Leveling the playing field:

Increasing student achievement through data-driven ability grouping and instruction.

Jami Sexton

January 6, 2010

Abstract

This action research project focuses on increasing student comprehension and achievement. The study examined the effectiveness of completing detailed item analysis of assessments for the purpose of placing students into different Language Arts classes and learning groups with in those classes. Research advocates placing students of similar ability levels for reading instruction as long as instructional practices are equitable. After student placement, the data was used then for instructional planning. Item analysis of weekly assessments and a unit test continued for the purpose of arranging learning groups and for instructional planning. The participants in this study consist of ten fifth-grade students. The students participated in lessons on a number of reading skills. Data were collected from fifth-grade Ohio Achievement Tests, Reading Benchmark tests, weekly assessments, a Unit Test, a reading survey, an interest survey, self-evaluation survey, conferences, and observational notes. The results of the study suggest a positive relationship between detailed item analysis and student placement. The results also show a positive relationship between using detailed data analysis for instructional planning and increasing student achievement.

Introduction

In response to No Child Left Behind, school districts across the country have been scrambling to try to find “scientifically-based research to support educational practices” perhaps “the time has come to revisit an old friend (or foe depending upon one’s personal persuasion): ability grouping” (Tieso, 2003, p.1). This is not your mother’s ability grouping, however. Today’s ability grouping is decidedly not the old tracking that placed students at a disadvantage by placing them into “low-achievement groups” with “self-

fulfilling low expectations" (apples4theteacher.com, 2004). Today, ability grouping allows classroom instruction to use the results of high-stakes testing for what it was originally conceived: driving instruction to provide quality education for all students. This paper is based on my experience grouping students into learning communities based not on an overall score on one high-stakes test, but reviewing the results of each question for each student on a high-stakes to determine how to group students and provide each child with an equitable, quality education.

The ability groups the children are placed in will be based on their results from individual Ohio Achievement Test questions, a Benchmark reading test and student interviews and surveys. Despite the fact that I am ability grouping, this data will be used to devise learning communities that may prove to be more heterogeneous than if I sorted the students only by overall score without any context concerning that score. Students will be placed into classes by specific need in order to address these needs. Once students are placed into these learning communities, I will be able to address students' specific reading needs and differentiate my instruction for the sole purpose of increasing student comprehension and achievement. During the course of my research, I will collect multiple pieces of data to demonstrate students' progress and achievement in the form of weekly and short-cycle assessments. In addition to quantitative research, I will also conduct qualitative research in the form of student surveys, observations and one-on-one interviews.

Ability grouping has been rejected as a 'best practice' grouping method because it has been associated with tracking students. The ability grouping of today addresses the academic needs of students. I believe that when ability grouping is well designed, there is

equality in instruction and it will “readily help” and “improve” (video: Expert Commentary: *How can I identify an area of focus for my research*) the reading abilities of my students. I am, however, acutely aware students self-esteem also needs to be monitored to ensure no child begins to feel as if they are in the “dumb” class. To address these concerns, I have devised an anonymous survey to allow students to share with me any negative feelings that may come from these groupings. I also plan to address these issues during my student interviews. Given the individual nature of the sorting process, I believe my classes will become more heterogeneous instead of homogeneous. I do not anticipate issues with self-esteem, but also believe I need to be cognizant of the warning signs and address these personal needs in a timely manner. As much as I am concerned with my students’ achievement, in the course of a lifetime, positive self-esteem is far more important.

Research Questions

1. What effect does sorting student groupings through analysis of individual testing results versus sorting only by overall score have on the make up of the student groupings?
2. What effect does differentiated instruction have upon students’ performance levels in their individual areas of need? Describe any differences in performance levels when the differentiated instruction was delivered in small groups (four students or less) or in the whole-class setting.
3. What effect does placing students in ability groups have upon their self-esteem?

Literature Review

Tieso, C.L. (2003). *Ability grouping is not just tracking anymore, 26. The Questia Online Library.* Retrieved May 13, 2009, from http://www.questia.com

This article fully embraces and justifies both my educational practice of ability grouping and my desire to use a combination of high-stakes testing, student ability and instruction based on students' needs to make "significant improvements in students' achievement" (Tieso, 2003, p. 1). This is exactly what I want to do with my action research project.

The article is interesting because it doesn't shy away from addressing the critics of ability grouping. The author doesn't dance around or try to excuse the very real inequalities that existed in some practices of ability grouping when it was used to track students and there was no room for the students to "escape" their perceived level. However, as far as Tieso will go in conceding any faults with ability grouping. Instead of throwing the baby out with the bath water, she encourages educators to use ability grouping to address the different needs of students in a manner that will allow them to achieve.

There are several practices included in the article, some of which I had already considered and some that would be impossible in my current setting. I am encouraged I have been on the right track in my thinking about better ways to serve the needs of all of my students, not just the ones with letters (IEPs) and numbers (504s) attached to their names.

Mitchell, R. (2006, March 30). Effects of high-stakes testing on instruction [Review]. *The Center for Public Education.* (2006. May 30). Retrieved May 13, 2009, from http://www.centerforpubliceducation.org

This article addresses a number of issues I am interested in for my research project and my teaching practice. One issue addressed in this article is the research on the effects of testing. I thought it was very interesting most of the research has been about why we should or shouldn't have high-stakes testing and little research has been done to address the actual effects testing has on classroom instruction. The research has been done "shows teaching a curriculum aligned to state standards as feedback produces higher test scores than an instructional emphasis on memorization" (2006, p. 2). No offense to the researchers, but they could have learned in any basic teaching methods course memorization is not a good teaching practice.

The article includes some very interesting information about attitudes towards testing from the general public, teachers, counselors and students. I found it intriguing the public had concerns about validity of the testing, but wanted to know their school district was performing well. The teachers' attitudes emphasized their concern other teachers would only teach to the test and not properly instruct their students, but they would not do that themselves. I was most surprised by the report most students do not suffer from test anxiety. The data from this survey though may not be valid because the focus group was very small. Test anxiety is one of the factors definitely skewing a child's test results that would definitely have an impact on my analysis of test data, so I will need to find more information about this factor.

Finally, the article addressed the issue because widespread testing is relatively new, there is little research on the use of test data, but there is a model for analyzing student data. The article only mentions Todd McIntire's guide, but I am now trying to

find out more information about what he determined were best practices for helping all students to meet the state standards.

Schullery, N.M., & Schullery, S.E. (2006). Are heterogeneous or homogeneous groups more beneficial to students?, 30(4), 542-556. *Journal of Management Education*, (2007). Retrieved May 13, 2009, from http://www.sagepublications.com

This article states the argument between heterogeneous and homogeneous groupings by comparing the classroom setting to a work setting where everyone is a mix of abilities. The authors make compelling points about life outside of the school setting. I think many of the points presented are interesting and though provoking because they are different from my own. This article will help me to consider a different point of view about how I group in the future.

The article does state some successful ways homogeneous groupings have worked. It is these successes and how they were achieved I need to keep my focus on as I develop my action plan for student success.

Anderson, K.M. (2007). Tips for Teaching: Differentiating instruction to include all students. *Preventing School Failure*, *51*(3), 49-54. (2007). Retrieved May 14, 2009, from http://www.heldref.org

This article focuses on strategies to help all of the students in the classroom through a variety of means. Although the author doesn't specifically state 'ability grouping,' but she describes sorting students into groups which are flexible in nature allowing students to work toward mastery.

This article will help me as I develop my action plan for student success. There are tips for getting started with differentiating instruction and there are steps to take to implement the strategies in the classroom. The article describes how these strategies

worked for a real teacher in a real classroom. This information is the most helpful because it clearly shows these techniques are doable in a classroom setting. I am finding there are many strategies available, but I have serious reservations as to their validity in a classroom setting with real children.

Torgesen, J., Schirm, A., Castner, L., Vartivarian, S., Mansfield, W., Myers, D., et al. (2007). *National Assessment of Title I: Final Report* (pp. 1-276, Rep. No. NCEE 2008-4013). Washington, D.C.: U.S. Department of Education.

This reference has an overwhelming amount of data and materials about different reading intervention for struggling readers. I will use some of the techniques mentioned in the report to help my own students. Some of the testing models used are outside of the scope of what I could possibly do with my students, but generally, much will be quite useful.

There is a great deal of data included in the report. Not a great deal of it is useful to my cause, but I have found the models to be very interesting and will use some of the data analysis practice in my own research.

Educational Sources. (2004, November 7). Math and Reading Ability Grouping in Elementary Schools. In apples4theteacher.com. Retrieved May 18, 2009, from http://www.apples4theteacher.com.

This article discusses the differences between ability grouping within the classroom setting and from class to class. There are several theories addressed within the article outlining different grouping plans and their potential effectiveness.

The first theory discusses regrouping only for one or two subjects. The suggestion is grouping can “improve student achievement,” but “the level and pace of instruction must be adapted to achievement level” (Educational Sources, 2004, p.2).

The second theory is the Joplin Plan. I've read about this in other publications. I agree with this plan in theory to have heterogeneous groupings throughout the day, but "across grade levels for reading instruction" interesting perspective seems to be outside of the realm of what most school districts would be able to do. I also question having three different grade levels in one class without it having a negative impact on students' self-esteem.

The third plan is the Non-graded Plan. Students are placed into groups based on performance level rather than a grade. Much like the Joplin Plan, I question having students at different ages in the same class period without there being any effect on student's self-esteem.

The last plan is within-class ability grouping. The article states this is most effective for math instruction. The article corroborates what I have been finding in my own research: not enough is known about the effectiveness of any one type of grouping for reading instruction because not enough research has been done concerning reading instruction because of the difficulty with determining a control group. This is where I am also struggling to determine on what basis will I know my plan is effective. This article raised more questions for me than it answered, but in and of itself is important.

Muir, M. (2007). Tracking and ability grouping [Research brief]. *The Principals' Partnership*. (2007, January 8). Retrieved May 18, 2009, from http://www.principalspartnership.com

This article outlines the differences between ability grouping and tracking. Much like the other articles I have found, tracking is criticized for a multitude of sins against education. Ability grouping is discussed differently when done only for reading

instruction. Again, the focus is to make sure all students are receiving the same level of instruction to ensure equity.

Equity in education is the main focus of this article. This is where ability grouping goes very wrong when equity is not strived for and subsequently achieved. The article gives some solid advice on how to achieve this and I plan to try to develop a graphic organizer to keep myself in check to make sure I am indeed achieving equity in my instruction.

Westchester Institute of Human Services Research. (2003). High-Stakes Testing. *The Balanced View. 7*(1).

This article focuses on the positives and negatives of high-stakes testing. For my purposes, the article provides information about using high-stakes testing to form school curriculums to ensure that schools, and subsequently classroom teachers, are teaching what the state requires.

The article states there is little research on the impact high-stakes testing can make in a classroom setting. There is a word of caution about teaching to the test, but at the same time, the article encourages teachers to use their test data in a way to effect a positive change upon their own practices. I am hopeful I will able to make a positive change upon my practice so I can help my students become better readers.

Schumm, J., Moody, S., & Vaugh, S. (2000). Does one size fit all? *Journal of Learning Disabilities* (2000, September 1). Retrieved June 2, 2009, from http://www.accessmylibrary.com.

This article presents grouping from the perspective of children with learning disabilities. This is very important to me because I do have children in my classes with

IEPs and 504 plans. One pitfall I want to avoid is students thinking there is a ‘dumb class’ and causing their self-esteem to suffer.

The authors stress the need for inclusion reading classes and stress equity in instruction. They also state when small groups are made within these classes, the students should be placed into heterogeneous groups. Although I have always felt uncomfortable with our Small Group Tutor pulling students from class for reading instruction, I am guilty of placing those students with IEPs in the same small reading groups so she can work with them within the classroom setting.

I need to be aware of how I form my small reading groups. I have never kept track before as to whom I have in groups together versus how students perform on reading assessments, but that might something to consider as I conduct my research.

The National Center for Fair and Open Testing. (2001, January 1). Testing plus real accountability with real results. *FairTest.* Retrieved June 2, 2009, from http://www.fairtest.org

This article focuses in what the real purpose of high-stakes testing is: making schools and teachers accountable for instruction and provides information for how to use test data to improve instruction. The authors state teachers are given a stack of numbers, but are rarely told how to use the data. They suggest looking at individual questions rather than overall scores to alleviate the guesswork as to why a student scores what they did on a particular test. This is exactly the premise of my research.

The article goes on to stress before teachers can be expected to have students who perform well on these tests, administrators need to do their job first by aligning the curriculum with state standards and then by providing teachers with the materials needed

to provide instruction. Neither of these two issues are problems I face in my own classroom, but I am surprised there are schools would follow their own curriculum instead of following what the state has outlined.

What I will use this article for in my own practice are the sections which give real examples of what other educators have done in their classrooms to improve instruction. There were many interesting ideas from teachers in Rhode Island and California, but I really like an idea from a school in Colorado in which student progress on certain benchmark skills is tracked, but students also keep a journal tracking their progress. I like the idea of students taking ownership for their achievement. I'm not certain how I would implement this, and I really wish the article would have provided the organizer that is used, but I truly think it is worth taking the time to figure out.

Research Process

My research process started with my data collection matrix (see Appendix A). My first question was whether or not specific item analysis would make a difference in placing students into classes. Rather than place students in ability groups simply based on an overall score, I wanted to determine if there would be a difference in groupings if I chose to analyze individual skill performance. To answer this question, I had to start with the students' Ohio Achievement Test in Reading results. I reviewed how students did on each question to see if I could determine specific skill deficiencies.

Due to many factors that might have influenced the OAT scores, I believed it was important to have a second reading test to analyze. At the beginning of the school year, I administered a Benchmark test from our reading series. I have administered this test in the past and have determined that although it is a challenging test, the results are both

valid and reliable. I used these test results to cross-reference skill levels and make decisions about placement into ability groups.

One of the factors that could have influenced students' scores on the OAT is the students' attitude concerning the test. To determine if attitude could be a variable in the student's results, I surveyed students about their feelings concerning the state-mandated test and what it means to them. Test anxiety, nervousness or apathy all could negatively impact test scores.

Once skill needs are determined, I wanted to know the affect of differentiated instruction would have upon students. I started by surveying students about their perception about their reading skills and what they believe are their areas of need. This allowed me to get to know my students better, and also allowed me the opportunity to be able to understand how they feel about reading so I would know if I would have self-esteem issues to deal with from my students.

Actual differentiated instruction was next. Since I knew what areas students struggled with on past assessments, I could make changes to my lesson plans for not just the different learning groups, but to form groups within my different classes. Certain classes or groups would receive more modeling and scaffolding while others were given less, but all groups were given the same assessments to determine if they were able to achieve on the same levels.

The assessments I used to determine if differentiating by individual reading skill were daily homework practice, weekly quizzes, surveys and unit tests about how the students believe they were progressing. Daily practice is important to help me determine if students are ready to move on to more difficult skills or if more modeling is needed.

The weekly quizzes provide an assessment opportunity for students to demonstrate what they can achieve independently. These weekly quizzes also provide me with additional data to be able to plan future lessons.

The student surveys provide a vital service for both the students and myself. The survey forces the students to think critically about the skills presented from the Unit and then decide where they honestly think they are at in their progress. I use these surveys to track what areas students still feel uneasy about as we approach the Unit assessment. I can use the students' perceived areas of need with what I have concluded to be areas of need.

The Unit test provides a format for students to be demonstrate their skill levels when the skills are no longer presented in isolation, but are combined with a wide variety of other reading skills. Following the completion of this assessment, the students and I can revisit their surveys to decide if they accurately gauged their skill levels before the assessment. I can also use the students' scores on individual skills to continue differentiating my instruction.

When ability grouping, it is extremely important to make sure that students' self-esteem stays intact and no one group or individual thinks they are in the 'dumb' group. To accomplish this task, students are surveyed about working with others in groups, how they believe the best groups are formed and what they believed are the most effective groups. This information will allow me to be well aware of my student's feelings and possible anxiety about working with others.

As the school year progresses, I will continue to keep my students' self esteem under consideration by conducting student interviews and keeping my own observational

notes as to how students are working together, progressing and developing as students and young people.

Data Analysis

Officials who set educational policy and standards may state they are concerned about increasing student achievement, but those achievement levels are currently only determined by one high-stakes test. Classroom teachers are in the precarious position of keeping state test scores high, but also meeting the educational needs of their students. If achievement testing is where meeting state-mandated achievement levels end, then that is also where determining student needs should begin.

Rather than looking at one score and placing a student in a homogenous group or ignoring the scores altogether, what would happen if detailed analysis of the high-stakes test results occurred? Instead of spending weeks trying to figure out what skills students need help with in order to become better readers, what if an educator had this information before the school year even started and designed their lessons based on the students' abilities? Too often teachers plan lessons based on an ideal merely to finish a textbook or a workbook. Student needs and interests are somehow left out as if merely covering a skill will be enough.

My first area of research was to determine if analysis of students' scores by individual question rather than an overall score would make difference in the manner in which I placed them into different Language Arts classes. To determine this, I had to start by placing students into classes using the manner in which I have in the past, simply by overall score. Those students with Advanced scores, 460 or higher, were placed together (group A), Accelerated scores, 435 to 459, were placed together if there were enough

students (group B). Otherwise, I would fill in this class with students who scored Proficient, but were close to Accelerated. The next class would be students who were within the Proficient range, 400 to 434, (group C). The last group consists of students who scored within the Basic, Limited or fifteen points or less into the Proficient level (group D).

Next, I analyzed the scores of each child's fifth-grade Ohio Achievement Test. Once this analysis was complete, I moved two children with low (407 and 416), but Proficient scores to Group A because they had done well on all of the multiple-choice sections of the test with the exception being the questions about the poem. These two students also consistently did poorly on the extended response questions. I moved three children from Group C to Group D because although they had scored well within the Proficient range (436, 429 and 423), they had defined areas where they consistently missed questions about specific skills no matter the format.

Once my class rosters were initially set, I wanted confirmation the information from the Ohio Achievement Tests was accurate. To accomplish this task, I administered a Reading Benchmark Test from our reading series. I have administered this assessment for the last three school years and have found it to be both valid and reliable. The scores on this assessment were consistent with the scores from the OAT and I was able to evaluate the students' extended response answers with much more scrutiny since with the OAT I am only given an extended response score without explanation. After completing the analysis of the Reading Benchmark Test, I only moved one student from Group B to Group D. His OAT score was quite high, 468, but he had taken the test with

accommodations. He has since been moved from an IEP to a 504. Without the test accommodations, he did very poorly on the Reading Benchmark Test.

In addition to administering another assessment model, I also surveyed the students concerning their attitudes and opinions about the OAT. I wanted to determine if any student's score could have been skewed by test anxiety or apathy. Throughout all of my students, seventy-three children, I only had five students who reported test anxiety, but of those students only three thought it affected their scores. I was surprised by the number of students, thirteen, who reported they 'didn't care' about how they did on the OAT. Reasons given for these attitudes ranged from 'they don't count for a real grade' (8 students) to 'I didn't like my teacher and thought she would get into trouble if I did bad' (2 students).

Even though I would complete the data analysis for all of my students, my next task was to choose the students I would evaluate for this project. I choose ten students (See Appendix, chart B for biographical information) from my Group D because, although their test scores on the OAT were varied (see Appendix, chart C), they did poorly on the Benchmark test (See Appendix, chart D) and consistently missed questions based on the same skills on both assessment models. In addition to choosing these students, I also was aware of what skills we would need our focus: main ideas and details, problem/solution, cause and effect, context clues and summary.

After my evaluation of whether or not individualized data analysis would make a difference in my groupings, I began the task of planning for reading instruction. To assist me in my planning, I surveyed the students about their own preferences (See Appendix, Chart E). I ask about their interests and hobbies outside of school first. How do they

spend their free time? Is there anything they are passionate about? A marked majority of students listed sports as the activity they love the most to participate in during free time. Forty percent of the students listed a sport as their passion. Other free time interests included playing with animals, playing video games, watching TV, drawing and reading. I was pleasantly surprised only Student G listed playing video games as their passion. I was disappointed, but not surprised, only one student, E, listed reading as a passion. Understanding what my students enjoy outside of school helped me to find reading materials, both primary and supplemental, that would be of high-interest to them.

I also surveyed the students about their attitudes about reading. I asked if they like to read, if they believe they are good at reading, attitudes about reading groups, what they find troublesome about reading and, when they do read, what is their favorite genre. When asked if they like to read, eight of the students said they did like to read, one said they did not like to read at all and one student checked ‘sort of’ and added it depended on the reading material making my selections of reading materials that much more important.

When asked if they thought they were good readers, only Student G checked ‘yes’ on the survey. The other nine stated they were ‘sort of’ good readers. When questioned about what they perceived as their biggest area of weakness, seven of the ten students said fluency yet of those seven, only four said they didn’t like to read aloud. Two of the other students identified their weaknesses as comprehension based and neither of those students said they liked to read aloud.

When asked about how they feel about reading in groups, only Student H said he would feel uncomfortable reading aloud. All ten children stated they felt uncomfortable

reading aloud in Science and Social Studies. Specific explanations varied, but each student expressed they were worried about their fluency in front of their friends. When asked about what made them feel most comfortable reading in a group all ten students responded they are most comfortable when they read with other children who have the same reading skills and weaknesses they do, so they don't feel 'dumb.' Only Student E made any mention of reading with a 'low' group and did not elaborate on her survey as to what she though constituted a 'low' group.

Of the weaknesses mentioned, one student, G, said his problem area was staying in the same spot when reading. During our first reading conference, he established that he tries to use his finger to stay on track, but it doesn't always work, so he ends up rereading sections or skipping sections entirely especially when he is excited about a passage because he tries to read too fast.

Finally, I asked about what they enjoy reading to help me choose reading materials. Seven of the ten students listed adventure-based stories as their favorite type of fiction to read. The other three students listed fantasy as their number one choice. It should be noted that all three of these students are girls and asked what genre the Twilight series would be considered. Non-fiction was more varied. Not surprising, five students said they liked reading about sports stories, while three students said they liked reading about space and other Science-based books. Only one student, C, stated she liked reading biographies and the final student, B, stated she did not like reading any type of non-fiction. Further analysis of her assessment scores would support this, as it was on the non-fiction selections that she had the worst scores. In fact, she did not even attempt a four-point extended response for a non-fiction selection and in during our conference

admitted she didn't even finish reading one of the non-fiction selections on the OAT once she realized it was non-fiction and just filled in blanks for the multiple choice.

The first week of instruction, we read a realistic fiction selection. The reading skills for that week included basic story elements and the vocabulary skill was one of our target areas: context clues with multiple meaning words. As we continued to work on these two skills throughout the week, I selected a realistic fiction story about a boy playing baseball and another about a girl playing soccer. Additionally, I we also read non-fiction selections about finding fireflies and amusement parks. An assessment was administered on Thursday and the results were mixed (See Appendix: charts F-O for all quiz results) with four students scoring 90% or higher, two students scoring in the 80% range, one student scoring in the 70% range, and two students scoring failing grades. These two students missed both of the multiple-meaning questions. One of them missed two of the three multiple-choice questions about character, setting and plot. The other missed the multiple-choice question about setting and missed two of four points on the extended response because they did not state three examples.

Using the information from the assessment, my Friday lesson plans included reviewing the assessment to review where the students had found success and where they needed additional skills practice. First, we completed a review of multiple-meaning words using a thesaurus based on what the students had missed. We completed additional practice using a reading one selection about Monarch butterflies and another comical story about middle-school students at an assembly.

Week Two instruction revisited story elements focusing on the skill the students had struggled with the week before: setting. The vocabulary skill for this week was

compound words. The main reading selection was a realistic fiction adventure story about the discovery of Machu Picchu. Due to the difficulty level of the vocabulary in this story (both Spanish and Incan words are used), I chose to have the students listen to the story as a whole group and then reread it with a partner of my choice. Additional selections included a non-fiction selection about the Inca and another adventure story set in Egypt.

The results of the weekly assessment was again mixed. Two students had perfect scores, but then the scores dropped into the eighties. Two students scored 83%, five students scored 73% and one student had a failing score because he chose not to complete the extended response. The students all missed the same two compound word multiple-choice questions. I had noted these two questions had proved to be problematic the last couple of years I have administered this assessment. The other sixth-grade Language Arts teachers find these two questions to be extremely difficult for students as well. I believe the question is very poorly worded and the answer choices are also poorly worded.

Our Language Arts Department decided these two questions are not valid or reliable and we will not use them in their current form in the future. When these two invalid questions are removed, the scores increase dramatically. There are now four students scoring 100% and five students scoring 90% or higher. The student who initially failed the assignment would raise his score, but would still be in the failing range. Even though I determined the errors on the assessment were more an issue with the assessment than the students, our Friday lesson consisted of a review of compound words.

Week three skills included a review of context clues and the introduction of main idea and details. Our reading selection for the week consisted of a short non-fiction story about how scientists studying animals to invent things to help humans. We also read

selections about cockroaches and spiders. Main Idea and detail was a new skill, so I modeled finding the main idea using stories about dolphins and penguins.

That week's assessment proved to be a great achievement. Three students scored 100%. Four students scored in the 90% range. Two students scored in the 80% range. The lowest score was a 73%. This student missed both context clues and main idea and detail multiple-choice questions, but had a perfect score on the extended response, which I found to be puzzling.

With these successes, our Friday lessons consisted of independent work. I created individual skill packets for students based on their assessment performances. During this independent work time, I had a conference with each student to be able to discuss their progress so far and touch base on how they thought they were progressing. During our discussions, each student said they liked the stories we had read so far. Three students (F, G and H) said they had looked ahead in the reading book to see what we were going to read next. The student who stated she hated non-fiction (Student B) said what we had read so far wasn't 'too bad' and she liked reading about using spider webbing in bulletproof vests. Not all of the conferences were completely positive. A student who had expressed he did not enjoy reading aloud (Student H) stated he still did not like reading aloud and didn't want to do it. We worked out a deal where he had one free pass per week, but that was it. He could only opt out once. He was satisfied with this compromise.

The student who had refused to complete an extended response question (Student J) also proved to be a difficult conference. When I inquired as to why he had chosen not to complete the response even after being asked to do so, he stated he didn't feel like it and thought he would see what I would do. I asked him to tell me what I did when he

refused, he told me I simply reminded him it was his grade he was hurting and I wouldn't know for my records if he could do it or not and walked away. I asked him what he thought of his decision now. He stated he wasn't sure, but he couldn't believe I didn't yell at him and make him do it. I explained that I'm not a person who would do such a thing, and at twelve-years old he should be able to make such decisions to complete an answer on a test since the consequences all fall on him. He looked at me strangely and asked if we were done. I haven't completely figured out what he was trying to learn about me, but he has not refused to complete anything since our conversation.

Week four skills included more context clues, but this time with restatement. Our reading skill was cause and effect. Our story selection was a fun folktale about a rabbit and a chameleon. Folktales were also part of Social Studies this week as they read about oral tradition.

We read the story selection as reader's theatre, so we did not read any other selections in Language Arts. There were additional folktale selections read in Social Studies including the story of Gilgamesh and an excerpt from Beowulf.

I wasn't sure what to expect with this week's assessment since I had done little modeling and nearly all of the student's practice was based on preparing for a performance. There were two students who scored 100% and five more who scored in the 90% range. Two students scored in the 80% range, but their scores were lowered because they did not answer all of the parts of the extended responses. One student scored 73%. He was 0/2 on the context clue multiple-choice and missed a cause and effect question. He did choose the second best answer for both context clues questions, and this has not been an area of concern on past assessments. When we conference, he said he was having

an off day. When I asked him to read over the selection again and correct his answers, he was able to with 100% accuracy. I noted this in my observational records for this student to be aware of this in the future.

Week five skills again reviewed context clues and main idea, but introduced synonyms. We had ‘fun with the Thesaurus’ lessons to work with synonyms. This week’s spelling lesson also focused on synonyms, so the students had even more practice with them. The reading selections were non-fiction selections about conservation of animals. We first read about the Clouded Leopard in Thailand and then the Leatherback Turtle.

That week’s assessment was surprising to me after the success the students had been experiencing. There was only one student with 100% and two students who scored in the 90% range. Three students scored in the 80% range and one in the 70% range. I was surprised by the three students who scored in the 60% range, a failing score in our district. All three students struggled with one particular context clues question and the main idea and detail multiple-choice questions. Two of these students only scored 2 out of 4 on their extended responses because they did not include enough details. These two students have struggled with extended responses in the past, so I decided to work with them again in a small group to help them with their extended responses.

Typically if students had not achieved to my expected levels, I would have retaught the lesson and we would have worked together on the skills. Instead, I decided to take a step back and allow the students the opportunity to practice independently. I presented the class with three reading selections all with the same theme of conservation, but each story was about a different animal. Students were able to choose what they

wanted to read about, and then they could choose someone to work with who had the same selection. As the students worked, I made notes about their work habits and ability

My district requires short-cycle assessment be administered every six weeks. As a Language Arts department, we decided the Unit tests provided by the textbook company that already match the skills we are teaching in our classrooms and are modeled after the OAT format. Before I administered the assessment, I surveyed the students about how they felt about their skill level thus far. They were to rate their skill level as ‘good’ meaning they were ready for the assessment, ‘maybe’ meaning they were comfortable completing the skill, but would like more review, or ‘help’ meaning they were not at all comfortable with this skill.

All of the students except one rated their main idea and detail skills as ‘good.’ That one student (Student J) rated their skills as a ‘maybe’ was the student who had struggled with this skill on previous assessments, so I was pleased he recognized a need. Context clues skills were rated pretty high with eight ‘goods’ and two ‘maybes’. Again, the student who had struggled all along with context clues chose ‘maybe’ and the other ‘maybe’ came from the student who had done well with context clues up until the last assessment. When rating multiple-meaning words, there was a near even mix of six ‘goods’ and four ‘maybes’. Compound words were given all ‘goods’ from the students. Cause and effect ratings were a surprise because only five students rated it ‘good’ while four students rated it a ‘maybe’ and there was one ‘help’ (Student J). The scores on the assessment with cause and effect were among the highest all school year. When I asked why this was an area of concern, the students responded it was because we only spent one week on cause and effect, so they didn’t feel as comfortable.

I used these surveys to plan our review. An obvious area of focus was cause an effect followed by multiple-meaning words, compound words and context clues. We reviewed skills for two class periods working with the whole group, partners and small groups. I chose stories about spiders, ants, the Great Lakes (we live on Lake Erie), a basketball team and a school play.

The scores from the Unit test were an overall success, but there were some surprises. Five students scored in the 90% range. Two students scored in the 80% range. The last two scores were in the 60% range. One of the scores was a surprise because the young man had done well on previous assessments. Upon further review, we realized he had made a mistake on his scan card and put two scores on one line throwing off five answers. When this was rescored, his score increased to a 77%, still below what I believe he is capable of, but an improvement nevertheless.

The other student scoring below passing was Student J who has struggled all school year. It should be noted that this student has an IEP for a learning disability related to reading comprehension, fluency and an assortment of behavior issues. While this research was taking place, this student was re-evaluated and it was determined he had been incorrectly placed in a regular education setting and should be serviced in a resource room. He will be removed from my class and will be placed in a resource room after Winter Break.

Following the Unit Assessment evaluation, I determined the class as a whole still needs to work on multiple-meaning words, context clues and main idea and details. There were three multiple-meaning questions on the test and each student missed at least one of them. There were five context clue multiple-choice questions and each student missed at

least one of them. The most commonly missed context clue question had two answers which both worked when placed back into the sentence. There were three main idea and detail multiple-choice questions and each student missed at least one of them as well.

Switching gears from a different story selection each week, I introduced a novel, Sign of the Beaver by Elizabeth George Speare. Despite departing from the reading textbook, I am still using some of the materials from the reading series. I still focused on the skills I outlined at the beginning of the year as areas of need.

Unit Two, Week One skills included problem and solution and context clues with restatement again. The class evaluated the different problems the main character, Matt, faced and how he dealt with them. They also drew some conclusions about Matt based on how he solved his problems.

Because of the amount of reading involved, our weekly assessment will no longer be on a set day. This week, the assessment was administered on Friday. The same type of skills' assessment was used to evaluate students' skill level as before the novel reading. In the first assessment, students did not fair well. There were only two students who scored in the 80% range. Four students scored in the 70% range. Three students scored in the 60% range. The problem and solution multiple-choice questions proved to be a problem. All ten students missed at least one of these questions and seven of the students missed 2 out of the 3 questions. The biggest issue with this assessment was the extended response. Six students missed 2 of the 4 extended response points because they identified the problem and the solution, but did not give a detail for each. Using context clues with restatement was not a problem for students on this assessment. Of the four vocabulary

questions on the assessment, only three students missed one of these multiple-choice questions.

Before moving onto the next part of the novel, we reviewed the quiz and students made corrections on the extended response. We read a selection about the Stone Ages and an archaeologist to review problem and solution. The practice assignment was about a Native American boy. The assignment was formatted to look exactly like the assessment. The students did exceptionally well. On all ten multiple-choice questions, there were only two students who didn't score 100%. Those two students only missed one question. I do not have any conclusions to explain the discrepancies in scores.

Week Two skills included main idea with details and idioms. The weekly assessment for this section was given on a Wednesday. Student scores improved considerably. Four students scored in the 90% range. Four students scored in the 80% range. Two students scored in the 70% range. The one question that the students had difficulty answering was a detail question that asked which answer would not be included. These types of questions have traditionally been difficulty for students.

There are no specific skills from this section that warranted immediate reteaching. I do plan to revisit idioms again in the future because this lesson is one of the first times students have been introduced to this skill.

Week Three skills included inflectional endings and cause and effect again. Inflectional endings lessons were addressed during our grammar study. At this part of the novel, there is a great deal of action, so finding causes and effects was not difficult.

The assessment for this section was given on a Thursday. The students did very

well on this assessment. Two students scored a 100% on the quiz. Four students scored in the 90% range. Three students scored in the 80% range. Only one student scored in the 70% range. No one missed any of the inflectional ending questions. There was not one multiple-choice cause and effect question in particular that proved to be the one that was problematic. The error made by seven students was not completing the extended response question with a cause and a detail and an effect with a detail.

The last week of the novel reintroduced synonyms and introduced making generalizations. The weekly assessment for this section was given on Thursday. The students scored very well. Four students scored in the 90% range and the other six students scored in the 80% range. Seven students missed one of four synonym multiple-choice questions. Three students missed one of the three making generalizations questions. Five students missed a point on the extended response because they did not give all three details.

Before administering the novel/unit 2 skills' assessment, I surveyed the students in the same manner I did for the first unit/short-cycle assessment this time I focused on the skills identified as areas of concern at the beginning of the school year: context clues, problem/solution, main idea with details and cause and effect. This time eight students rated context clues 'good' and two reported 'maybe' ratings. Students were a bit more hesitant about problem and solution. Four students rated it as a 'maybe', leaving only six students to give it a 'good' rating. Main Idea with details were rated with seven 'goods' and three 'maybes.' Cause and effect was rated with nine 'goods' and only one 'maybe'.

The selections for our review were all non-fiction since we had primarily been

reading historical fiction for four weeks. We reviewed the different skills for three days before I administered the assessment.

The students' achievement on the assessment was acceptable. Seven students scored in the 80% range and the other three students scored in the 70% range. The only wide-spread area of concern was a making generalizations extended response in which eight of the students only scored one out of two points because they didn't answer both parts of the questions, a detail multiple-choice question that seven students missed.

My final piece of data was administering an additional Benchmark Reading test. This test includes a wide-variety of skills including making generalizations, using context clues, idioms, inflectional endings, problem/solution, main idea with details, cause and effect, synonyms, and story elements.

We reviewed these skills for two days before the assessment. The students' Performance (See Appendix, chart P) was satisfactory. Three students scored in the 90% range, six students scored in the 80% range and one student scored in the 70% range. The only questions that were decidedly problematic was one character multiple-choice question that was an 'all of these except' question six of the students missed, a multiple-meaning multiple-choice question five of the students missed, and a multi-step cause and effect extended response seven students did not answer both parts.

During my research, I completed detailed analysis of the students' fifth-grade OAT scores, a score from a Reading Benchmark Test, and weekly skills' quizzes to determine specific deficiencies in the students' reading skills. In addition to the assessment models, I surveyed the students to glean information about their interests outside of school and their attitudes towards reading. I used this list of deficiencies and

the students' interests to drive my instructional plans. Once these plans were in place, I chose reading materials best suited the students' interests.

I determined that there was value in completing the detailed analysis because I changed the placement of students to be able to address their needs with other students who had the same instructional needs. This allowed me to differentiate my instructional plans for maximum student achievement. Although not all of my students had a passing score on each assessment, there was definitive improvement in the students' scores.

Action Plan

My action plan is to continue with this data analysis. The work involved in analyzing the data from the OAT is daunting and I was frustrated by the lack of timeliness on the part of the State of Ohio in releasing the detailed results of each child's test. That being said, I cannot imagine beginning a school year without knowing this information about my students. I will complete this analysis for all of my students in the future and place my students based on their skill level and not just their score.

If I were to change something, I would devise an organizer for the students so they can keep track of their progress, so when I survey the students about their achievement, they can review their weekly assessments and accurately evaluate their skill levels and areas of need. The students would also be able to take more responsibility for their learning by being able to accurately tell me what they need to continue to practice.

I also need to continue to re-evaluate the data entry process I use in order to attempt to lessen the amount of time I spend compiling all of this information. At times, the data analysis can be overwhelming and I may not be able to keep such detailed records and have the time to teach, too. Having a more efficient way to compile

information would help to give me more time to spend planning my instruction to benefit my students.

Currently, all of the teaching teams in my building use different methods for sorting students into class rotations. I would like to be able to share how much more effectively I was able to start my instruction and truly hone in on what my students needed because I had put in the time completing the analysis. Being able to focus on student needs has lead to increased scores on weekly assessments and short-cycle assessments. I am also hopeful I will see improvement in my students' OAT scores as well. If other teachers in my building embraced this practice as well, I think it could make a positive change in their teaching practices.

The Director of Curriculum in my district is my also former principal, and given that he has always been a proponent of data-driven instruction, he has been very interested in the final outcome of this research. I have already presented this plan and my data to the Language Arts team in my building. If there is an improvement in these students' OAT scores, then our Curriculum Director is going to have me present to the district's Language Arts Curriculum Committee.

Finally, I plan to submit this plan to an educational journal as an alternative method for dividing students into ability groups.

References

American Counseling Association. (2004). Position statement: High-stakes testing. Retrieved May 18, 2009, from http://schoolcounselor.org

Anderson, K. M. (2007). Tips for teaching: Differentiating instruction to include all students. *Preventing School Failure, 51*(3), 49-54. Retrieved May 14, 2009, from http://www.heldref.org

Educational Sources. (2004, November 7). Math and reading ability grouping in elementary schools. In *applesfortheteacher.com.* Retrieved May 18, 2009, from http://www.apples4theteacher.com

Jerald, C.D. (2005). Using data: The math's not the hard part. *Journal for Students Placed at Risk*. 10(7), 295-308.

Mills, G. (2007). *Action research: A guide for the teacher researcher* (3rd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Mitchell, R. (2006, March 30). Effects of high-stakes testing on instruction [Review]. *The Center for Public Education.* (2006, May 30). Retrieved May 13, 2009, from http://www.centerforpubliceducation.org

Muir, M. (2002). Tracking and ability grouping [Research brief]. *The Principals' Partnership*. (2007, January 8). Retrieved May 18, 2009, from http://www.principalspartnership.com

Robinson, J.P. (2008). Ability grouping benefits some. *Educational Evaluation and Policy Analysis*. (2008, June 12). Retrieved May 18, 2009, from http://www.teacheffectively.com

Schullery, N.M., & Schullery, S.E. (2007). Are heterogeneous or homogeneous groups more beneficial to students? *Journal of Management Education*. *30*(4), 542-556. Retrieved May 13, 2009, from http://www.sagepublications.com

Stringfield, S. & Wayman, J.C. (2003). Teacher-friendly options to improve

teaching through student data analysis. *Center for Social Organization of Schools, John Hopkins University*. Retrieved May 18, 2009, from http://www.csos.jhu.edu

Teachscape, Inc. (Producer). (2007). Expert commentary: *How can I identify an area of focus for my research?* [Streaming Video]. San Francisco, CA: Producer.

Teachscape, Inc. (Producer). (2007). Expert commentary: *How do I implement an intervention as part of my research project?* [Streaming Video]. San Francisco, CA: Producer.

The National Center for Fair and Open Testing. (2001, January 1). Testing plus real accountability with real results. *FairTest*. Retrieved June 2, 2009, from http://www.fairtest.org

Tieso, C. L. (2003). Ability grouping is not just tracking anymore. *Ability Grouping is Not Just Tracking Anymore*, *26, The Questia Online Library*. Retrieved May 13, 2009, from http://www.questia.com

Westchester Institute for Human Services Research. (2003). High-Stakes testing. *The Balanced View*, 7(1).

Appendix

Chart A: Data Collection Matrix

Research Questions	Data Sources	Data Sources	Data Sources
1. What effect does examining students' individual score results have on student groupings?	Ohio Achievement Test in Reading	Benchmark-reading test	Student survey about OAT.
2. What effect does differentiated instruction have upon students' performance levels in their individual areas of need?	Reading comprehension survey about perceived areas of need.	Weekly skills practice *Modeled instruction *Independent practice (homework)	Weekly skills assessments and Short-cycle assessments Unit surveys about successes and perceived areas of need.
3. What effect does placing students in ability groups have upon their self-esteem?	Student survey about grouping	Student interviews	Anecdotal notes and observation of students

Chart B: Biographical Data

Student A	Girl, 11-years old. Discharged from Speech Therapy in Fall 2008. Final 5th grade Language Arts grade: B-.
Student B	Girl, 11-years old. Final 5th grade Language Arts grade: B.
Student C	Girl, 12-years old. Final 5th grade Language Arts grade: A.
Student D	Girl, 11-years old. IEP for reading comprehension and written expression. Final 5th grade Language Arts grade: C+.
Student E	Girl, 13-years old. Retained previous year. Had attended one year of public school before attending an Arts Academy for three years and was home schooled for two years. Entered district as a 6th grader, but skill level warranted placement in 5th grade class. IEP for reading comprehension and written expression. Final 5th grade Language Arts grade: B.
Student F	Boy, 12-years old. Diagnosed with Autism and Aspberger's Syndrome. IEP accommodations for reading comprehension and written expression. Occupational Therapy for handwriting. Final 5th grade Language Arts grade: B.
Student G	Boy, 11-years old. IEP for reading comprehension, written expression and organization. Occupational Therapy for handwriting. Final 5th grade Language Arts grade: B.
Student H	Boy, 12-years old. IEP for reading comprehension and written expression.

	Final 5th grade Language Arts grade: B.
Student I	Boy, 11-years old. Father passed away when he was six-years old and Mom has Stage 4 cancer. Frequent absences. Final 5th grade Language Arts grade: D.
Student J	Boy, 12-years old. IEP for reading comprehension, written expression and spelling. Will be moved to Resource Room setting for Language Arts and Math in January 2010. Final 5th grade Language Arts grade: C.

Chart C: OAT Scores

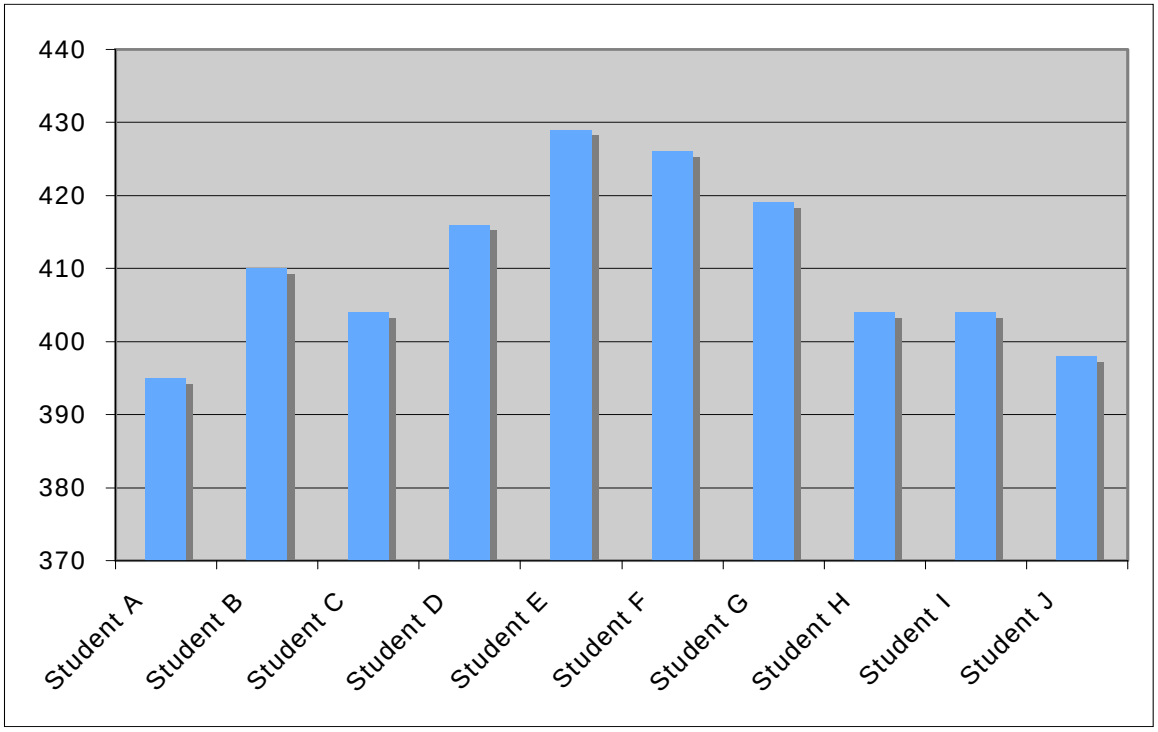

Chart D: Reading Benchmark A scores

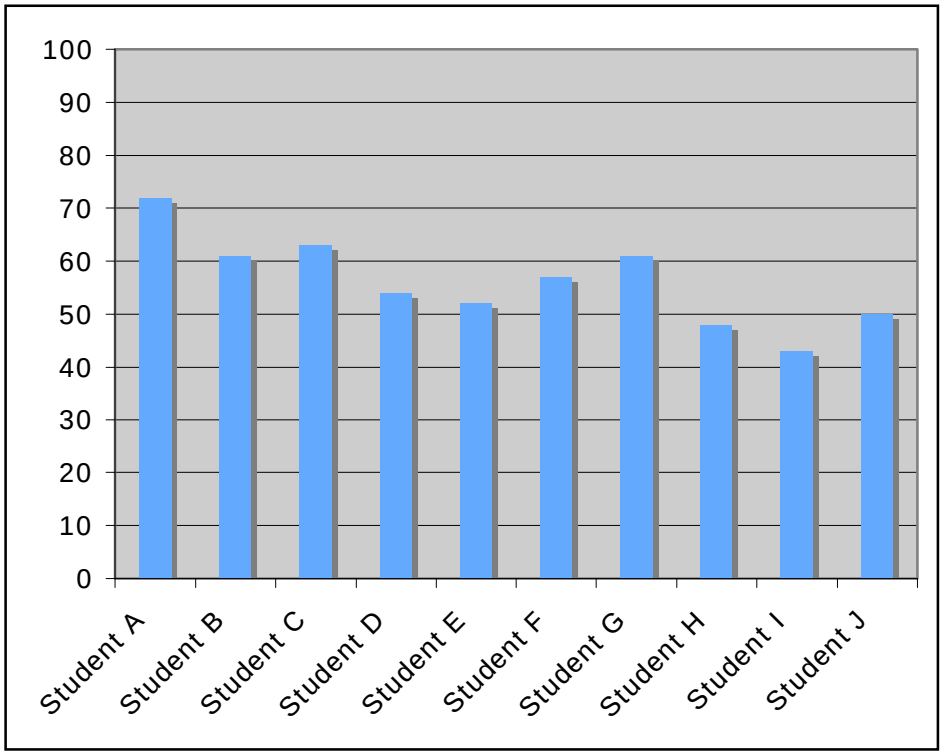

Chart E: Student Interest and Hobbies

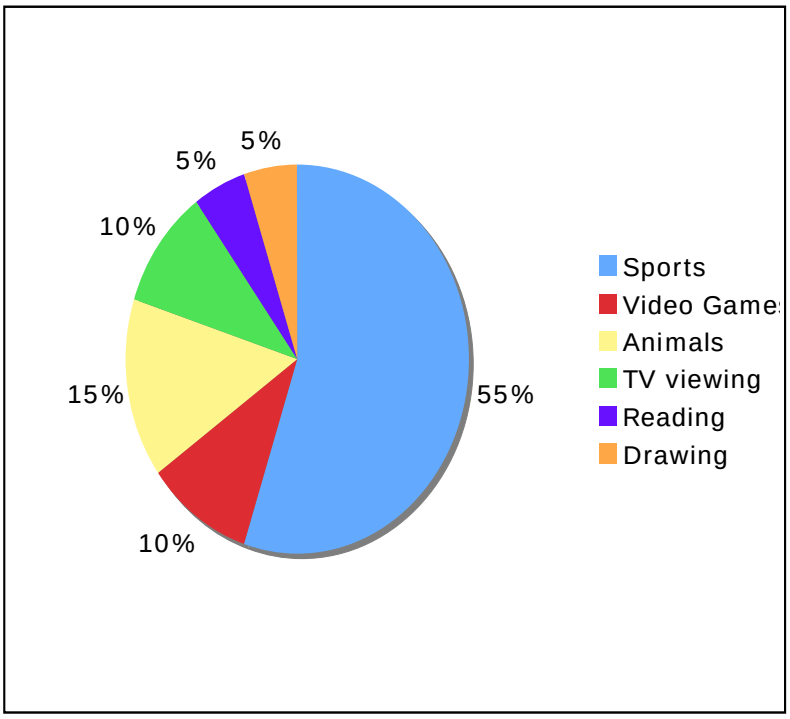
5%
5%
10%
15%
10%
55%
Sports
Video Game
Animals
TV viewing
Reading
Drawing

Chart F: Assessment Scores: Student A

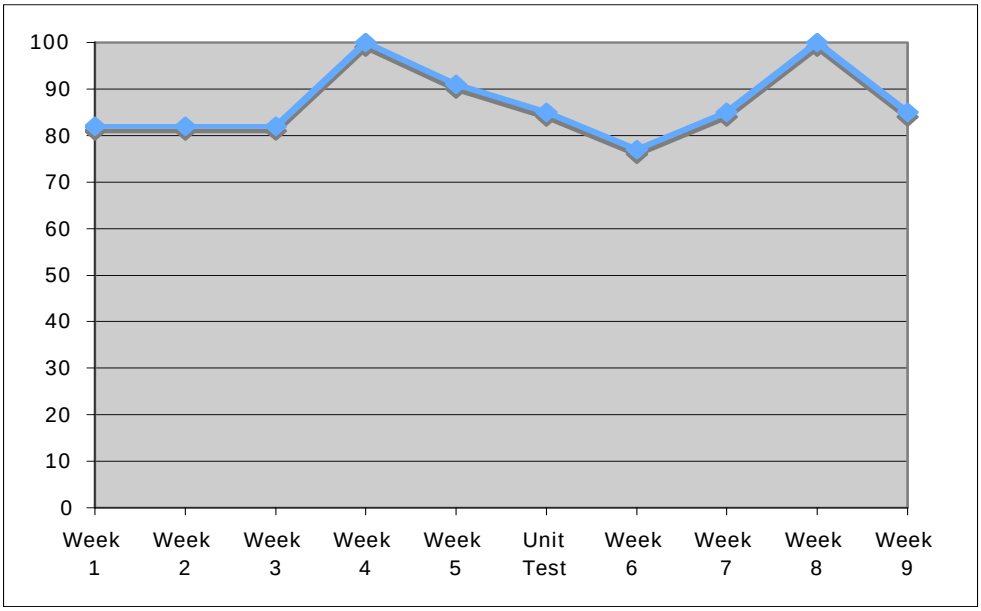

Chart G: Assessment scores: Student B

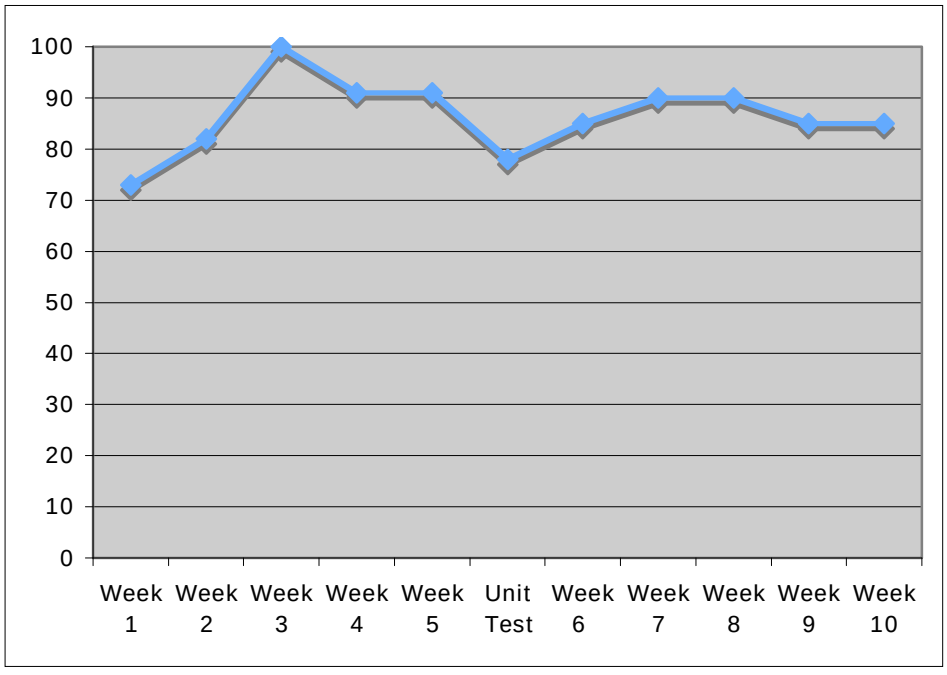

Chart H: Assessment scores: Student C

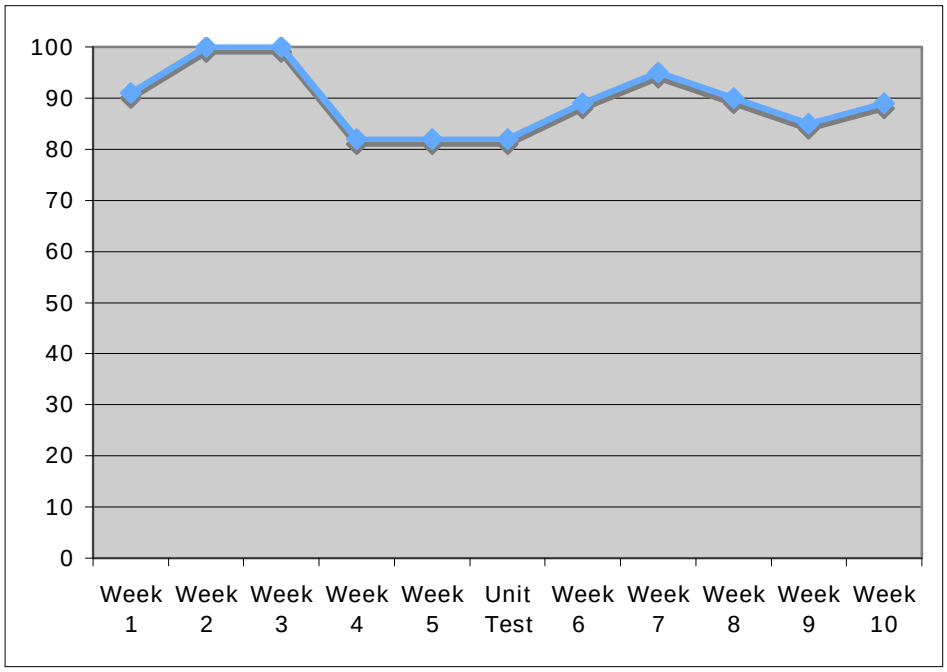

Chart I: Assessment scores: Student D

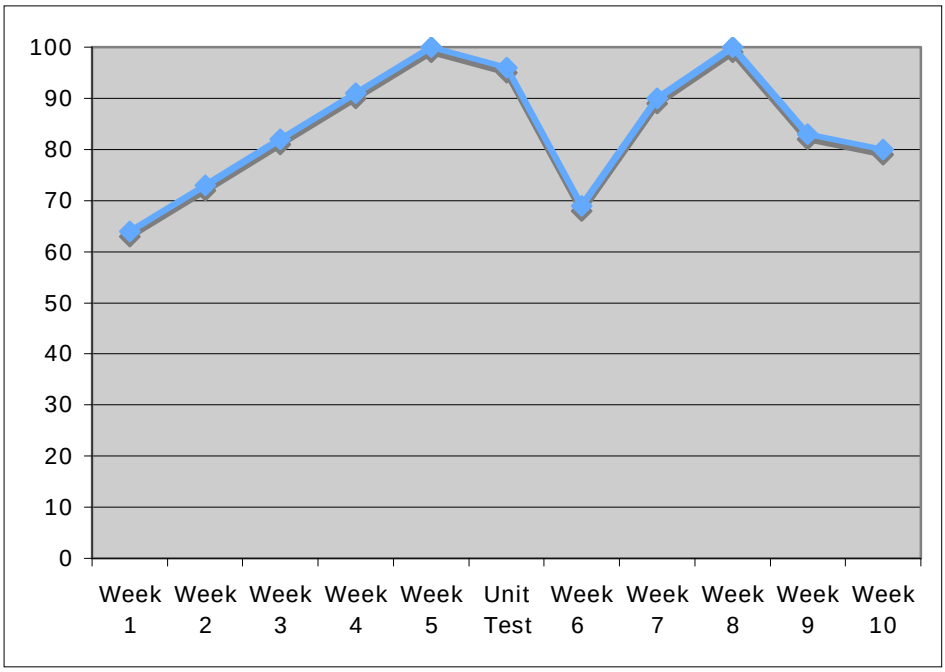

Chart J: Assessment scores: Student E

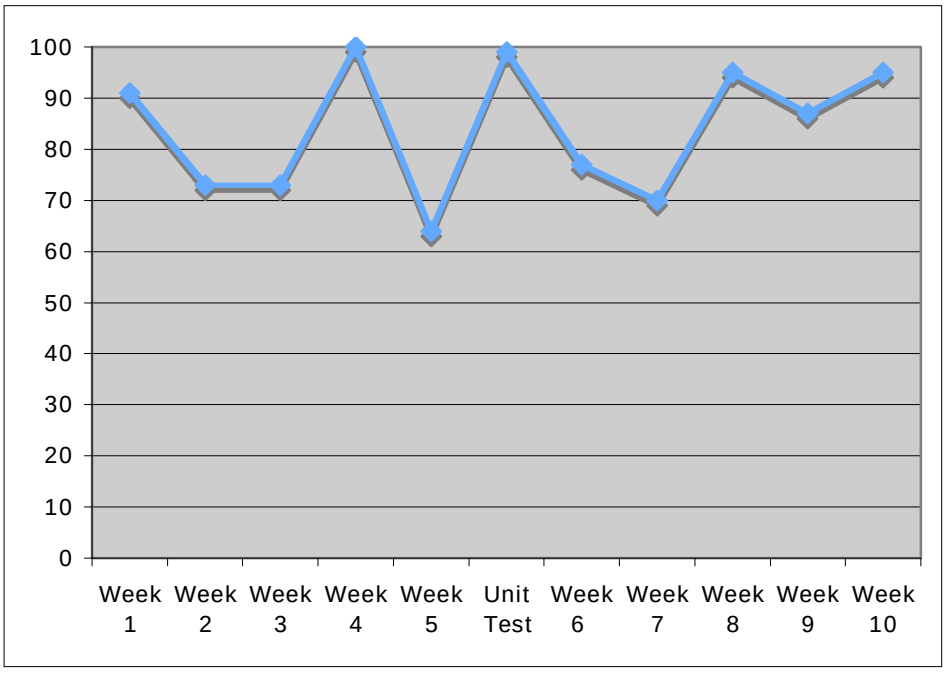

Chart K: Assessment scores: Student F

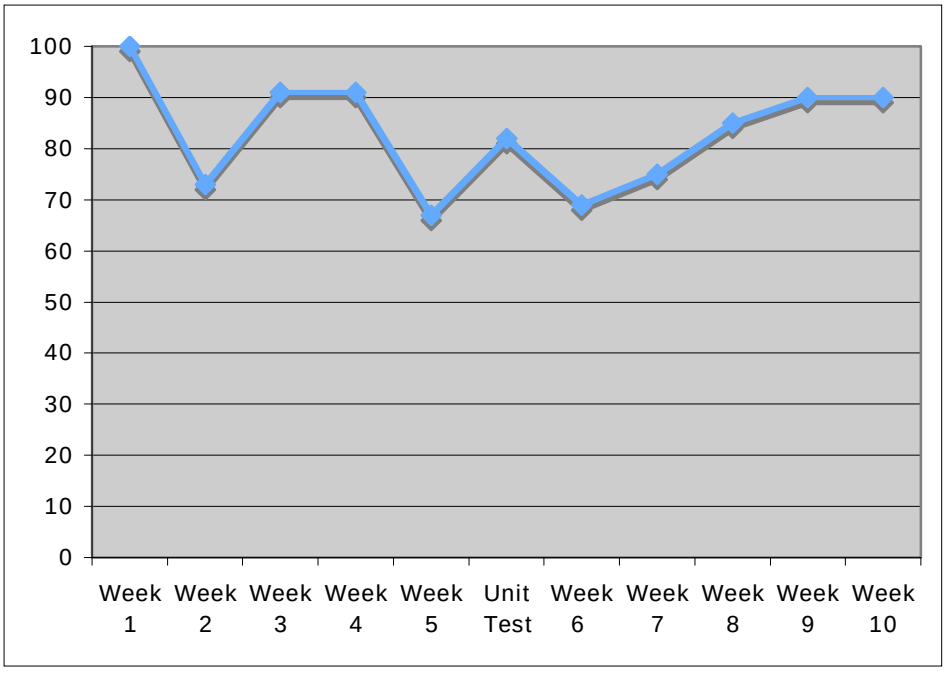

Chart L: Assessment scores: Student K

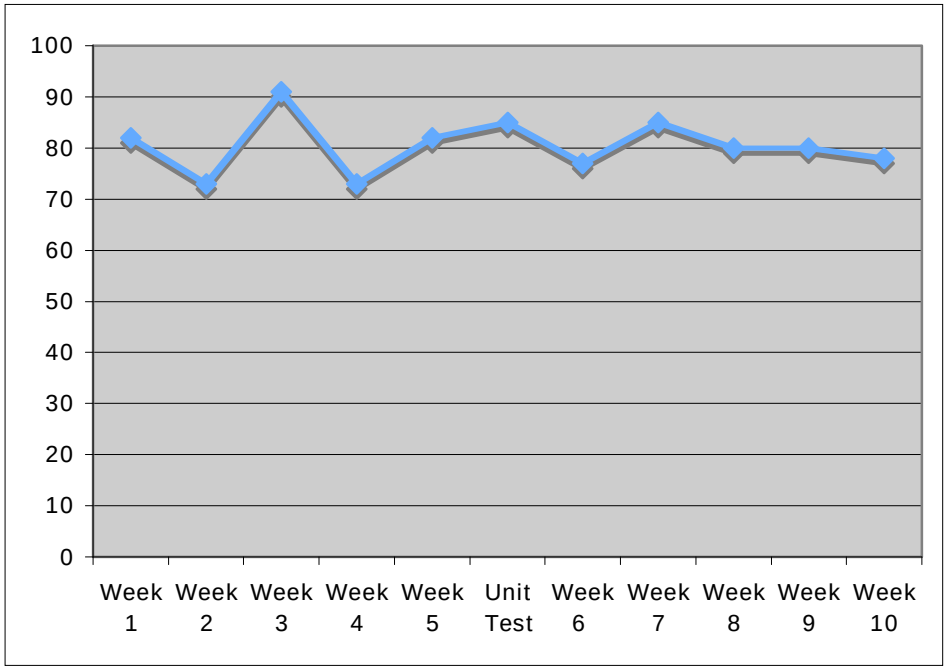

Chart M:
Assessment scores: Student H

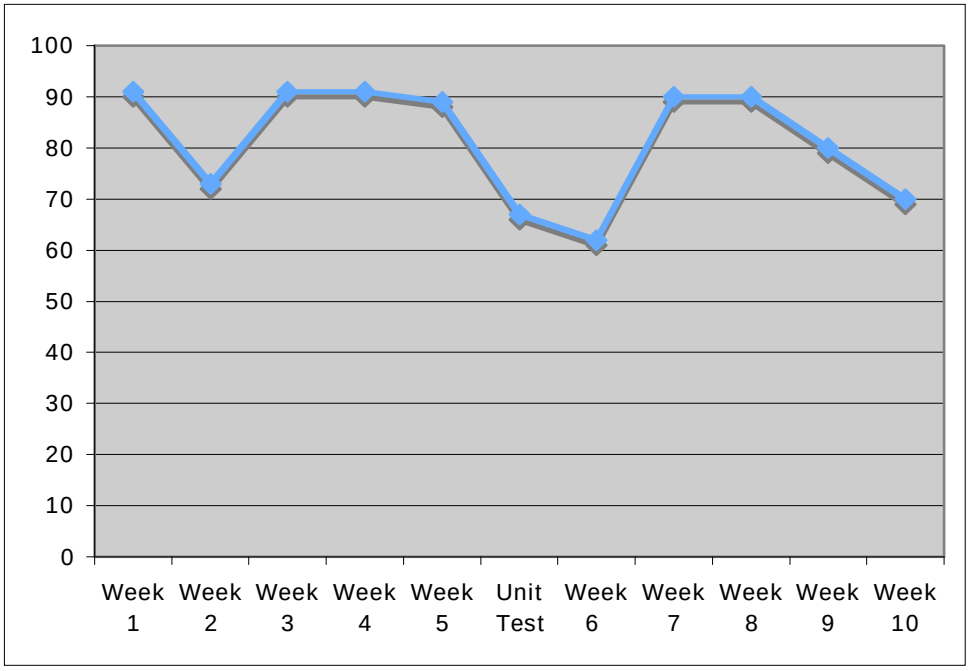

Chart N: Assessment scores: Student I

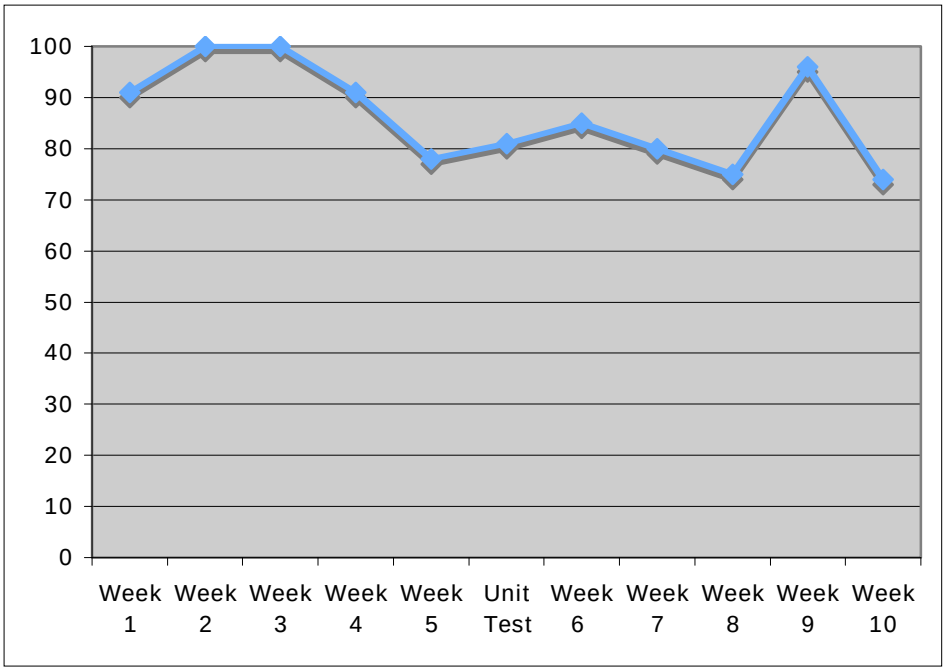

Chart O: Assessment scores: Student J

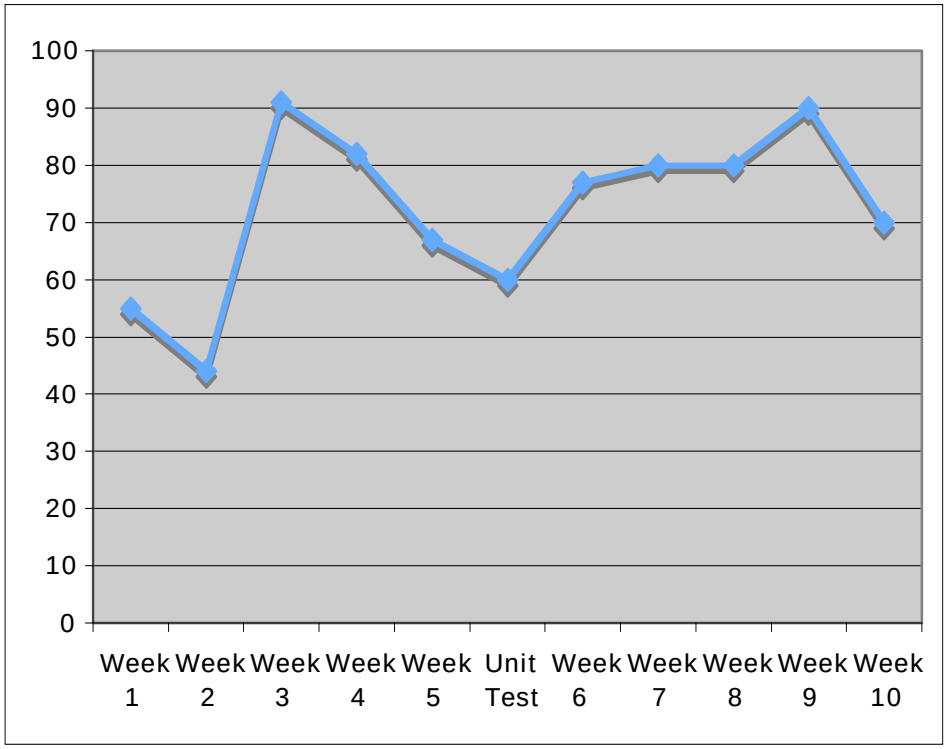

Chart P: Final Benchmark Test