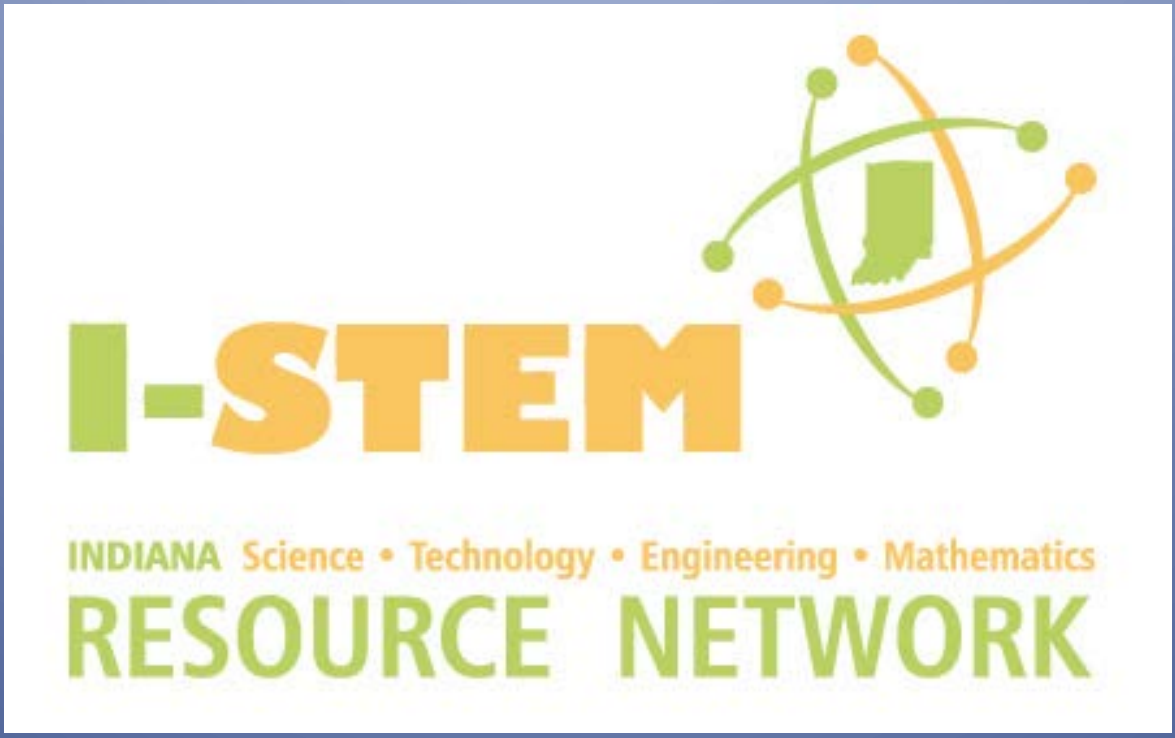
Report 2015-03



Indiana Science Initiative Update: The Impact of the Indiana Science Initiative on Students' Science, Mathematics and English Language Arts Knowledge Evaluations

October 2015

The I-STEM vision is for Indiana to be a national leader in student achievement and to demonstratively improve college and career readiness in the STEM disciplines.

The evaluation of the Indiana Science Initiative has analyzed data related to multiple areas of student level outcomes and the impact of professional development on these student level outcomes. These student level outcome data were collected from the Acuity assessments and the ISTEP+ and the findings are detailed below.

Acuity

Acuity assessments, developed by CTB McGraw Hill, provide diagnostic measures for grade 3-8 students in all subject areas. The assessments are aligned to the Indiana standards and designed to support an educator's ability to inform instruction at the student, class, school, and corporation level. While not explicitly tied to the science curricula used by the ISI teachers, the ISI team and TERC evaluators felt this instrument would adequately measure science content knowledge gains as a result of ISI. Thus, the analysis of these data measured:

- increases in student achievement for ISI students on Acuity physical, earth/space, life science, and technology/engineering assessments.
- the extent to which ISI supports equity in these areas of study.

Data: There is no state requirement for schools to use Acuity assessments. Therefore, the study relied on available 2012-13 Acuity data for grades 3-8 that could be secured from the Indiana Department of Education. Data were organized within four science standards, each of which covered a science discipline: physical, earth/space, life, and technology/engineering. The ISI Research Associate merged teacher and student data, identified which teacher IDs were associated with ISI, and stripped all names from the set (to adhere to human subjects protections). In order to examine the impact of ISI, TERC evaluators employed a quasi-experimental design using nearest neighbor propensity score matching. Matching was done at the school level, and was based on student demographics and school size. This process resulted in a sample of 42 ISI classrooms and 41 comparison classrooms that did not use science kits. On average, there were 19 students per classroom.

Analysis: For each of the four science standards, we developed hierarchical linear models (HLMs) to account for the fact that students were nested within teachers' classrooms. Student-level characteristics were entered into the model. These included:

Race (coded as white and nonwhite)

Gender (coded as male or female)

Free/Reduced Lunch (coded as no FRL and qualifies for FRL)

Exceptionality/special needs (coded as no exceptionality and identified exceptionality)

From these models, we determined which student-level characteristics were associated with Acuity scores, and also, we determined which random effects were statistically significant. Only student characteristics and random effects that reached statistical significance were retained in subsequent models.

In a second step, teacher-level characteristics were added to the model. These included:

ISI (coded as ISI teachers or comparison teachers)

Grade (coded as elementary school or middle school)

Study Results

Initial testing

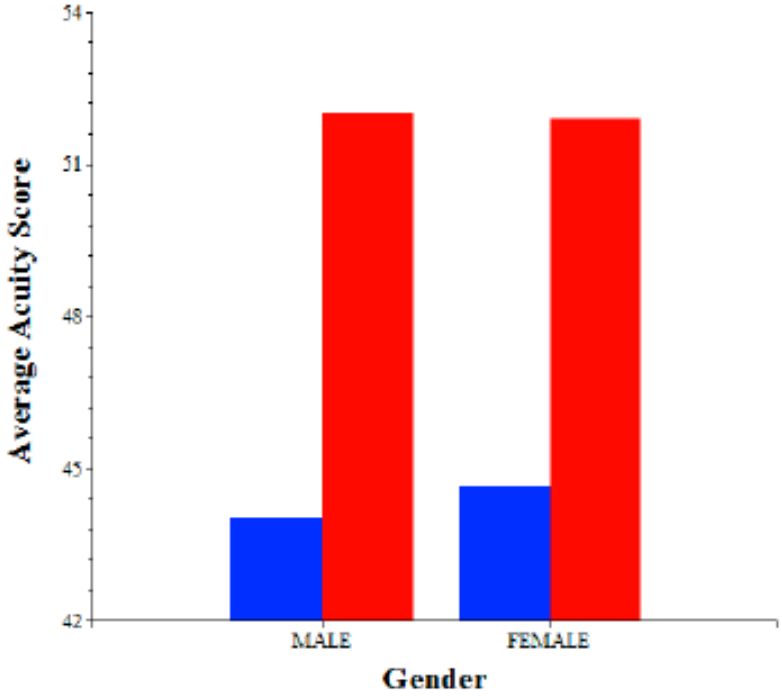
This testing was conducted to determine whether each student group needed to be analyzed separately by race, gender, FRL, and exceptionality. Similarly, we tested whether Acuity scores differed for students in elementary versus middle school grade levels.

- *student characteristics:* Although some student characteristics were associated with Acuity assessment scores, none of the random effects were significant. Because of this, we could measure the effect of ISI on all students in one model for each standard, EXCEPT for gender in standard 2 (see findings for Standard 2, below).
- *grade level:* Acuity scores didn't differ systematically by grade level. Therefore, we could measure the effect of ISI on both elementary and middle school students in one statistical model for each standard.

Findings for Standard 1: Physical Science

Elementary and middle school students in ISI classrooms had *significantly higher Acuity scores in physical science* as compared with students in non-ISI classrooms. This statistically significant positive effect ($t(59)=2.42$, $p = 0.02$) was true for boys and girls, students of all races, and for those with and without FRL status and identified exceptionalities.

Physical Science Scores by Gender (ISI is in red)



ISI professional development focusing on physical science content and instruction followed by classroom implementation of physical science kits began in the summer of 2010. Thus, many ISI teachers had several years to develop knowledge and skill as well as to increase the amount of physical science instruction in their classrooms. Findings suggest this was beneficial to student learning, as measured by the Acuity assessment.

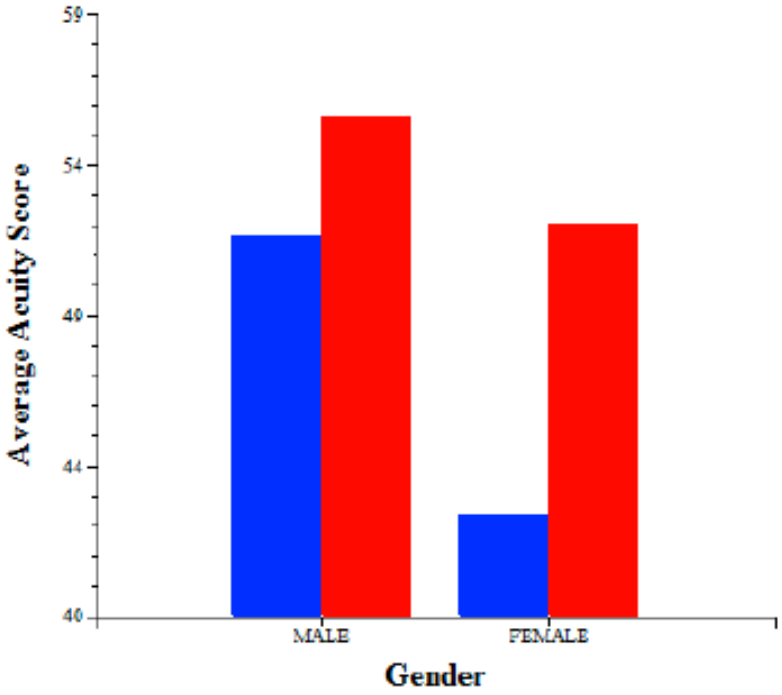
Findings for Standard 2: Earth and Space Science

*Elementary and middle school **female** students in ISI classrooms had significantly higher Acuity scores in earth and space science* as compared with female students in non-ISI classrooms.

- This positive effect ($t(32)=1.89$, $p = 0.059$) was true for female students of all races, and for those with and without FRL status and identified exceptionalities.
- While there was *not* a statistically significant difference between ISI and comparison group male students, there was a *positive data trend* (an overall higher score) for ISI elementary and middle school boys of all races, and for those with and without FRL status and identified exceptionalities ($t(32)=1.68$, $p = 0.102$).

In our analysis of Standard 2 scores, gender matters. Males in grades 3-8 in both ISI and comparison classrooms scored higher than their female counterparts. However, *being a girl in an ISI classroom mitigates this gender difference and helps to close the gender gap in earth and space science knowledge*.

Earth and Space Science Scores by Gender (ISI is in Red)



Most ISI teachers participated in earth and space science PD followed by kit implementation in 2011, two years prior to the collection of this set of Acuity data, once again giving teachers more time and opportunities to gain knowledge and implementation experience. Findings suggest this was beneficial to student learning.

Findings for Standard 3: Life Science & Standard 4: Technology and Engineering

We report on these two standards together, as results were the same for each one. *There was no statistically significant difference in Acuity scores for elementary and middle school students in ISI classrooms in life science* ($t(74)=0.26$, $p = 0.80$) *or in technology and engineering* ($t(82)=0.38$, $p = 0.70$) as compared with students in non-ISI classrooms. This was true across all student groups.

Such a finding for life science and technology/engineering was not especially surprising, given that ISI professional development and kit implementation in these two areas began just prior to the 2012-13 school year when these Acuity data were collected. Thus, there was limited time in which teachers could translate learning from ISI professional development into practice, and for many, it was the first year they instructed with an ISI life science kit or conducted investigations in technology/engineering.

Furthermore, with limited curricula available specifically for technology and engineering for the elementary and middle grades, ISI supported teachers in meeting the goals of Standard 4 via investigations in the physical, earth/space, and life science kits. As a result, there may have been less focus on or time to explore technology and the engineering design process as fully. Given students lack of familiarity with the engineering as compared with other areas of science, more emphasis on engineering may be needed.

ISTEP+

The ISTEP+ is the state level assessment given to students in grades 3 through 8 covering mathematics and English language arts (ELA). Additionally, the ISTEP+ test science at grades 4 and 6. Multiple analyses have been run to measure the impact of the ISI and its teacher professional development on student outcomes. One of these studies done by TERC compared three groups of students:

1. Students in non-ISI comparison classrooms
2. Students in ISI classrooms where teachers are using curriculum provided via the project
3. Students in ISI classrooms where teachers are using curriculum *and* have participated in 14 or more hours of professional developed offered via the Indiana Science Initiative

Hierarchical linear models (HLMs) were developed to account for the fact that students were nested within teachers' classrooms. Since ISI and comparison groups were matched on gender and race, the following demographics were controlled for at Level 1 in the model:

Free/Reduced Lunch (coded as no FRL and qualifies for FRL)

Exceptionality/special needs (coded as no exceptionality and identified exceptionality)

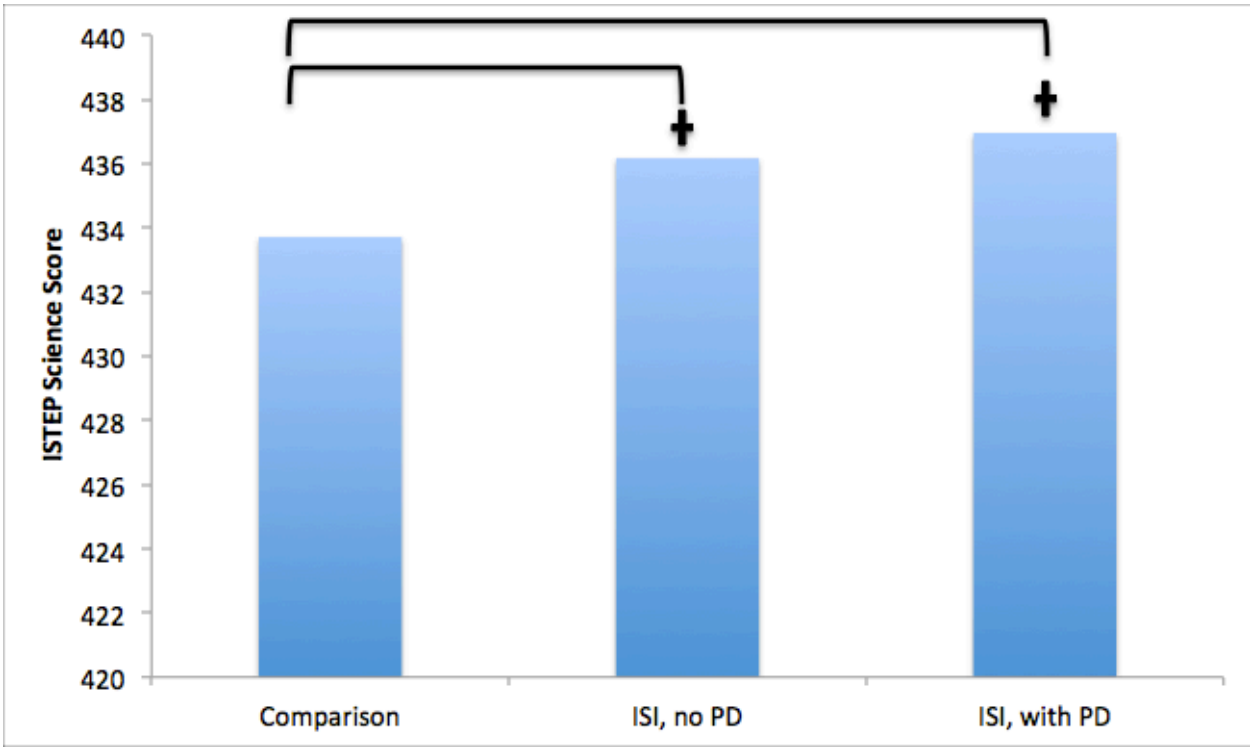
ELL Status (coded as medium to high proficiency (0) and low proficiency (1))

Gifted/Talented Status (coded as 0 not identified as gifted and 1 identified as gifted)

Teacher-level characteristics that were controlled for in the model included years of experience and grade-level taught (elementary or middle school).

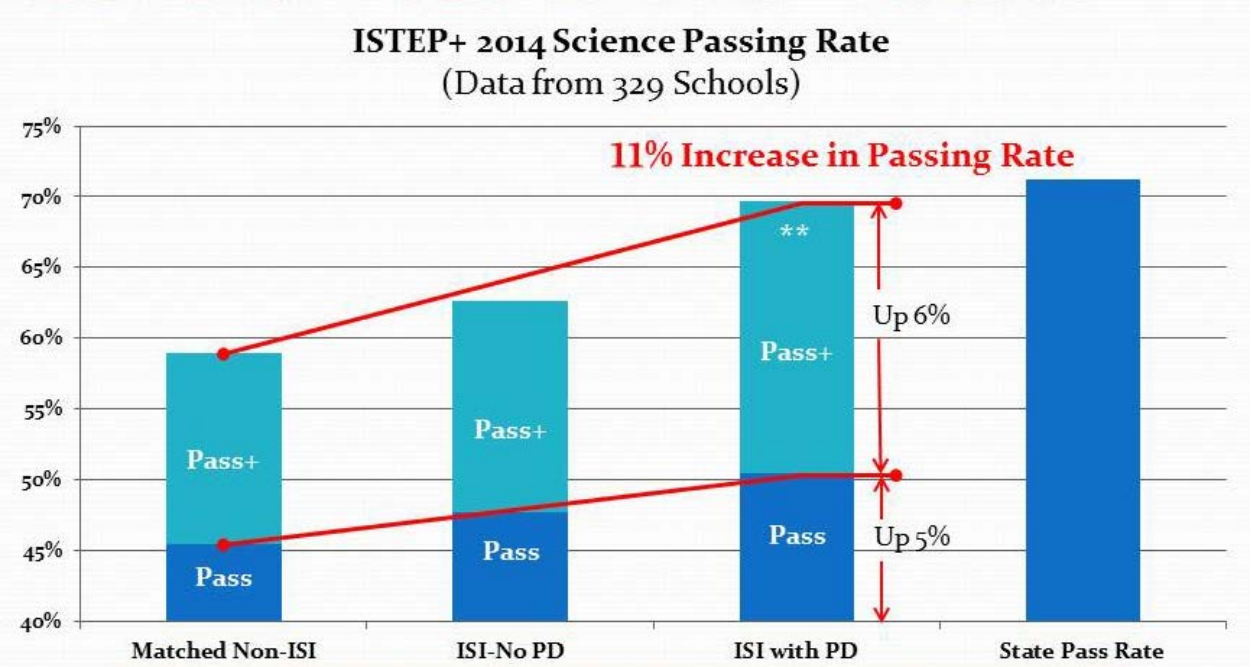
ISTEP+ Science

- Fourth grade and Sixth grade students in ISI classrooms scored higher on the science ISTEP+ assessment than comparison students, however, this positive trend was marginally significant ***(****t*(333) = 1.88, *p* = 0.06; *t*(333=1.59, *p* = 0.11***).***
- Students in classrooms with ISI teachers who received curriculum + PD scored higher on the science ISTEP+ than students of ISI teachers who did not receive PD, however this difference is not statistically significant (*p*>0.50).
- Our analysis of subgroups on the science ISTEP+, indicated that students in each group (e.g., students with FRL, exceptionality, ELL, and gifted status) were scoring lower than their non-subgroup counterparts (students without FRL status, for instance). This difference exists across all three main groups and is not affected by ISI.



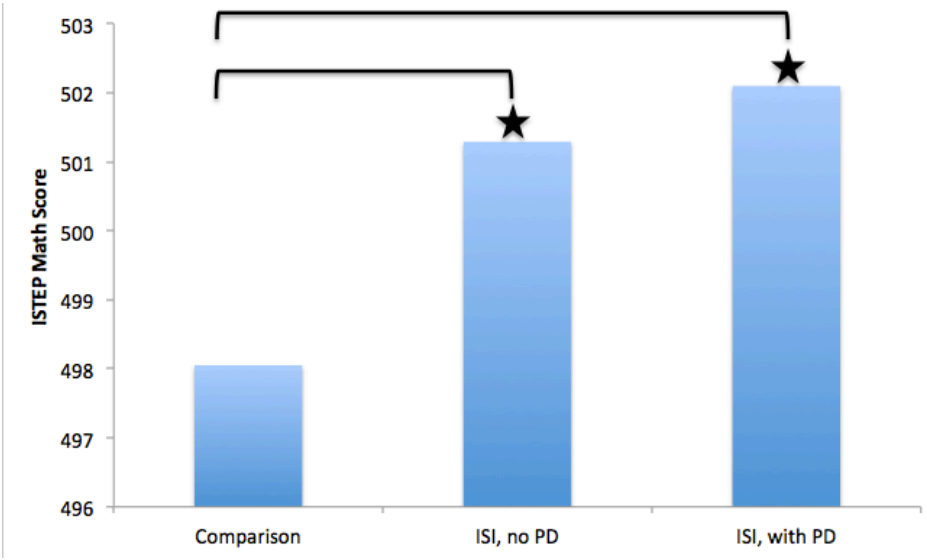
These data were also examined looking at the passing rates for each of these groups controlling for any differences between that were not accounted for in the matching process. The passing rates in science for students of teachers with ISI PD were significantly higher that the passing

rates of ISI students with teacher with no PD as well as the comparison group. While the passing rate for students of teachers with ISI PD was slightly lower than the state passing rate, this difference was not statistically significant. Additionally, the free and reduced lunch rate for these ISI students was over 10% greater than the state average. The ISI students whose teacher had ISI PD had a pass rate over 11% (6% Pass+ and 5% Pass) greater than the control group and 7% greater than ISI students whose teacher did not attend PD.

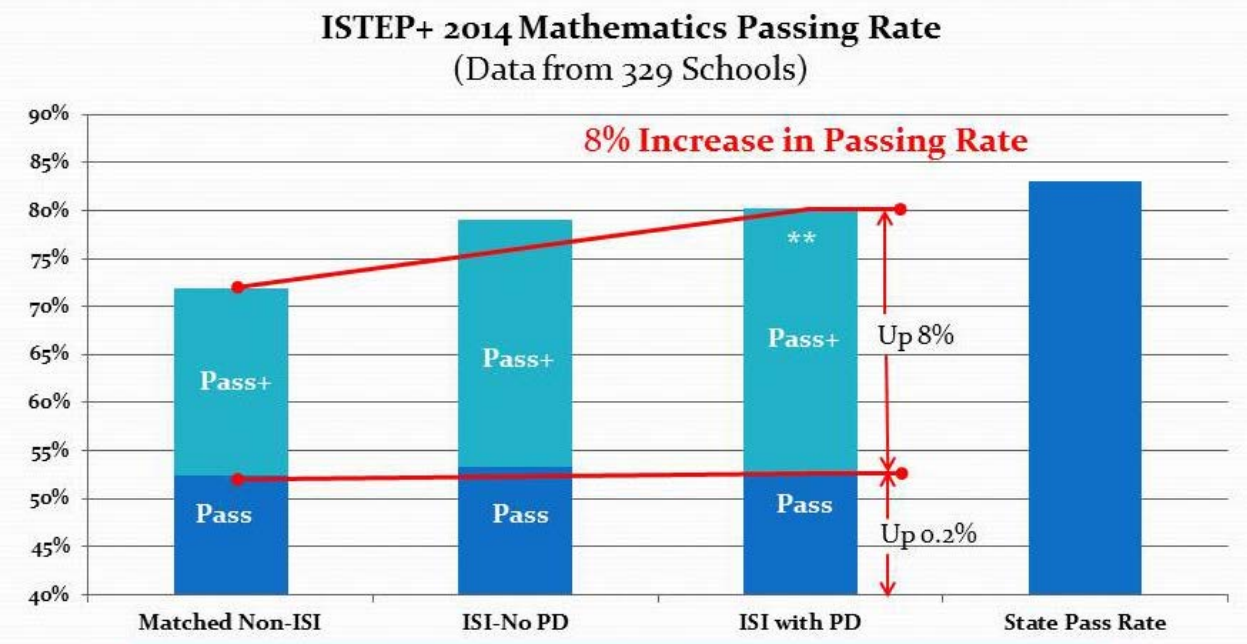


ISTEP+ Math

- ***Elementary and middle school Students in ISI classrooms had significantly higher mathematics ISTEP+ scores*** as compared with students in non-ISI classrooms ($t(874)=2.15$, $p<0.05$; $t(874)=2.17$, $p<0.05$).
- Students in classrooms with ISI teachers who received curriculum + PD scored higher on the mathematics ISTEP+ than students of ISI teachers who did not receive PD, however this difference was not statistically significant ($p>0.50$).
- Our analysis of subgroups on the mathematics ISTEP+, indicated that students in each group (e.g., students with FRL, exceptionality, ELL, and gifted status) were scoring lower than their non-subgroup counterparts (students without FRL status, for instance). This difference exists across all three main groups and is not affected by ISI.

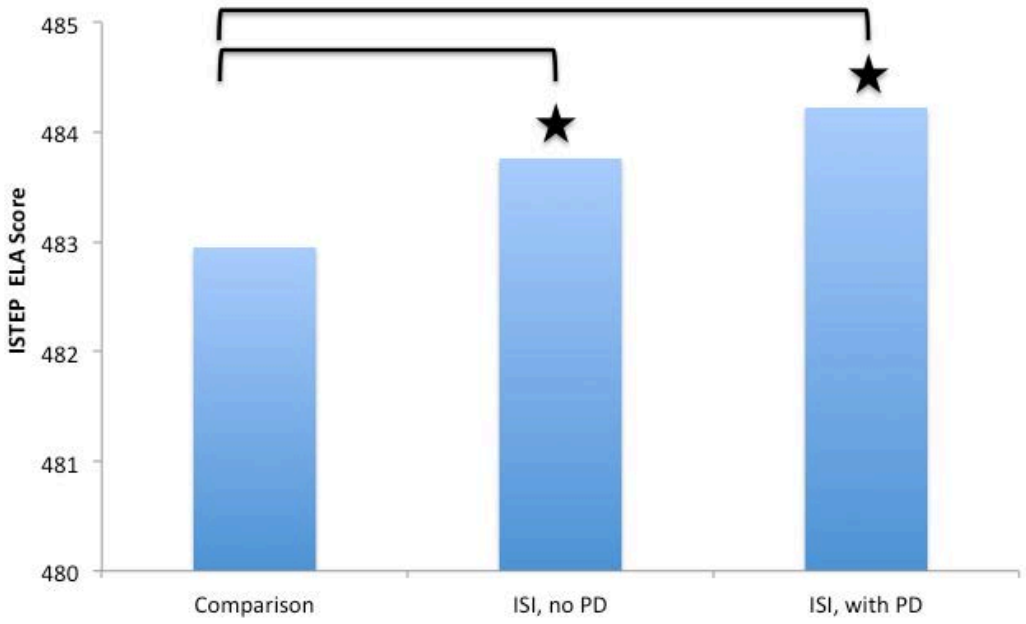


As with science, the passing rates of these same student groups were examined. ISI students whose teachers had ISI PD had a statistically significant higher pass rate than the comparison group in mathematics ($p<.01$). The ISI students whose teacher had ISI PD had a pass rate over 8% (8% Pass+, .02% Pass) greater than the control group.

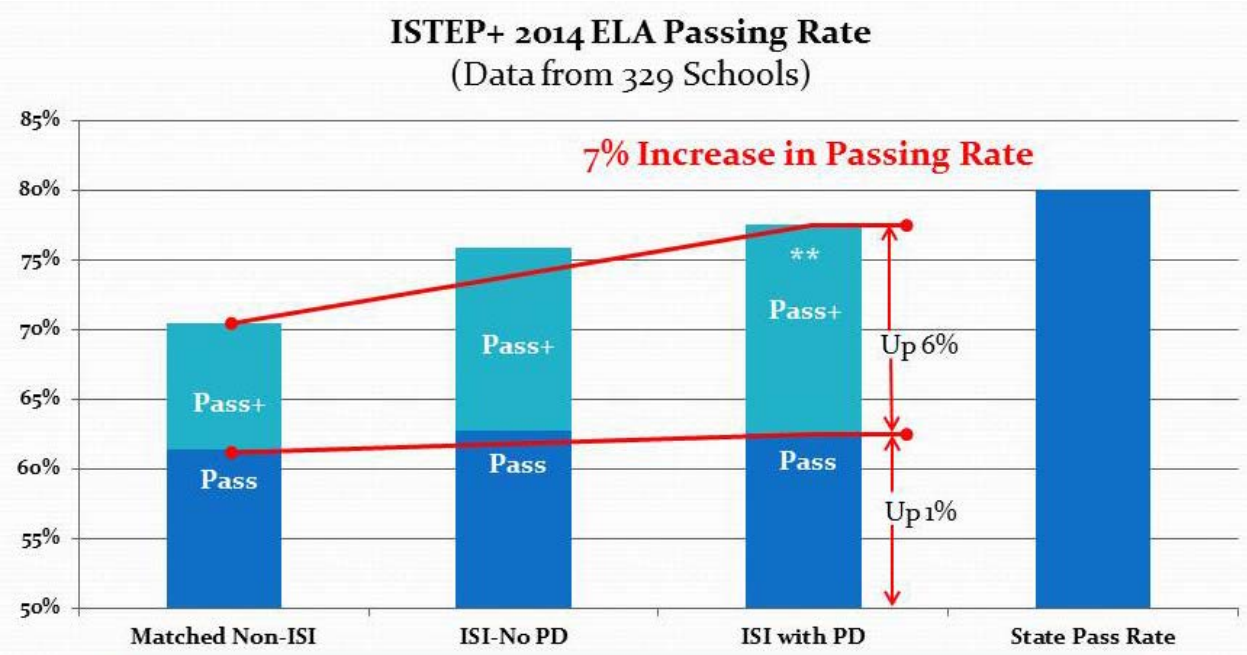


ISTEP+ English Language Arts (ELA)

- Elementary and middle school Students in ISI classrooms had significantly higher ELA ISTEP+ scores as compared with students in non-ISI classrooms ($t(874)=2.18$, $p<0.05$; $t(874)=1.80$, $p = 0.07$).
- Students in classrooms with ISI teachers who received curriculum + PD scored higher than students of ISI teachers who did not receive PD, however this difference is not statistically significant ($p>0.50$).
- Our analysis of subgroups on the ELA ISTEP+ indicated that students in each group (e.g., students with FRL, exceptionality, ELL, and gifted status) were scoring lower than their non-subgroup counterparts (students without FRL status, for instance). This difference exists across all three main groups and is not affected by ISI.



As with mathematics and science, the passing rates of these same student groups were examined for ELA. The passing rates in ELA for students of teachers with ISI PD were significantly higher that the passing rates of the comparison group. The ISI students whose teacher had ISI PD had a pass rate over 7% (6% for Pass+, 1% for Pass) greater than the control group.



The Impact of ISI Professional Development on Student ISTEP+ Performance

To further investigate the impact of ISI PD on student performance, ISI teachers were clustered by the *amount* of professional development (PD) they received. Categories were determined in consultation with I-STEM researchers, and they are listed below.

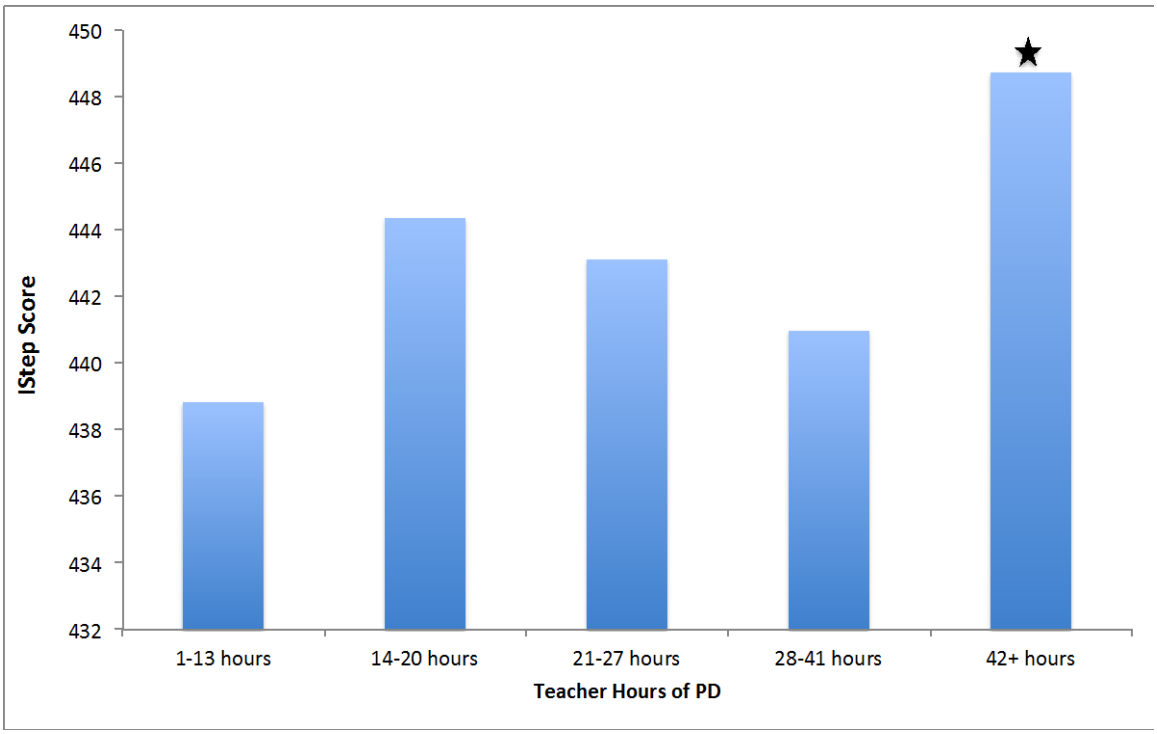
- Very low = 1-13 hours
- Low = 14-20 hours
- Moderate = 21-27 hours
- High = 28-41 hours
- Very high = 42+ hours

It is important to note that these designated labels (very low to very high) were based on the amount of professional development (PD) available to teachers within this data set. Thus, the "very high" label is not an indication that 42+ hours is considered very high within the field of education. Instead, it is simply a very high number of ISI PD hours for this group of teachers. In fact, a recent report states that "The duration of professional development must be significant and ongoing to allow time for teachers to learn a new strategy and grapple with the implementation problem" (Gulamhussein, 2013), and underscores earlier studies suggesting that teachers need more than 50 hours of instruction, practice, and coaching.

Below are results that compare whether different levels of ISI PD impact student performance on Science, Mathematics, and ELA ISTEP+ scores. Since there was no significant difference

between the performance of students in elementary and middle school classrooms, the findings apply to *all* students in ISI classrooms.

Science



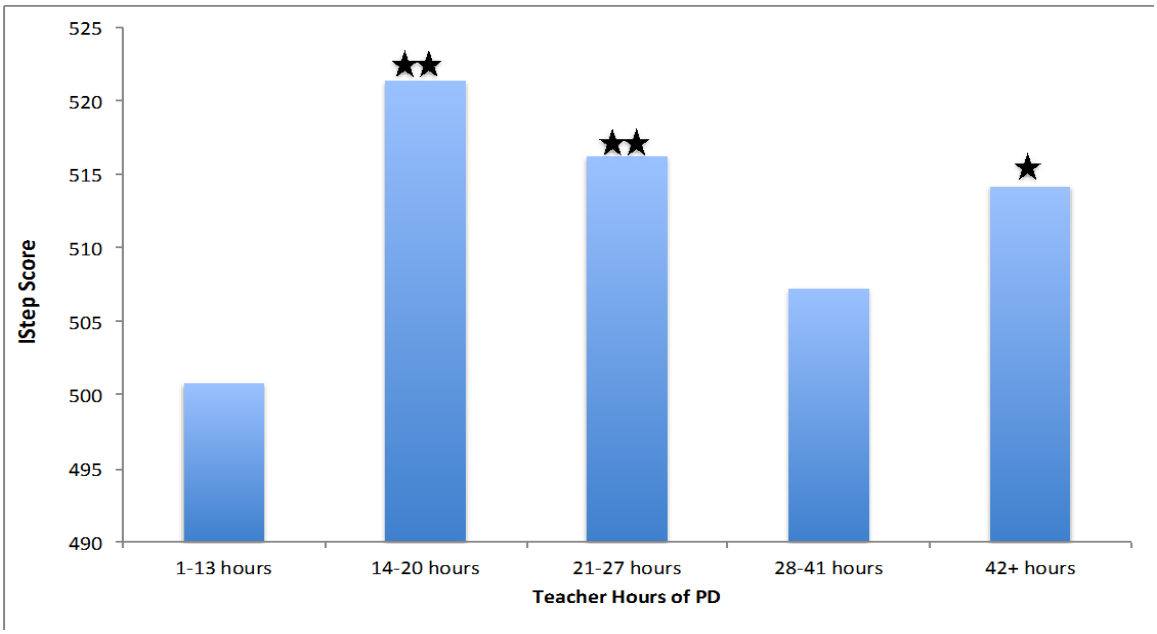
** indicates significance at the p=<.05 level controlling for school level (middle or elementary).

* indicates a data trend at the p=<.1 level controlling for school level (middle or elementary).

•Students in classrooms with ISI teachers who received 42+ hours of PD scored higher on the *science* ISTEP+ assessment than students with ISI teachers who received very low levels of PD. This positive data trend was near but not at the <.05 level of significance ($t(307) = 1.83$, $p<0.07$).

Comment The ISTEP+ science assessment includes questions pertaining to physical, earth-space, and life science. Preparing to use science kits in all of these areas requires more hours of PD. Therefore, until teachers have completed 42+ hours of training, they may not be using kit curriculum for all science domains and/or they may not be as prepared to teach all of the topics as well.

Mathematics



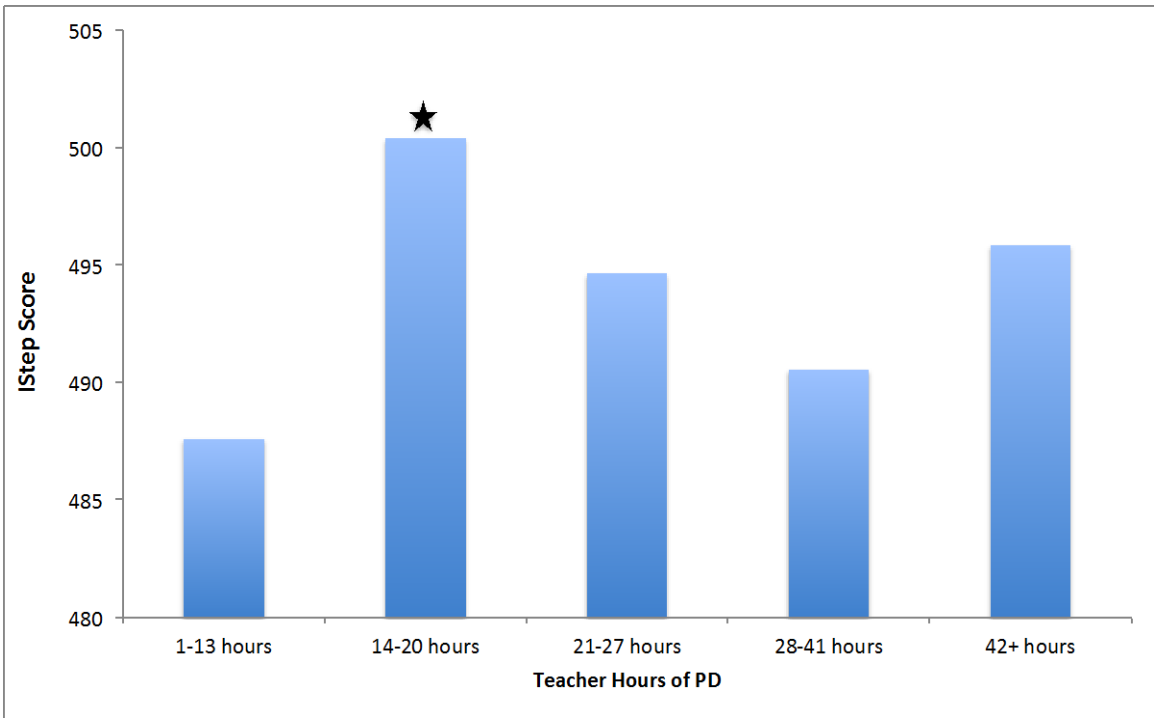
** indicates significance at the p=<.05 level controlling for school level (middle or elementary).

* indicates a data trend at the p=<.1 level controlling for school level (middle or elementary).

- Students in classrooms with ISI teachers who received 14-20 hours of PD scored significantly higher ($t(307) = 2.20$, $p<0.05$) on the *mathematics* ISTEP+ assessment than students with ISI teachers who received very low levels of PD.
- Students in classrooms with ISI teachers who received 21-27 hours of PD scored significantly higher ($t(307) = 2.22$, $p<0.05$) on the *mathematics* ISTEP+ assessment than students with ISI teachers who received very low levels of PD.
- Students in classrooms with ISI teachers who received 42+ hours of PD scored higher on the *mathematics* ISTEP+ assessment than students with ISI teachers who received very low levels of PD. This positive data trend was near but not at the <.05 level of significance ($t(307) = 1.85$, $p<0.07$).

Comment We have no statistical explanation for why teachers with 28-41 hours of PD did not score significantly higher (even at the <.1 level) than the very low level teachers. We examined residuals to see whether there were outliers (whether a few classrooms within this group scored extremely lowed and pulled the group average down). This was *not* the case.

English Language Arts



** indicates significance at the p=<.05 level controlling for school level (middle or elementary).

* indicates a data trend at the p=<.1 level controlling for school level (middle or elementary).

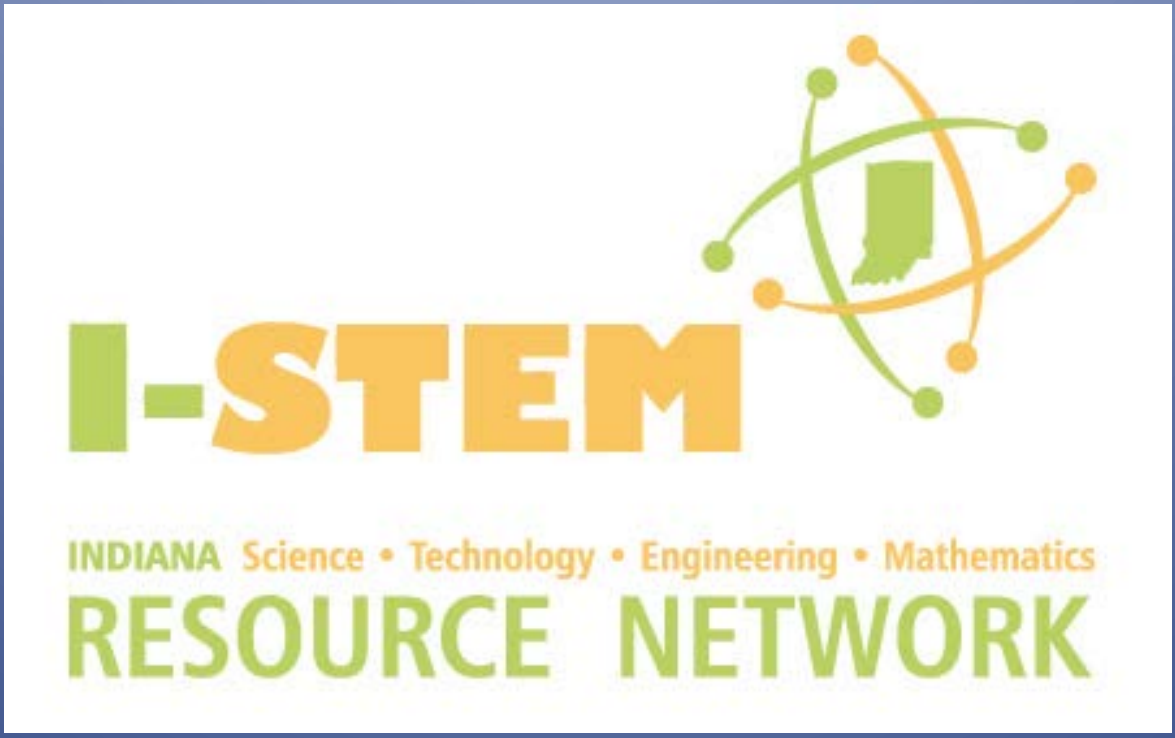
• Students in classrooms with ISI teachers who received 14-20 hours of PD scored higher on the ELA ISTEP+ assessment than students with ISI teachers who received very low levels of PD. This positive data trend was near but not at the <.05 level of significance ($t(307) = 1.89$, $p<0.06$).

Comment While science notebooking and making meaning conferences support verbal and written language practices generally, their primary purpose is to strengthen scientific thinking. Thus, we might expect minimal impact simply from receiving more PD. In addition, our analyses of ISI sustainability (ongoing) suggest that teachers receive a great deal of PD in ELA as compared with science. Therefore, all ISI teachers are likely to have similarly high levels of knowledge and skill related to language arts, making it difficult to detect any added benefit of science notebooking on ELA.

References

Gulamhussein, Allison, "Teaching the Teachers," Center for Public Education, 2013.

I-STEM Resource Network is supported by the Lilly Endowment, the Lilly Foundation, BioCrossroads, the Indiana Department of Education, The Indiana Commission for Higher Education, and Purdue University.



For more information, please contact:

Paul J. Ainslie, Ph.D.
Managing Director, I-STEM Resource Network
Purdue University
Rm 152 MMDC
700 Ahlers Dr.
West Lafayette, IN 47907
Office: 765-494-0557 Mobile: 317-531-7301