# An Initial Evaluation of the Iowa DE Project to Enhance Students' Authentic Intellectual Work

# Fall 2007 to Fall 2011



May 2, 2012

This report was developed by the Iowa Department of Education, which commissioned Elizabeth Weinstein of Weinstein and Associates, Inc. to conduct the case studies and analyze focus group transcripts. Weinstein and Associates, Inc., is a training and consulting company based in Des Moines. M. Bruce King assisted the DE in the design of the task analysis and Fred M. Newmann and King consulted with the Department in the design of the student achievement study and reporting of results. King is professor of Educational Leadership and Policy Analysis at the University of Wisconsin-Madison. Newmann is professor-emeritus of Curriculum and Instruction at the University of Wisconsin-Madison.

#### **Executive Summary**

# **Project Overview**

In September 2007, the Iowa Department of Education (DE) began the Authentic Intellectual Work (AIW) project of professional development with high school administrator and teacher teams to enhance the intellectual quality of student work in Iowa schools. AIW is a research-based framework focusing, not on specific teaching techniques, but on intellectual demands that teachers present to students to prepare them to successfully respond to the challenges of the modern world, post-secondary education, and the workplace.

## **Authentic Intellectual Work Definition and Indicators**

AIW is defined by three criteria: *construction of knowledge* through *disciplined inquiry* to produce discourse, products, and performance that have *value beyond school*. The AIW framework establishes criteria for teaching that

- Maximize expectations of intellectual challenge and rigor for all students,
- Increase student interest in academic work,
- Support teachers in teaching for in-depth understanding rather than superficial coverage of material, and
- Provide a common conception of student intellectual work that promotes professional community among teachers of different grade levels and subjects.

These criteria are applied to student work, instruction, and assignments through standards as identified in the chart below.

# Criteria and Standards for Authentic Pedagogy and Student Work

Criteria for Authentic Intellectual Work	Instruction	Assessment Tasks	Student Work
Construction of Knowledge	Higher Order Thinking	Construction of Knowledge	Construction of Knowledge
Disciplined Inquiry	Deep Knowledge and Student Understanding Substantive Conversation	Elaborated Communication	Elaborated Communication
Value Beyond School	Value Beyond School	Value Beyond School	

These criteria and standards were derived from research conducted by Fred M. Newmann and colleagues at the University of Wisconsin-Madison and elsewhere from 1990 to 2003. Through a number of studies, researchers found that the achievement of students who experienced high levels of authentic instruction and assessment exceeded the achievement of their peers who received lower levels (the studies are summarized in Newmann, King and Carmichael, 2007). The findings were consistent in grades 3 through 12; in language arts, mathematics, science, and social studies; in schools from urban, suburban, and rural settings; and with diverse groups of students.

Using this research as a foundation, Newmann, Dana Carmichael, and Bruce King assisted the DE in designing a professional development project that focused on improving teachers' ability to design instruction and assessments to increase student authentic intellectual work. Schools apply voluntarily to the project and teams of teachers and administrators participate in:

- beginning-of-the-year kick-off institutes to introduce teachers and administrators to AIW critera and standards,
- regular on-site team meetings to critique and improve teachers' assignments, assessments, and lessons,
- periodic on-site coaching by external coaches trained in AIW, and
- mid-year institutes where teams from different schools continue their professional development through subject area and grade alike workshops.

The rapid growth of the AIW program, stemming largely from "word of mouth advertising" from one teacher and administrator to another, is one indicator that lowa educators value this approach to professional development. During its initial year, teams from nine schools included 76 teachers who participated in the program. During 2008-09 an additional 165 teachers joined their peers in implementing AIW at those nine schools, and teams from eleven other high schools and one middle school began professional development in AIW, bringing the total number of teachers participating to 336. Expansion in schools already practicing AIW, into other schools within those districts, and the addition of ten new schools brought the total of teachers participating in AIW during the 2009-10 school year to 1102. In 2010-2011, 22 more schools joined the AIW statewide initiative, with more than 2000 teachers participating in AIW professional development. In 2011-2012, 35 new schools and districts joined the project. With 106 schools (including expansion schools within a district already doing AIW) having participated by spring 2012, this makes AIW the largest Department supported professional development initiative in the state.

An important aspect of the project is designed to build the state's internal capacity to support and sustain AIW professional development in the schools. Each year, consultants from several AEAs and the DE become new AIW coaches through summer institutes and ongoing mentoring. At present, 13 AEA and DE consultants serve as AIW coaches, and 9 new coaches began their year of training in summer 2011. The intent is to develop a cadre of coaches within the AEA system to provide this service to

schools and districts. In 2011-2012, the DE also began supporting the development of 24 AIW teachers and administrators as local, school-based and district-based AIW coaches.

# **Initial Evaluation Design and Summary of Results**

To this point, project resources have concentrated on program development, but an initial evaluation was undertaken through four studies with the following objectives and methods:

- To understand the professional development process within schools, case studies of implementation of the AIW framework in four schools have been completed.
- To understand administrators' reactions to the program, two focus group discussions were held with administrators and curriculum directors and the results summarized.
- To understand the extent teachers' intellectual demands were affected by professional development emphasizing feedback from colleagues on the quality of their assignments (tasks) for students, an analysis compared their tasks before and after receiving feedback.
- To understand whether participation in the AIW project is associated with higher achievement on the Iowa Test of Basic Skills and Iowa Test of Educational Development, test scores in AIW schools were compared with scores of students in matched schools not in the program.

The evaluation to date indicates consistent positive results for teachers and students, as well as challenges that need further work. The results are elaborated for each of the studies in the four sections of the full report following this executive summary.

# Impact on Teachers' Practice

Focus groups and case studies describe the changing nature of instruction from the teacher-as-deliverer of facts to teacher-as-facilitator of student thinking, in-depth understanding, and skill development that is meaningful and valuable. The quality of classroom discussions has been at a much deeper and more thoughtful level. Expectations for students have been increased and curriculum is now more closely connected to students' lives, making lessons more challenging and, simultaneously, more meaningful. Because students are more engaged, they are more persistent in problem solving. The review of teachers' tasks show that high school teachers who participate in AIW professional development are able to implement assessment tasks that scored significantly higher in the standards for authenticity. Effect sizes ranged from medium in science and social studies to large in mathematics.

# Change in Professional Culture and Leadership

Administrators referred to the level of collaboration among teachers as "unprecedented." Using common protocols and criteria, teams of teachers within and across disciplines meet to improve their practice. Teachers examine their practice through the lens of the AIW framework, individually and collectively asking questions such as, "Will this lesson provoke students' higher order thinking and substantive conversation?" or "Does this unit lead students to apply and understand knowledge in contexts beyond school?" or "Will this assessment task require students to show an in-depth

understanding of an important concept?" AIW teachers value the opportunities AIW professional development provides to make their instruction better. AIW schools also experience more sustained focus for their PD. AIW has improved the collaborative spirit between administrators and teachers, according to those interviewed in focus groups and case studies. Because administrators are part of the learning team, they find themselves giving teachers more relevant feedback. Also, AIW provides teachers with more leadership opportunities.

## Student Achievement

Students in AIW schools across grade levels and subjects (reading, mathematics, science, social studies) usually scored higher on the ITBS/ITED than students in non-AIW schools and had higher percentages of students scoring proficient (i.e. the 41<sup>st</sup> percentile and above). For grades 4, 8, and 11 – the grades for which lowa schools must report annually, AIW students scored significantly higher in 8 of the twelve comparisons (3 grades x 4 subjects) and AIW had higher percentages proficient in all 12 comparisons. The percentile advantage to AIW students was 5 points or higher in 8 of the 12 comparisons. The results across all grades 3-11 were similar. Of the four subjects, AIW students posted the most consistently higher scores in mathematics and showed consistently smaller differences in social studies.

While this evaluation offers initial positive indicators of the program's success, we hope that in the future, it will be possible to gather more comprehensive quantitative data that will show the extent to which professional development affects individual teachers' instruction and assessments which in turn then affects their students' achievement.

# **Summary of Case Studies**

#### Introduction

Case studies of Authentic Intellectual Work (AIW) in the Iowa project included four high schools that have been practicing this professional development model for the last five years. The schools, AHST Secondary School, Cedar Falls High School, Spencer High and Middle Schools and Valley High School were visited in April, 2011. The visits included observations of AIW team meetings and interviews with administrators and teachers. Of the four schools, AHST, serving four rural communities, was the smallest school with 150 students seventh through twelfth grade. Valley High School in West Des Moines was the largest with 1,900 tenth through twelfth grade students.

Each of the four schools used a phased-in approach to implementing AIW. A core group of teachers was initially trained in AIW, piloted the system, and then additional groups were added until all staff members are now included in AIW.

Five major themes were identified in the case studies.

- Administrators and teachers interviewed said they believe that student achievement and learning has improved because of AIW.
- Teachers and administrators interviewed said that student engagement has improved because of AIW.
- AIW teams enhance teacher collaboration by using a common language to score and discuss instruction, assessment tasks or assignments, and student work. This has resulted in an improved professional culture within the schools implementing AIW.
- Teachers involved on AIW teams have improved their practice.
- A number of important challenges to AIW implementation should be addressed.

#### **Successes**

The four schools cited multiple examples of how AIW has been advantageous to students, teachers, and administrators. AIW has provided a framework for regular communication among teachers. According to many of those interviewed, teachers are improving instruction through collaboration with their colleagues. In most instances, a high level of trust has been established in the scoring teams. Members give honest feedback and receive comments from their colleagues with grace. In addition, the perspective of teachers, who are in different disciplines, has helped teachers better identify potential student questions and has enriched their curriculum and instruction. As one teacher from Cedar Falls stated, "I think with a 'science mind', as do some of my students. It helps me to see how a student with an 'English language mind' thinks by hearing an English teacher share her view point."

Through the AIW scoring process, tasks, student work and instruction have improved. According to a teacher at AHST, even a simple change of wording in how a question is posed can have a huge impact on

student learning. Teachers, because of AIW, tend to facilitate the acquisition of knowledge, understanding, and application, as opposed to the former model of teacher as a conveyor of facts. At Valley High School, staff reported that teachers have raised expectations for students and now challenge them to synthesize and analyze information to solve problems. At Spencer High School, a science teacher shared that she has changed how she teaches. Instead of sharing many facts about the subject, she begins her lesson by asking students what they know about the topic. She then teaches the lesson, and at the conclusion of the class, asks the students if their thinking about the topic has changed and if so, how.

Teachers communicate with each other at a more professional level, according to a teacher at Spencer High School, and now also use a common language when discussing student learning opportunities. Construction of knowledge, substantive conversation and higher order thinking are all terms that are regularly used when teachers interact with each other.

Students at AHST are also familiar with the AIW language. These students have been "let in on the secret", and are so familiar with the terms that they challenge their teachers by asking, "Is that a HOT question?" At AHST, professional development sessions for students have been held twice a year to explain AIW's purpose and process.

As a result of AIW, the four schools reported other significant changes in student behavior. Students are more engaged in learning. According to teachers at Valley High School, students are expected to use higher order thinking skills, where they are not simply absorbing facts, but seeking meaning and making connections with other pieces of information. In addition, linking curriculum to student life beyond school has made it more relevant and of greater value for students.

Since AIW has been implemented, because students are more engaged, they contribute more to discussions than in the past. "Students often bring something of interest, which I would never have thought of," said one teacher at Cedar Falls High School. Another teacher at Spencer High School stated that, by teachers being open to student input to discussion, it shows "that we value what they have to say. There has been a shift from teacher focus to student focus."

Another success evident at AIW schools is that is inclusive. Teachers and administrators alike stated that AIW has created a framework that involves all staff who collaborate on a regular basis to improve the way they work with students. In addition, school personnel believe that AIW is an initiative that can be sustained over time and, as Principal Mueting at Spencer High School said, "It is here to stay."

## Scoring

In two of the four schools visited, scoring teams are organized across disciplines. At Valley High School, scoring teams are composed of five teachers from the same discipline. At Spencer High School, during the first two years of AIW, teachers were organized in multi-discipline teams for scoring. Now, however, the school has switched to teams composed of teachers in the same discipline.

The major advantage of the multi-disciplinary approach is that tasks, student work and instruction can be viewed from the perspective of the non-expert. Consequently, the questions asked by scoring team members could be similar to those that students want to ask, but might not express. Also, multi-discipline teams allow for linking curricula across disciplines, so that students can benefit from a more holistic approach to learning. On the other hand, a disadvantage of multi-discipline teams is that, at times, teachers are unable to give appropriate feedback because they do not know the fundamentals of the subject. For example, at Cedar Falls High School, a physics teacher mentioned that a teacher of physical education might believe that student work in physics requires higher order thinking skills when, in fact, it requires simple rote learning. The teacher is unfamiliar with the language and concepts and, consequently, cannot gauge the level of complexity of the work. Another example that was shared at Cedar Falls was of a teacher whose scoring team had not been very helpful. The team is composed of two special education teachers and several other members, who do not teach core subjects. Although the members of the scoring team are well intentioned, often they are unable to provide helpful feedback to the teacher who teaches core classes.

Organizing scoring teams of teachers in the same discipline does have an advantage. No time needs to be spent explaining terms, concepts, or models because members are subject matter experts. One of the disadvantages of same-discipline teams, of course, is that the variety of viewpoints is lacking. At Cedar Falls High School, science department faculty now score student work in a same-discipline team, and this has proven to be a beneficial approach because no explanation of the content is necessary for the scoring process. At Spencer High School, the intent for 2011-2012 is to have teachers choose whether they will be in a same-discipline or multi-discipline team.

At each of the schools, scoring protocols have been established. Teachers select which item they bring to their team to be scored. At Valley High School, for example, the Rules of Thumb guidelines are used to help teachers select a task, student work, or instruction. They select an item to be scored that they plan to use again; believe could be improved; and are open to receiving feedback. At Cedar Falls High School, an AIW Task Guide is utilized. This tool helps teachers plan the focus of their request to their team for scoring and feedback.

In the scoring teams at each of the schools, teachers present the item to be scored, explain its purpose and how the task, student work, or instruction is organized. The teacher also states the AIW standard(s) he/she would like to strengthen in the item. When instruction was being scored at Spencer High School, a Parking Lot sheet was utilized. On a large sheet of paper, members of the scoring team listed factors that could have influenced their scoring of the instruction. Too much noise in the room, a classroom visitor, and technical difficulties with the video-camera are examples of incidents that could have impacted the scoring.

Members score the presented item and give a rationale for their score. A discussion about possible improvements follows. At each of the schools, a record of the scoring meeting is kept and the document is entered into the school's shared drive and/or forwarded to the school principal.

Scoring instruction appears to be more complex than scoring tasks or student work because a greater degree of organization is required and more obstacles have to be addressed. At all four schools, teachers scored instruction after they had first scored tasks and student work. At Cedar Falls High School, the preferred method for scoring instruction is by observation in the classroom. The observation can be organized at a time when several teachers have a preparation period, or by using substitute teachers to cover the classes while the observation and scoring is taking place. However, because of time constraints and limited resources, scoring instruction is most often accomplished by videotaping the lesson being scored. At AHST, videotaping is the primary method for scoring instruction because a small number of staff members make flexibility in scheduling difficult. Videotaping has its own challenges, according to teachers at the four schools visited. These include the following: equipment that fails to give an accurate visual of the classroom or delivers poor audio; the video only records a portion of what is occurring in the classroom; multiple conversations or activities in the room cause confusion; and viewing only five or ten minutes of instruction cannot truly reflect the whole teaching session.

Scoring is an evolutionary process, as one teacher described so aptly. There are many factors that impact a scoring session. A new AIW team operates differently than a more mature team. At AHST, a teacher recalled that during the first year of scoring everyone was very uncomfortable bringing work to be scored. According to a team member at Valley High School, new teachers tend to be enthusiastic about sharing their items. However, mention was also made of an experienced teacher who was admired for "humbly receiving feedback from his peers."

## Leadership

It was apparent from the school visits that for AIW to become a reality there have to be leaders at the helm who strongly believe AIW can benefit teachers and students. Leaders must be convinced that AIW will result in improved student scores and a school climate that is conducive to learning. When there is a change in leadership, as happened at AHST, this can cause an interruption in support for AIW. The new leader has to become fully cognizant of the benefits that can be gained from AIW, and realize that there is strong teacher support for this professional development framework. With Dr. Poole's retirement from Valley High School, it might prove challenging to maintain AIW at its current levels. However, the advantage of changes in leadership is that it provides a test of the value and sustainability of AIW. The merits of AIW will help to ensure its continuation.

# Challenges

Implementing AIW is not easy. It takes time for training and for practicing the principles of AIW. In addition, it requires teachers to be open to giving and receiving feedback, which can be an intimidating experience, especially early on. It appears that AIW may be more applicable to certain disciplines than others. At Valley High School, one of the math teachers stated that AIW framework is not easy to use because "there are so many fundamental skills that need to be taught first." For example, the math teacher contended that elaborated communication is a natural fit for language arts but it is more of a challenge to include when teaching math. An art teacher at Cedar Falls High School stated that AIW was

a "real struggle" because there was no rubric for her subject, and the social studies rubric that was used for scoring "didn't fit". With the addition of the "Other than Core" rubric, this has helped.

There appear to be challenges, according to the teachers and administrators interviewed. At Cedar Falls High School, a teacher mentioned being bored with AIW and that the process is becoming too mechanical and "easy". One of the challenges with AIW is to determine how to keep the momentum going; how to sustain the high level of enthusiasm and commitment to practicing AIW.

Another challenge with AIW is the lack of data that pinpoints AIW as positively impacting student learning. As one teacher at Cedar Falls High School stated, "Anecdotally, we can see positive changes, but we need some hard data to prove it." Reliable and valid data is needed to convince school boards that AIW makes a strong impact on student learning and is, consequently, worth funding.

## **Future**

Both AHST Middle and High School and Valley High School are experiencing leadership transitions. Changes in leadership result in challenges to the existing system. However, because teachers at the schools believe so strongly in the value of AIW, and there is evidence that AIW promotes a culture focused on student and teacher learning, it is believed that AIW will continue to function. At Valley High School, Dr. Poole, who was soon to retire, stated that the School Improvement Team was about to develop goals for the building for 2011-2012 and would be sure to "follow the same path" with AIW. With the support of Superintendent Hopkins at AHST for AIW, and the commitment of staff, the new principal has the opportunity to build upon the past successes of AIW.

At Cedar Falls High School, scoring teams will be meeting on a much more consistent basis than in the past. Every week, teams will convene for ninety minutes for AIW professional development activities. At Spencer High School, one of the administrators stated that "AIW is strategic in nature and continues to be valuable overtime. It is not an initiative like others that have come and gone. It will remain an integral part of the school system in the future."

#### **Summary of Focus Group Discussions**

## Introduction

Two focus groups comprising 20 principals and three curriculum directors from AIW schools met in February, 2011, to discuss how AIW has impacted their schools. The leaders' responses to nine questions resulted in the identification of six major themes: student benefits; changes in professional culture; impact on instruction; the scoring process; the school leadership role; and support that is needed in the schools for implementing AIW.

#### **Student Benefits**

AIW, according to the focus group participants, has resulted in students being more engaged in the classroom, not only with their teachers, but also with other students. In addition, the quality of classroom discussions has been at a much deeper and more thoughtful level. Expectations for students have been increased and curriculum is now more closely connected to students' lives, making lessons more challenging and simultaneously, more meaningful. Because students are more engaged, one school leader had observed individuals being much more persistent in solving complex issues.

## **Change in Professional Culture**

Teachers from all subject areas and grades may participate on AIW teams at a school, and this has created a professional development model in which everyone is invested. The establishment of scoring teams has brought about a major shift in the professional culture of the schools implementing AIW. Instead of teachers working alone, in their classrooms, they are now, through the AIW structure, working together to help each other improve instruction. This has lead to successful collaborations amongst teachers at an unprecedented level, according to the focus group participants. Teachers also have a better understanding and appreciation of their colleagues' skills and abilities across all disciplines.

# The Impact on Instruction

The learning climate has transformed from the teacher as the deliverer of facts, with the expectation that information will be recited back by the student, to more of a facilitator role where the focus is on knowledge that is meaningful and valuable. To an observer of a classroom where AIW is practiced, it might appear noisy and confusing. This is a contrast to a classroom in the past, as one school leader put it, when "little soldiers" would sit in a row of desks as the teacher provided information, but little opportunity for active learning.

Teachers have come to view instruction through the AIW lens. Experienced teachers practicing AIW are now unable to design or deliver curriculum without asking themselves, "Will this provoke higher order thinking?" and "Does this have value to the student beyond school?" and "Will this lesson, as designed, lead to substantive conversation?"

# **The Scoring Process**

The school leaders concurred that the AIW scoring process has improved teacher tasks, student work and instruction. They also stated that the scoring process was helpful but more as a stepping off point for discussion than being an end in itself. In general, teachers have learned to bring items to their teams to be scored that they will use in the near future, as opposed to something that was recently taught and will not be used for another year.

# **Leadership and Change**

The school leaders stated that AIW has brought about a profound change in how teachers deliver instruction. As leaders, they have been part of the change process. A collaborative spirit now exist, not only between teachers, but also between teachers and administrators that focuses attention on what improvements can be made in the classroom. Because administrators have been involved in the scoring process, and, as one administrator stated, "We're all professional educators working toward the same goal," their credibility as instructional leaders has been enhanced.

The school leaders' role of evaluating teachers has changed. There has been a shift in approach to teacher walkthroughs. Observations of teachers are no longer focused on classroom management but on what students are learning. The administrators believed that they are providing more relevant feedback to teachers because of the AIW framework. There is a movement from some of the administrators to use AIW as a formal vehicle for teacher evaluation.

The administrators admitted that a major cause of the change brought by AIW is that it is teacher driven. It is apparent that once teachers are part of the AIW project they become committed to it as a professional development framework. They view AIW as an extension of what they are already doing only it provides a means for making assessment tasks and instruction better. They share with other teachers the benefits to be gained. Consequently, teachers not in AIW, "are banging on the door ready to get going," as one administrator described what was occurring at his school. The school leaders want to be sure that the enthusiasm for AIW can be sustained.

## **AIW Support**

The school administrators emphasized that support for AIW in their buildings is needed in the areas of training and data collection. School staff must continue receiving AIW professional development to help keep skills honed, and to inspire them to continue their efforts to improve. In addition, enhancing the data collection process via the Technical Assistance Report system would be a significant help. Reliable data is also needed to convince school boards and other stakeholders that the investment in AIW has delivered on its promise.

#### **Summary of Effects of Professional Development on Task Quality**

Teachers participating in the Authentic Intellectual Work initiative bring assignments and assessment tasks they've designed for students to their teams for scoring and discussion, including suggestions for improving the task. Discussion is based on the standards and rubrics for construction of knowledge, elaborated communication, and value beyond school that appear in *Teaching for Authentic Intellectual Work: Standards and Scoring Criteria for Teachers' Tasks, Student Performance, and Instruction* (Newmann, King, and Carmichael, 2009).

One of the expected outcomes of AIW professional development is that teachers will improve their ability to develop and implement tasks that align more closely to the AIW standards. To evaluate this, the Department gathered 27 literacy, 25 mathematics, 30 science, and 30 social studies assessment tasks, both the original and revised versions, from its AIW Website. The tasks were all developed and revised for use in grades 9 through 12. The main research question was "In a sample of AIW high schools teachers in the core academic areas, do revised tasks demonstrate significant improvement in the level of authenticity compared to original tasks?" Fifteen high school English, mathematics, science, and social studies teachers already participating in AIW received training and scored the tasks for this phase of the evaluation in June 2011.

Both the original and revised versions were scored independently by pairs of raters. Raters were not told if the task was an original or a revised version, who the teacher was or what school they were from. The scoring included each of the standards appropriate to task scoring: Construction of Knowledge, Elaborated Communication, and Value Beyond School. Each original and each revised task received a single composite score of 3 to 10 points for the all three standards. According to procedures commonly used in this type of research, scores of the two raters on any one task that differed by one point were averaged; scores that differed by more than one point were discussed by the two raters to arrive at a consensus score.

The review of tasks, both pre- and post-AIW team review and revision, showed that this sample of high school teachers participating in AIW improved the authenticity of their tasks in all subject areas. The improvements were statistically significant in the subject areas of mathematics, science, social studies, and for all tasks. See Table 1.

Since paired quantitative analysis is available, Pearson's correlation is used to determine the effect size. Pearson's r can vary in magnitude from -1 to1, with -1 indicated a perfect negative relationship, 1 indicated a perfect positive relationship, and 0 indicated no relationship between two variables. Cohen (1992) gives the following guidelines for the social sciences: small effect size, r=0.1; medium, r=0.3; large, r=0.5. Given that definition for effect size, math shows a large effect while science, social studies, and the overall results have medium effects. We can conclude that participation in AIW professional development has contributed to teachers' knowledge and skills in improving the intellectual quality of work expected of students in their tasks. It is highly unlikely that these tasks would have been revised to align more closely to the AIW Standards without professional development.

# **Iowa Department of Education**

Table 1: Task Review Analysis

Content Area	N	Original Mean	Revised Mean	Mean Differences	Pearson's r
Literacy	27	6.48	6.69	.20	
Mathematics	25	5.68	6.42	.74	0.59
Science	30	5.07	5.78	.72	0.30
Social Studies	30	5.53	6.33	.80	0.40
ALL TASKS	112	5.67	6.29	.62	0.46

# **Estimating the Effect of AIW on Student Performance**

# Design

An ideal design to estimate the effect of AIW professional development on student performance would include data on a) individual teacher's degree of participation in AIW professional development, b) the quality of each participating teacher's implementation of AIW as measured by their scores on revised tasks, and c) a database that permitted the linking of test scores for specific students to their teachers. Such a design could show the effects of AIW professional development on teachers' practice and the effects of teachers' practice on student achievement. Since the AIW-lowa project to date has not been able to link individual students' scores to their teachers or to record each teacher's actual level of participation in AIW professional development or the quality of their implementation of AIW standards, the initial evaluation used a simpler design. The effect of AIW on student performance was estimated by comparing student achievement in a sample of AIW schools, compared to achievement in a matched sample of non-AIW schools.

The AIW schools selected for the study by the DOE were those that had all teachers engaged in AIW as their primary professional development for one full year prior to the date of testing. Given that condition, data from 16 schools representing 10 districts implementing AIW were used. These districts and schools were matched to another set of Iowa schools and districts (not implementing AIW) using the following criteria:

- same grade structure
- within 10 percent on each of the following variables
  - o total student enrollment
  - o percent white
  - o percent Low Socio-Economic Status (SES)
  - percent English Language Learners (ELL),
  - o percent Individualized Education Plan (IEP).

As a result of using these criteria, data from 17 schools representing 12 districts were selected as the Non-AIW schools. See Table 2.

Test data in 2010-11 from the Iowa Tests of Basic Skills for grades 3 through 8 and the Iowa Tests of Educational Development in grades 9 through 11 in reading, mathematics, science, and social studies were compared. Across all grades and subjects, the scores of a total of 3,908 students in AIW schools and 4,060 students in non-AIW schools were compared.\* The total number of students disaggregated by grade level and by subject area can be found in Appendix A.

\*School districts in lowa retain local control of when they administer state accountability assessments (ITBS/ITED). Based on when students were tested during the year, their test score may be standardized to fall, midyear, or spring norms. These three norm groups are not directly comparable with one another. To report results that minimize differences due to the time of testing, each student's score was statistically adjusted using linear equating between fall, midyear, and spring testing within subject and grade level. The formula for this equating is given in Appendix B.

Table 2: Numbers of Districts, Schools and Students by Grade Level in the Analysis

Participants	AIW	Non-AIW
# of Districts	10	12
# of Schools	16	17
# of Students	3,908	4,060

Grade	Total # of AIW Students	Total # of Non-AIW Students
3 <sup>rd</sup>	150	159
4 <sup>th</sup>	137	152
5 <sup>th</sup>	334	267
6 <sup>th</sup>	423	266
7 <sup>th</sup>	443	450
8 <sup>th</sup>	450	485
9 <sup>th</sup>	683/193*	781/303*
10 <sup>th</sup>	551/224*	703/299*
11 <sup>th</sup>	737	797
Total	3,908	4,060

<sup>\*</sup>The smaller number of students listed at the 9<sup>th</sup> and 10<sup>th</sup> grade level represents the number of students for which we have comparable scores in the content areas being compared. If schools did not administer the test in the spring of 2011 we were unable to use their data in the analysis of means and percentile advantage.

Percentile advantage was calculated by taking the difference between means and dividing it by the standard deviation which is a measure of the amount that student scores are spread out on either side of the mean in the total distribution of student scores. In a normal distribution, which we assume here, about 34% of the scores are below the mean and 34% above the mean (that is, one standard deviation in each direction). A student scoring one standard deviation above the mean would be in the 84th percentile of the total distribution of students. So, for example, if on a test the AIW mean exceeds the non-AIW mean by 5 points and the standard deviation in that group of scores is 20, that difference represents 25% of a standard deviation. The average AIW student would score 25% of a standard deviation higher than the average non-AIW student. Since 25% of 34 percentiles equals 8.5 percentiles, the average AIW student would have an achievement advantage of 8.5 percentiles over the non AIW student.

In addition to comparing AIW and non AIW schools on test scores, the evaluation examined the percentage of students who scored proficient according to lowa's cut off points in the distribution of scores – students scoring at the 41<sup>st</sup> percentile or higher were considered proficient.

# **Results**

Tables 2-5 show average test scores for students in AIW and non-AIW schools in the four subject areas in grades 4, 8, and 11 – the grades for which lowa schools must report annually and which are included in lowa's Condition of Education Report. Results for all grades 3-11 are presented in Tables C2-C5 in

Appendix C. For these tables, positive numbers indicate AIW achievement is higher and \* indicates the mean difference in scores is statistically significant at the .05 level (i.e., the probability of observed difference being due to chance is less than 5 in 100).

Since neither the test scores nor numerical mean differences, even when statistically significant, convey a practical sense of "how big" these differences are, the fifth column of the tables expresses these differences as the advantage in percentiles, if any, that AIW students demonstrate. This indicates how many percentiles higher an average AIW student would score, compared to an average non-AIW student.# The fifth column of the tables gives these results.

lowa classifies the test scores into 3 proficiency levels: **below proficiency** includes all students scoring below the 41st percentile, **proficient** includes all students scoring at the 41st percentile to the 89th percentile, **advanced** includes all students scoring at the 90th percentile and above. The last column of the tables reports the percent of students in AIW and non AIW schools who scored proficient (41<sup>st</sup> percentile and above) on the lowa Tests.

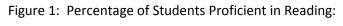
Figures 1-8 summarize the data in Tables 2-5 in graph form.

#### **READING:**

Table 3: Student Achievement in AIW and Non AIW Schools in Reading

	Non AIW Mean Test Score	AIW Mean Test Score	Difference in Means*	Percentile Advantage to Students in AIW Schools**	Percent AIW	Proficient Non-AIW
Grade 4	218.797	225.936	7.140*	9.2 percentiles	88.3%	83.4%
Grade 8	266.755	272.458	5.703*	5.5 percentiles	81.6%	78.1%
Grade 11	293.096	296.712	3.616	2.9 percentiles	82.0%	80.0%

<sup>\*\*</sup>Compared to an average student in a Non-AIW school (i.e., who scores at the 50% percentile), an average student in an AIW school would score approximately this many percentile points higher.



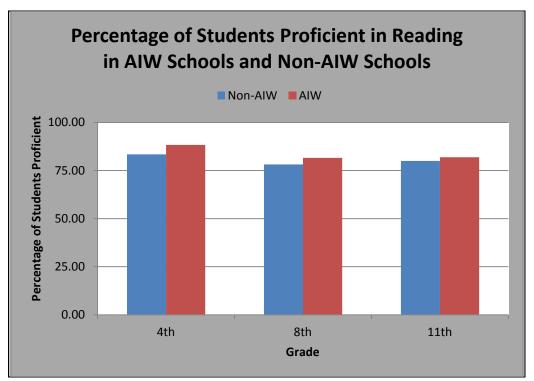
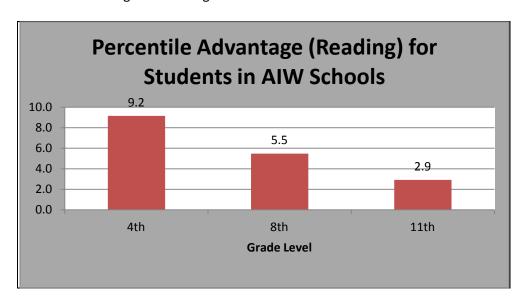


Figure 2: Percentile Advantage for Reading

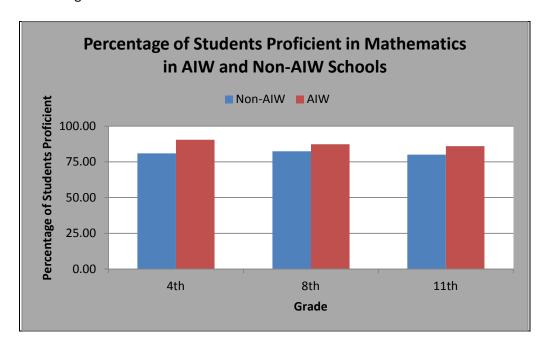


## **MATHEMATICS:**

Table 4: Student Achievement in AIW and Non AIW Schools in Mathematics

	Non AIW Mean Test Score	AIW Mean Test Score	Difference in Means*	Percentile Advantage to Students in AIW Schools**	Percent AIW	Proficient Non-AIW
Grade 4	216.567	224.036	7.47*	11.4 percentiles	90.5%	80.9%
Grade 8	269.915	277.467	7.55*	8.1 percentiles	87.3%	82.4%
Grade 11	293.747	305.814	12.07*	11.2 percentiles	86.0%	80.1%

Figure 3: Percentage of Students Proficient in Mathematics:



<sup>\*\*</sup>Compared to an average student in a Non-AIW school (i.e., who scores at the 50% percentile), an average student in an AIW school would score approximately this many percentile points higher.

**Percentile Advantage (Mathematics)** for Students in AIW Schools Percentile Point Advantage 14.0 11.4 11.2 12.0 10.0 8.1 8.0 6.0 4.0 2.0 0.0 4th 8th 11th **Grade Level** 

Figure 4: Percentile Advantage for Mathematics

## **SCIENCE:**

Table 5: Student Achievement in AIW and Non AIW Schools in Science

	Non AIW Mean Test Score	AIW Mean Test Score	Difference in Means*	Percentile Advantage to Students in AIW Schools**	Percent AIW	Proficient Non-AIW
Grade 4	218.964	222.801	3.84	4.8 percentiles	88.2%	83.4%
Grade 8	275.716	281.973	6.26*	6.5 percentiles	90.4%	87.6%
Grade 11	304.345	312.149	7.80*	6.7 percentiles	88.6%	84.4%

<sup>\*\*</sup>Compared to an average student in a Non-AIW school (i.e., who scores at the 50% percentile), an average student in an AIW school would score approximately this many percentile points higher.

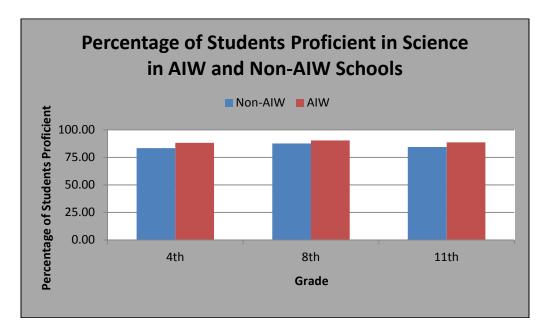
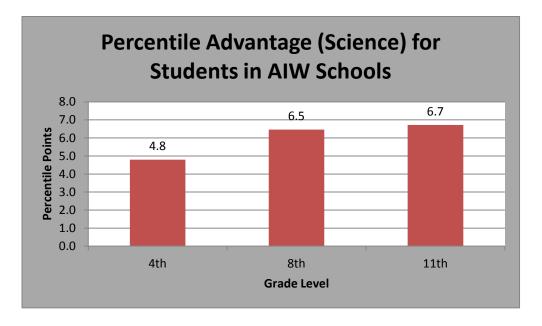


Figure 5: Percentage of Students Proficient in Science

Figure 6: Percentile Advantage for Science



# **SOCIAL STUDIES:**

Results for social studies include only about half of the number of students (3000) compared to other subjects (6000). Schools don't have to report in this area so many schools don't administer this subtest or don't administer it at all grade levels. And since the state does not require that this area be reported, the Department of Education does not have formulas for equating social studies fall or winter scores to

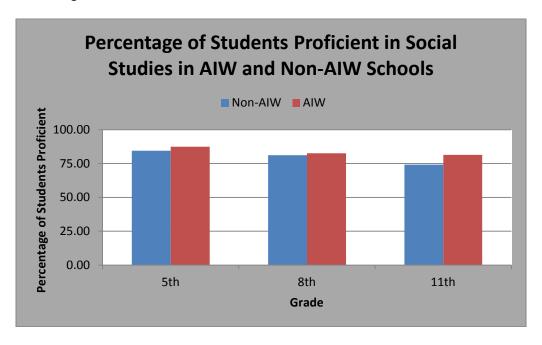
spring norms. Thus results here can include only schools that administered the social studies test and did so in the spring.

Because of the considerably lower number of tested students in social studies, we analyzed only grade levels with more than 100 students with social studies data in the AIW schools and the non-AIW schools. As a result, grade 5 (rather than 4), 8, and 11 are reported for social studies. Appendix C (Table C5) includes results for all grades, regardless of number of students tested.

Table 6: Student Achievement in AIW and Non AIW Schools in Social Studies

	Non AIW Mean Test Score	AIW Mean Test Score	Difference in Means*	Percentile Advantage to Students in AIW Schools**	Percent AIW	Proficient Non-AIW
Grade 5*	234.599	238.614	4.02	4.4 percentiles	87.4%	84.4%
Grade 8	269.273	268.603	-0.67	-0.6 percentiles	82.6%	81.2%
Grade 11	288.529	298.662	10.13*	8.0 percentiles	81.4%	74.2%

Figure 7: Percentage of Students Proficient in Social Studies



<sup>\*\*</sup>Compared to an average student in a Non-AIW school (i.e., who scores at the 50% percentile), an average student in an AIW school would score approximately this many percentile points higher.

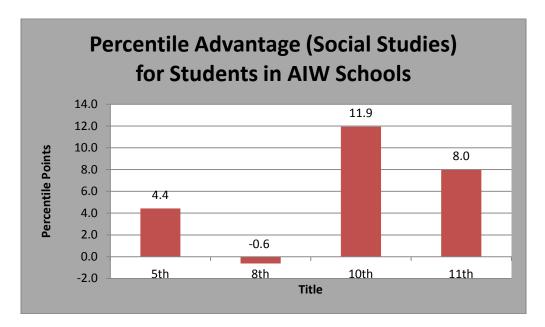


Figure 8: Percentile Advantage for Social Studies

## **Summary and Implications of Achievement Data**

Students in AIW schools across grade levels and subjects usually scored higher than students in Non-AIW schools. For the three grade levels reported, AIW students scored significantly higher in 8 of the twelve comparisons (3 grades x 4 subjects) and AIW had higher percentages proficient in all 12 comparisons. The percentile advantage to AIW students was 5 points or higher in 8 of the 12 comparisons. The data across all 9 grades (3-11) in Appendix C shows similar results. That is, of thirty six comparisons (9 grades x 4 subjects), AIW students scored significantly higher in 26 comparisons and AIW had higher percentages proficient in 32 comparisons. The percentile advantages to AIW students was 5 points or higher in 26 of the 36 comparisons. Of the four subjects, AIW students posted the most consistently higher scores in mathematics and showed consistently smaller differences in social studies.

The data here offers no information to suggest explanations for achievement differences between grade levels and subjects, but the consistently positive achievement results for AIW vs non-AIW students across grades and subjects can be explained in at least three ways.

Prior research shows that students of teachers who practice authentic pedagogy at high levels score higher on both standardized and more authentic assessments than students of teachers whose practice reflects lower levels of authentic pedagogy (for a summary of this research, see Newmann, King and Carmichael, 2007). If teachers in these AIW schools practiced higher levels of authentic pedagogy than those in non-AIW schools, we would expect such results. However, since this evaluation did not include data on differences in AIW and non-AIW teachers' actual classroom practice, we do not know whether this explanation accounts for the results.

Although we cannot document differences in classroom practice between AIW and non-AIW schools, the AIW schools participated in intensive professional development on AIW which was not available in non-AIW schools. The evaluation design did not permit examination of the relationship between the extent of individual teachers' participation in the program and their students'

achievement, but since all teachers in each AIW school participated in the AIW program, it is possible that the AIW program motivated AIW teachers to work much harder than non-AIW teachers to boost their students' achievement. This explanation for the results, while consistent with the case study reports herein, cannot be confirmed because the evaluation did not collect systematic data on the differences in motivation or effort to teach between AIW and non-AIW teachers.

Finally, it is possible that teachers in those schools that volunteered to participate in the AIW program were more highly motivated to devote serious effort to their students' achievement than teachers in non-AIW schools, before they had any experience with AIW professional development. This selection bias could explain higher achievement in AIW schools, independent of participation in the AIW program. This explanation could be tested by examining achievement trends in AIW and matched non-AIW schools prior to the AIW schools' participation in the AIW program.

While none of the explanations can be confirmed through this evaluation, the consistently favorable results for students in AIW schools suggests that the AIW program should be continued and that future evaluations should be designed to assess the relationship between the extent of individual teachers' participation in AIW professional development activities, their actual classroom practice, and their student's achievement, preferably disaggregated by socioeconomic status, race/ethnicity, gender, and other demographic characteristics known to affect school achievement. To explore the possibility of selection bias into the AIW program, future evaluations should also examine achievement trends in AIW schools prior to entry into the AIW program.

# References

Cohen, J. (1992). A power primer. Psychological Bulletin, 112, 155-159.

Newmann, F.M., King, M.B. & Carmichael, Dana L. (2007). *Authentic Instruction and Assessment: Common Standards for Rigor and Relevance in Teaching Academic Subjects*. Des Moines, IA: Iowa Department of Education. <u>www.centerforaiw.com</u>.

Newmann, F.M., King, M.B. & Carmichael, Dana L. (2009). *Teaching for Authentic Intellectual Work: Standards and Scoring Criteria for Teachers' Tasks, Student Performance, and Instruction.*<a href="https://www.centerforaiw.com">www.centerforaiw.com</a>.

# Appendix A

# **Total Number of Students by Subject and Grade in AIW and Non-AIW Schools**

Table A1: Number of Students Tested by Subject and Grade in AIW and Non-AIW Schools

Grade	Rea	ding	Mathe	matics	Scie	ence	Social S	Studies
Crauc	AIW	Non-AIW	AIW	Non-AIW	AIW	Non-AIW	AIW	Non-AIW
3 <sup>rd</sup>	150	158	150	159	150	105	19	11
4 <sup>th</sup>	137	151	137	152	136	151	16	19
5 <sup>th</sup>	334	267	334	267	334	267	111	122
6 <sup>th</sup>	423	266	421	265	420	266	142	139
7 <sup>th</sup>	443	450	439	449	442	448	165	315
8 <sup>th</sup>	450	485	449	484	449	484	161	340
9 <sup>th*</sup>	193/681	303/781	193/682	303/781	193/683	295/772	74/311	150/363
10 <sup>th*</sup>	224/551	299/703	223/550	297/700	224/550	291/693	223/381	140/283
11 <sup>th</sup>	737	796	735	797	735	790	334	376

- For grades 9 and 10 the smaller numbers represent the number of students who took the sub test in the spring. We were not able to calculate equated scores for these two grades with students taking the subtest in fall or winter 2010-2011. For these students means, significance, and percentile advantage are based only on spring scores.
- For grades 9 and 10, the larger sizes represent the total number of students who took the subtest in the AIW and Non-AIW schools during all 2010-2011 school year. For these students, percentile ranks could be determined and used for identifying the percent of students proficient.

# Appendix B Linear Equating Formula

School districts in lowa retain local control of when they administer state accountability assessments (ITBS/ITED). Based on when students were tested during the year, their test score may be standardized to fall, midyear, or spring norms. These three norm groups are not directly comparable with one another. To report results that minimize differences due to the time of testing, each student's score was statistically adjusted using linear equating between fall, midyear, and spring testing within subject and grade level. The formula for this equating is provided below:

# Formula for Linear Equating of 2010-11 Fall and Spring ITBS Scores\*

Linear equating:

$$\frac{X_{1} - \overline{X}_{1}}{S_{1}} = \frac{X_{2} - \overline{X}_{2}}{S_{2}}$$

$$X_{1} = \frac{S_{1}}{S_{2}} X_{2} + (\overline{X}_{1} - \frac{S_{1}}{S_{2}} \overline{X}_{2})$$

For example:

$$X_{spring} = \frac{S_{spring}}{S_{fall}} X_{fall} + \left(\overline{X}_{spring} - \frac{S_{spring}}{S_{fall}} \overline{X}_{fall}\right)$$

For a fourth grade student who scores 210 on a fall ITBS RC:

$$X_{spring} = \frac{28.7}{25.9} \times 210 + (202.6 - \frac{28.7}{25.9} \times 193.8) = 220.6$$

<sup>\*</sup>Kolen, M.J. (1988). An NCME Instructional Module on Traditional Equating Methodology. National Council on Measurement in Education, Madison, WI.

# Appendix C Achievement Tables for All Grades

Table C2: Student Achievement in AIW and Non AIW Schools in Reading

	Non AIW Mean Test Score	AIW Mean Test Score	Difference in Means*	Percentile Advantage to Students in AIW Schools**	Percent AIW	Proficient Non-AIW
Grade 3	198.877	203.101	4.224	5.5 percentiles	82.0%	79.7%
Grade 4	218.797	225.936	7.140*	9.2 percentiles	88.3%	83.4%
Grade 5	233.191	241.194	8.003*	9.4 percentiles	89.8%	84.3%
Grade 6	241.447	248.957	7.510*	8.1 percentiles	83.5%	77.8%
Grade 7	260.701	263.103	2.402	2.3 percentiles	84.2%	80.9%
Grade 8	266.755	272.458	5.703*	5.5 percentiles	81.6%	78.1%
Grade 9	273.825	277.456	3.631	3.6 percentiles	81.9%	77.9%
Grade 10	273.706	285.353	11.647*	10.4 percentiles	79.5%	74.8%
Grade 11	293.096	296.712	3.616	2.9 percentiles	82.0%	80.0%

Table C3: Student Achievement in AIW and Non AIW Schools in Mathematics

	Non AIW Mean Test Score	AIW Mean Test Score	Difference in Means*	Percentile Advantage to Students in AIW Schools**	Percent AIW	Proficient Non-AIW
Grade 3	197.873	203.144	5.27*	9.2 percentiles	87.3%	80.5%
Grade 4	216.567	224.036	7.47*	11.4 percentiles	90.5%	80.9%
Grade 5	233.932	245.652	11.72*	13.4 percentiles	89.8%	84.6%
Grade 6	245.523	256.203	10.68*	12.3 percentiles	88.6%	85.3%
Grade 7	263.114	270.162	7.05*	8.2 percentiles	90.0%	83.5%
Grade 8	269.915	277.467	7.55*	8.1 percentiles	87.3%	82.4%
Grade 9	272.442	290.772	18.33*	19.1 percentiles	85.6%	79.3%
Grade 10	275.522	296.404	20.88*	19.7 percentiles	86.1%	75.3%
Grade 11	293.747	305.814	12.07*	11.2 percentiles	86.0%	80.1%

Table C4: Student Achievement in AIW and Non AIW Schools in Science

	Non AIW Mean Test Score	AIW Mean Test Score	Difference in Means*	Percentile Advantage to Students in AIW Schools**	Percent AIW	Proficient Non-AIW
Grade 3	195.332	205.607	10.27*	16.6 percentiles	93.3%	79.0%
Grade 4	218.964	222.801	3.84	4.8 percentiles	88.2%	83.4%
Grade 5	239.089	247.105	8.02*	9.2 percentiles	91.3%	86.5%
Grade 6	249.670	257.179	7.51*	7.4 percentiles	86.0%	84.6%
Grade 7	268.584	273.814	5.23*	5.0 percentiles	91.2%	86.8%
Grade 8	275.716	281.973	6.26*	6.5 percentiles	90.4%	87.6%
Grade 9	279.684	289.336	9.65*	9.2 percentiles	87.4%	82.5%
Grade 10	288.817	297.446	8.63*	8.1 percentiles	87.5%	82.8%
Grade 11	304.345	312.149	7.80*	6.7 percentiles	88.6%	84.4%

# Appendix C (continued) Achievement Tables for All Grades

Table C5: Student Achievement in AIW and Non AIW Schools in Social Studies

	Non AIW Mean Test Score	AIW Mean Test Score	Difference in Means*	Percentile Advantage to Students in AIW Schools**	Percent AIW	Proficient Non-AIW
Grade 3	198.990	206.498	7.51	10.8 percentiles	94.7%	63.6%
Grade 4	213.951	216.366	2.42	3.1 percentiles	81.3%	84.2%
Grade 5*	234.599	238.614	4.02	4.4 percentiles	87.4%	84.4%
Grade 6	247.167	250.733	3.57	3.8 percentiles	82.4%	82.7%
Grade 7	263.840	252.407	-11.43*	-10.9 percentiles	72.1%	82.5%
Grade 8	269.273	268.603	-0.67	-0.6 percentiles	82.6%	81.2%
Grade 9	265.040	267.608	2.57	2.0 percentiles	76.9%	68.6%
Grade 10	273.285	287.475	14.19*	11.9 percentiles	74.8%	69.6%
Grade 11	288.529	298.662	10.13*	8.0 percentiles	81.4%	74.2%

<sup>\*\*</sup>Compared to an average student in a Non-AIW school (i.e., who scores at the 50% percentile), an average student in an AIW school would score approximately this many percentile points higher