends on public support and donations, validating our internal results is essential when asking for the ongoing support of parents, taxpayers, politicians, and donors.

The TutorSmart Program Model:

TutorSmart administers the after-school tutoring program using professional teachers and literacy coaches to provide a best-practice approach to instruction[4],[5],[6]. The district pays for the instructors while TutorSmart administers the program. This includes enrolling and scheduling the students and teachers, providing a class schedule, providing the literacy coaches who provide feedback to the instructors, and maintaining the software licenses for the Mobymax system. TutorSmart provides access to educational resources, holds instructors and students to a rigorous instructional model, and tracks student progress with a comprehensive data collection system. The service is free to the school system, and students enroll at no cost to themselves.

Students on a reading improvement and monitoring plan (RIMP) are encouraged by the schools to participate. Instruction is tailored to the specific needs of students by using the MobyMax assessment tool that continuously monitors progress and tracks achievement [7], [8]. Interventions are strategically structured to promote academic growth through sequential

skill development and conceptual mastery, ensuring a methodical approach to learning progression. Students meet for an hour and a half three times a week with no more than four students per instructor, thus meeting the definition of high-dose tutoring[9], [10], [11]. TutorSmart provides specialized support through expert literacy coaching and data analysis assistance. This maintains high instructional standards in closing identified learning gaps. This approach enhances the effectiveness of tutoring sessions and may improve classroom instruction more broadly.

Objectives

This is a descriptive study combining the available data from Toledo Public Schools on standardized testing results with participation data from TutorSmart records. The objective of this study is to examine the association between performance on State testing and participation in TutorSmart. A secondary objective is to evaluate the performance of students who have different levels of participation in the program. A subgroup analysis would show the potential gains that could be made by the Program and may suggest an optimal level of participation.

MATERIALS AND METHODS

Inclusion criteria

In the 2023-24 school year, the TutorSmart model was implemented at ten school sites within the Toledo Public Schools district, eight of which served 3rd-graders. All 3rd-grade students on Reading Improvement and Monitoring Plans (RIMP) at these schools (n=224) were included in this study, with 54 of them participating in TutorSmart. The program is free to any student participating. The remaining 170 students served as the control group (Business as usual group).

The Toledo Public Schools district serves roughly 20,000 students, over 90% of whom are economically disadvantaged. Approximately 45% of students in the district are listed by the Ohio Department of Education and Workforce as “Black, Non-Hispanic”, followed by roughly 30% listed as “White, Non-Hispanic”, and 15% listed as “Hispanic”. Our study population included one school that is a bilingual English-Spanish magnet school, resulting in a slightly higher percentage of Hispanic students. Slightly more girls than boys participated (52% female, 48% male).

Data Collection and Analysis:

The primary data used in this analysis comprises student scores on two different standardized tests: i-Ready, a criterion-referenced adaptive assessment administered to the students three times in the 2023-24 school year, and Ohio’s State Test (OST) administered to the students twice in the 2023-24 school year. These scores for the relevant 3rd-grade students were obtained directly from Toledo Public Schools by formal data request. The data was deidentified and sent to an independent biostatistician for further analysis. Information pertaining to student attendance was collected by TutorSmart directly from the teachers at the instructional sites, which allowed two key TutorSmart metrics to be calculated: Dosage (how many hours of TutorSmart intervention a student has received) and Engagement (how often the student attended tutoring, shown as a %).

Statistical Methods

Test scores (specifically the “overall scale” scores) for i-Ready and for Ohio’s State Tests (OST) were described by group (TutorSmart vs. TPS) and time (since each test was administered to students more than once across the 2023-24 school year). Mean test scores for the two groups were compared at each time using two-tailed Student’s *t*-tests. The within-student change across time periods was calculated (Fall to Winter, Fall to Spring), and the mean change was tested for significance between groups using two-tailed Student’s *t*-tests.

A further subgroup analysis was performed on the 54 TS students depending on the degree to which they met certain program requirements. Students were considered to be ‘protocol compliant’ (pc) if they either attended 30 hours of tutoring (Dosage) or if they came to at least

two-thirds of the offered meetings during their period of enrollment (Engagement). Students who attained both these benchmarks were considered to ‘meet (the) model’ (mm). The mean, the standard deviation (SD), and the 95% confidence interval (CI) were described for each group. Because two different tests were conducted, we adjust for multiple comparisons by using a critical value of $p<0.025$ to indicate statistical significance (i.e., type I error level of 0.05 divided by 2 tests).

Students were classified as *removed from RIMP* if they had an OST Fall 2023 or Spring 2024 score of 700 or more. Group differences in the percentage of students removed from RIMP were tested using Chi-square tests. The odds of removal from RIMP between groups were expressed as an odds ratio with 95% confidence interval.

Data were analyzed with SAS v9.4.

RESULTS

TutorSmart vs. Business as Usual

Fifty-four (54) third-grade students were enrolled in the TutorSmart program. One hundred seventy (170) other third-grade students on RIMPs from the same schools served as the comparison group. In Fall 2023, the two groups had similar i-Ready scores (mean 451 vs. 446, p=0.51) and similar OST scores (mean 662 vs. 653, p=0.10). (Table 1). In Spring 2024, the TutorSmart students had significantly higher i-Ready scores compared to other TPS students (mean 493 vs 469, p=0.02). i-Ready scores improved more in the TutorSmart students from Fall to Spring compared to other TPS students (mean change 40 vs. 25, p=0.005).

There was not enough evidence to detect a difference in the rate of removal from RIMP between TutorSmart students and other TPS students (30% vs. 25%, p=0.53).

Subgroup analysis: dosage and engagement

Dosage (hours of tutoring) ranged from 10 to 76 hours with a median of 43 hours. Students were classified as high dosage if they attended at least 30 hours of instruction. Thirty-two (32) Of the 54 students in the study group met this criterion. There was insufficient evidence to demonstrate that the mean i-Ready and OST scores differed between these groups in the Fall or Spring. (Table 2) However, the within-student change in i-Ready scores from Fall to Spring was greater in the TutorSmart students than in the TPS group (mean change 47 vs. 28, p=0.04). The percentage of students removed from a RIMP was 31% in the high-dose group and 27% in the low-dose group. This difference was not a significant one.

Engagement ranged from 33% to 100% with a median of 72%. Thirty-two (32) TutorSmart students had high engagement (attended more than 66% of sessions), and 22 had low engagement (66% or lower). There was not enough evidence to show that the group had different mean i-Ready scores or fall OST scores. (Table 3) However, the high engagement group had significantly higher Spring OST scores (688 vs. 661, p=0.004) and higher within-student change from Fall to Spring (21 vs. 3, p=0.04) compared to the low engagement group. The percentage of students removed from RIMP was significantly greater in the high engagement group (41%) compared to the low engagement group (14%, p=0.03).

TPS vs. Protocol Compliant and Meets Model

Students were classified as protocol compliant (pcTS) if they met either one of the dosage or the engagement goals, and as meeting model (mmTS) if they met both. There were 43 protocol-compliant students. Compared to the control group, the pcTS students had significantly higher i-Ready Spring 2024 scores (469 vs. 495, p=0.013), and higher i-Ready within-student change from Fall to Spring (25 vs 41, p=0.003). Twenty-one (21) students were classified as met model. Compared to the control group, they had higher i-Ready Spring 2024

scores (469 vs. 501, p=0.039), and higher i-Ready within-student change from Fall to Spring (25 vs 49, p=0.001).

DISCUSSION

TutorSmart is a privately run nonprofit organization dedicated to supporting elementary education, similar to other organizations that have sought to fill a perceived need [12], [13]. It partners closely with the Toledo Public School system, the 6th largest school system in the state. At the study's inception, the 224 participating students

demonstrated limited proficiency with a mean OST score of 655, significantly below the career readiness threshold of 725. The established grade-level benchmark of 700 determined RIMP placement status.

In 2023, TPS was one of the 150 districts required to submit a reading achievement plan to the Ohio Department of Education. Given that Ohio requires a plan for reading achievement and monitors districts with struggling students[14], it is likely that students within the control group participated in other tutoring programs or alternative interventions. Rather than differentiating between specific interventions, these students were aggregated to represent a "business-as-usual" comparison group. Furthermore, our analysis encompassed all TutorSmart enrollees regardless of participation levels, reflecting real-world implementation conditions where student engagement varies due to external factors. In order to demonstrate the potential of the program, a subgroup analysis was done to track the performance of students who met the model's intended levels of participation. This subgroup, comprising students meeting both dosage and engagement goals, demonstrated nearly twice the growth rate of the control group, suggesting the program's optimal potential. Notably, the 42 students who met either of the attendance goals achieved comparable results, indicating that even consistent, if not maximum, participation can yield significant benefits.

TutorSmart's high-dosage tutoring model is primarily focused on gap closure, not to raise standardized test scores or to facilitate removal from Reading Improvement and Monitoring Plans (RIMPs). However, analysis of these secondary outcomes provides valuable insights into the broader educational impact. To validate the TutorSmart model against external benchmarks, this study examined data from two standardized assessments: the OST and the i-Ready (iR) assessment. The OST serves as the State of Ohio's primary tool for measuring student achievement against state learning standards and facilitating school performance comparisons[15]. The i-Ready is a state-approved diagnostic assessment, offering more frequent evaluation opportunities and providing both normative and criterion-referenced data[16]. This tool is designed to give a finer-grained skill assessment to teachers and is very consistent with the model TutorSmart employs in its cycles of assessment and intervention. We expect this is the reason the greatest gains were shown on the i-Ready assessment. While improvements in OST scores and RIMP removal rates were less pronounced, both metrics showed positive trends favoring TutorSmart participants.

Limitations: Several methodological constraints warrant consideration when interpreting this study's findings. This was a descriptive study, and students were offered a chance to participate rather than being assigned. Full participation represents a significant commitment in time after school. We expect this to be the primary reason

the study group was a third the size of the control group. This resulted in a limited power to reach statistical significance during subgroup analysis. Importantly, as well, this may have introduced some selection bias. The TutorSmart program's intensive structure—requiring three 75-minute sessions per week over a 23-week academic period—may be less accessible for some students. Despite the program's commitment to universal accessibility through teacher and principal referrals and the absence of financial barriers, the substantial time commitment could inadvertently favor students whose families possess greater capacity to facilitate regular attendance and academic engagement. Additionally, the study's implementation within a single school district may restrict the generalizability of results to broader educational contexts with different demographic compositions or resource allocations. These limitations highlight the need for future research incorporating randomized assignment or matched control groups to more definitively establish the program's efficacy across diverse student populations.

Future direction and conclusion. The trends in this study certainly suggest the efficacy of the TutorSmart model. Future work would include a randomized controlled trial and implementation in other districts with different demographics. Moving to an in-school model would certainly ameliorate the after-school time commitment. Finally, TutorSmart brought significant resources to the school system in an effort to close the achievement gaps of struggling students. An analysis of the cost of closing this gap is warranted.

REFERENCES

1. Butt, R.D., Litten, J., McClellan, P., Ryley, W., & Thomas, K. (2017, May). *Framework for tutoring services: a technical report submitted to the Toledo Community Foundation*. https://www.toledocf.org/wp-content/uploads/2019/09/A-Framework-for-Tutoring-Services_May-2017.pdf

2. Higgins, S. *The teaching and learning toolkit.* Education Endowment Foundation; The Sutton Trust: https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit

3. F. Yuan and D. K. Evans, *Equivalent Years of Schooling: A Metric to Communicate Learning Gains in Concrete Terms*. World Bank, Washington, DC, 2019. doi: 10.1596/1813-9450-8752

4. Institute of Educational Sciences. WWC Intervention Report: Start Making a Reader Today® (SMART®); U.S. DEPARTMENT OF EDUCATION June 11, 2007

5. Mackey EG. The Alabama coaching framework. Alabama State Department of Education (2020). The Alabama Coaching Framework. RMC Research Corporation. October 2020

6. Intervention brief: Reading Recovery. Institute of Education sciences. US Department of Education. June 2023. What Works Clearinghouse (WWC)

7. L. Brown and D. Johnson, “The impact of Moby Max, an individualized computer-assisted instruction, on math assessment scores of middle grade students.,” presented at the Society for Information Technology & Teacher Education International Conference, Association for the Advancement of Computing in Education (AACE), 2014, pp. 2228–2231. Accessed: July 05, 2025. [Online]. Available: https://www.learntechlib.org/primary/p/147350/

8. W. Koch, “MOBYMAX MATHEMATICS BENCHMARKER ASSESSMENTS AND STATE STANDARDIZED TESTS”. [White Paper] www.mobymax.com

9. Realizing the Promise of High Dosage Tutoring at Scale Preliminary Evidence for the Field MARCH 2024 Personalized Learning Initiative Research Team. UCHICAGO EDUCATION LAB: REALIZING THE PROMISE OF HIGH DOSAGE TUTORING AT SCALE , MARCH 2024

10. *High-impact tutoring: district playbook.* National Student Support Accelerator, in collaboration with BlueEngine. https://studentsupportaccelerator.org/sites/default/files/High_Impact_Tutoring_District_Playbook.pdf

11. Nikow, A., Oreopolus, P., & Quan, V. The promise of tutoring through pre-K-12 learning: a systematic review and meta-analysis of the experimental evidence. *American Educational Research Journal, 61(1)*, 74-107. https://doi.org/10.3102/00028312231208687

12. Surveying educational entrepreneurs: the headwinds and tailwinds to building new educational enterprises. *Edchoice.org. March 202*

13. Kim, J. Goodman, and M. R. West, “Kumon In: The Recent, Rapid Rise of Private Tutoring Centers,” 2021, doi: 10.26300/Z79X-MR65.

14. *Third Grade Reading Guarantee guidance*. (2024, October). Ohio Department of Education & Workforce. https://education.ohio.gov/getattachment/Topics/Learning-in-Ohio/Literacy/Third-Grade-Reading-Guarantee/Third-Grade-Reading-Guarantee-Guidance-October-2024.pdf.aspx?lang=en-US

15. *Understanding Ohio’s State Tests score reports 2023-2024.* Ohio Department of Education & Workforce. https://oh-ost.portal.cambiumast.com/content/contentresources/en/Understanding_State_Tests_Reports_2023-2024_Final.pdf

16. *i-Ready: Assessment for the Purpose of Instruction.* (2024). Curriculum Associates, LLC. https://ebooks.curriculumassociates.com/story/assessment-for-instruction/page/1

TABLES AND FIGURES

TABLE 1

Third Grade RIMP Students’ Outcomes with and without TutorSmart

	TPS Students		All TutorSmart Participants		p-value
No. Students	170		54		
	Mean (SD)	**95% CI**	**Mean (SD)**	**95% CI**	
i-**Ready**					
Fall 2023	446 (50)	438 - 454	451 (47)	438 - 464	0.51
Winter 2024	462 (57)	453 - 471	471 (54)	456 - 485	0.33
Spring 2024	469 (64)	459 - 480	493 (53)	478 - 508	0.02
Within-student change from					
Fall to Winter	17 (31)	12 - 22	20 (25)	13 - 27	0.59
Fall to Spring	25 (29)	20 - 30	40 (33)	30 - 49	0.005
OST					
Fall 2023	653 (35)	648 - 659	662 (33)	653 - 672	0.10
Spring 2024	673 (35)	667 - 679	676 (34)	667 - 686	0.54
Within-student change from					
Fall to Spring	20 (31)	14 - 25	14 (31)	5 - 23	0.27
Removed from RIMP, n (%)	43 (25%)		16 (30%)		

SD is standard deviation. CI is confidence interval.
P-values for test scores are from two-tailed Student's T-tests.
P-value for removed from RIMP is from a Chi-square test.
Removed from RIMP is defined as Ohio State test score of 700 or more.

TABLE 2

TPS versus TutorSmart Protocol Compliant and Meets Model

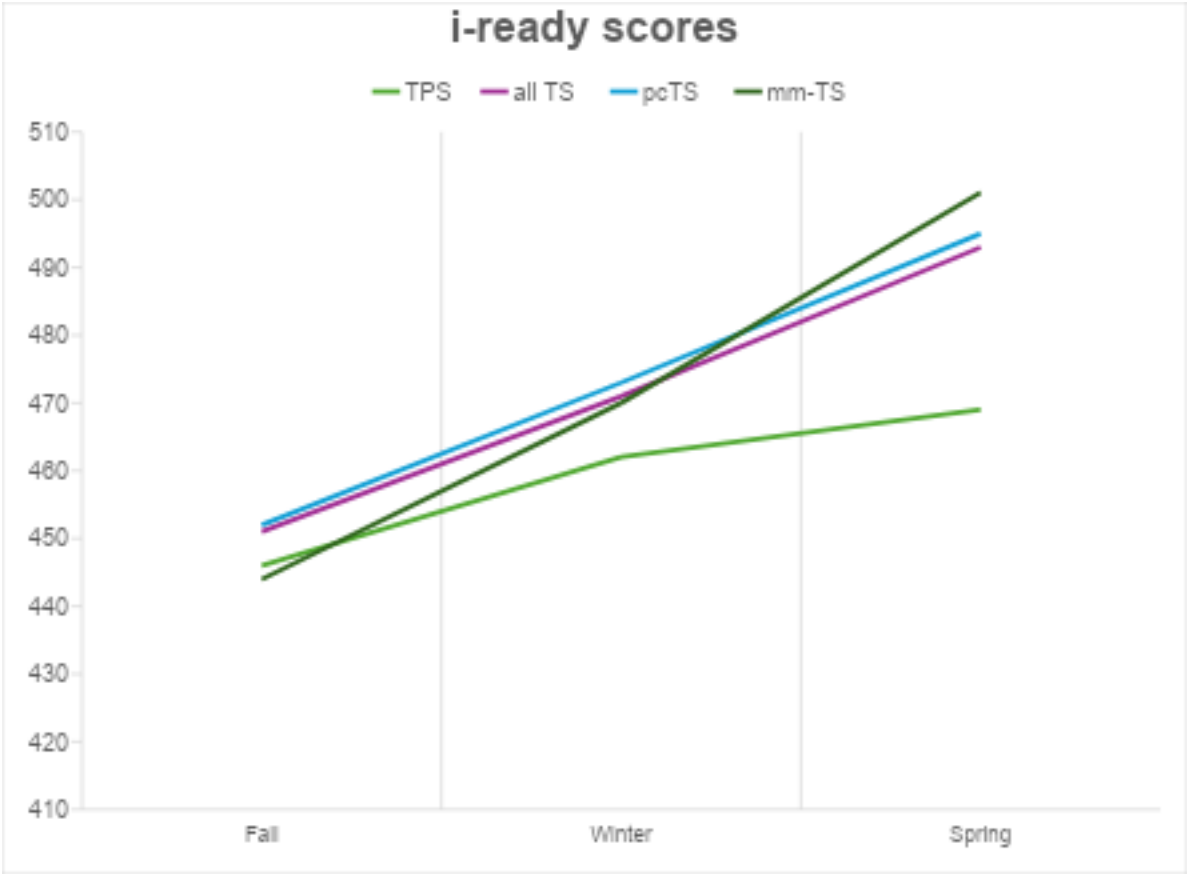
	TPS Students		TS Protocol Compliant >30 hours dosage OR >66% engagement		TS Meets Model >30 hours dosage AND >66% engagement	
No. Students	170		43		21	
	Mean (SD)	**95% CI**	**Mean (SD)**	**95% CI**	**Mean (SD)**	**95% CI**
i-**Ready**						
Fall 2023	446 (50)	438 - 454	452 (49)	436 – 468	444 (57)	418 – 471
Winter 2024	462 (57)	453 - 471	473 (58)	455 – 490	470 (66)	440 – 500
Spring 2024	469 (64)	459 - 480	495 (56)[a]	478 – 513	501 (55)[c]	475 – 526
Within student change from:						
Fall to Winter	17 (31)	12 - 22	21 (26)	12 – 29	25 (32)	10 – 40

Fall to Spring	25 (29)	20 - 30	41 (28)[b]	32 – 50	49 (26)[d]	36 – 62
OST						
Fall 2023	653 (35)	648 - 659	663 (35)	652 – 673	656 (35)	640 – 672
Spring 2024	673 (35)	667 - 679	681 (35)	669 – 692	686 (36)	669 – 703
Within-student change from Fall to Spring	20 (31)	14 - 25	17 (31)	7 - 27	27 (35)	11 - 44
Removed from RIMP, n (%)	43 (25%)		15 (35%)		8 (38%)	

Two-tailed p-value: TPS versus TS Protocol Compliant: [a] p=0.013, [b]p=0.003
Two-tailed p-value: TPS versus TS Meets Model: [c] p=0.039, [d] p=0.001

Appendix 2; Figures

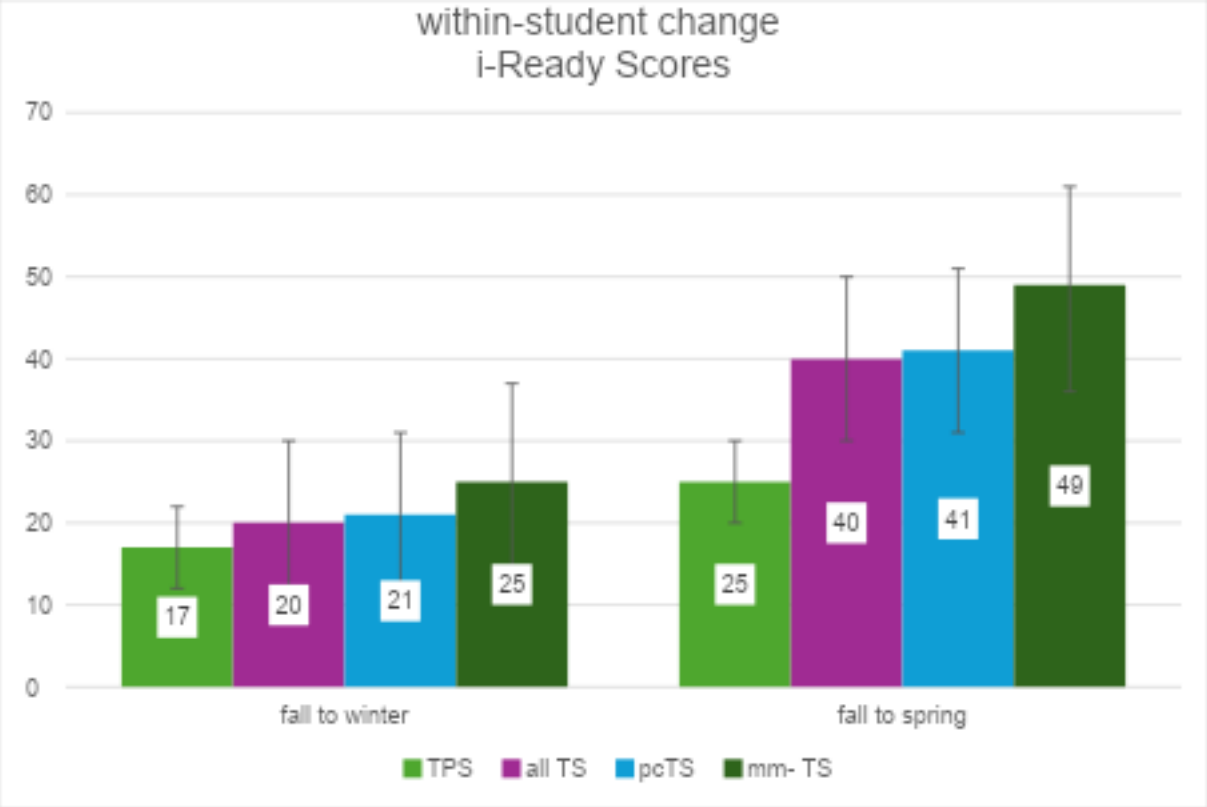
Figure 1.
i-Ready Scores by subgroup



TPS:	Toledo Public School students* not involved inTutorSmart
all TS:	ANY TutorSmart participation
pcTS:	protocol compliant (>30hrs dosage OR >66% engagement)
mmTS:	meets model (>30hrs dosage AND >66% engagement)

**all students in the study are enrolled in TPS and on a Reading Improvement and Monitoring Plan*

Figure 2.
Within Student Improvement by subgroup.



TPS:	Toledo Public School students* not involved inTutorSmart
all TS:	ANY TutorSmart participation
pcTS:	protocol compliant (>30hrs dosage OR >66% engagement)
mmTS:	meets model (>30hrs dosage AND >66% engagement)

**all students in the study are enrolled in TPS and on a Reading Improvement and Monitoring Plan*