CIMLE *Current Issues in Middle Level Education* (2010) 15(1), 10-18

Northwest Evaluation Association's Measure of Academic Progress (MAP): To What Extent Does the Use of MAP Assessment Data Affect Students' Language Usage Skills?

Danette Asperheim

The purpose of this study was to conduct action research on the use of assessment data to plan and instruct a middle level group of ten students between the ages of 12 and 13 years old. The research analyzed the Northwest Evaluation Association's (NWEA) Measure of Academic Progress (MAP) assessment data relative to student strengths and weaknesses exclusively dealing with language usage skills. The research attempted to determine whether or not use of immediate data from the MAP assessments to plan instruction according to their individual needs (differentiation) did in fact improve cognitive understanding. Some students' scores increased dramatically from the fall to the spring assessment, while others remained at the same level.

Relevance of Study

Differentiation is a big buzzword among educators these days. Creating lessons according to student interest, readiness or intelligences seems like a logical process; so why is it that educators, myself included, often times think that differentiation is an insurmountable obstacle comparable to moving mountains? Why do we shy away from reaching kids according to their interest, readiness, or intelligences? Is it because we always seem to be short on time, grades due by four o'clock or a parent meeting is scheduled during our planning time? Regardless of how busy we seem to be, is differentiation possible?

I believe it is imperative to create a differentiated classroom to adequately meet the emotional, social, and physical needs of the students. When differentiated lessons include hands-on manipulation, interaction, and exploration of the world, all students can be successful learners and when taught using effective differentiation strategies that focus and encourage their learning, their potential is limitless. That is why I conducted the following study. I wanted to see if differentiation made a significant difference in standardized test scores, student behavior, and ultimately, student success with language usage skills.

The following study highlights the related literature, the process of analyzing and implementing data from the Northwest Evaluation Association's (NWEA) Measure of Academic Progress (MAP) assessment, and the interventions I implemented with my students. It is my hope that educators and administrators will find manageable techniques that help them use standardized tests to facilitate differentiation in classrooms around the country.

Review of Related Literature

Assessment is one of the most critical facets of one's curriculum. Effective assessment is an essential part of effective teaching and is an important aspect of effective learning. "Classroom assessment and grading practices have the potential not only to measure and report learning but also to promote it" (McTighe & O'Connor, 2005, p. 11). Assessment should "measure performance in ways that will enrich learning, rather than restrict it" (Robertson & Valentine, 2000, p. 1). Therefore, if assessment isn't used or isn't used properly, it is extremely hard to gauge whether or not the instruction is valuable and whether or not learning and understanding is actually being achieved.

Many researchers of education believe that the validity and accuracy of assessments enables educators to determine the level of understanding of the students. Summative assessment is often the most used form of assessment to gauge student understanding of a main concept.

> Although evaluative [summative] assessments command the attention of students and parents because their results typically 'count' and appear on report cards and transcripts, by themselves, summative assessments are insufficient tools for maximizing learning. Waiting until the end of a

> teaching period to find out how well students have learned is simply too late. (McTighe & O'Connor, 2005, p. 11)

Therefore many practitioners use formative assessment in their everyday instruction, as they believe the students will benefit and gain more understanding of the content; because, the instruction is continuously being adapted according to their needs (which are being evaluated/analyzed from in-class assessments). Traditional formative assessments continue to be one of the ways teachers assess this learning. Although, "a standardized [traditional] approach to classroom assessment may be efficient, it is not fair because any chosen format will favor some students and penalize others" (McTighe & O'Connor, 2005, p. 15).

Research and common sense show that some students simply aren't good at taking traditional paper and pencil tests. Instead, students often feel more comfortable expressing their knowledge and understanding through presentation, discussion, demonstration and additional performance based tasks (Hurren & Rutledge, 2005).

> If students are able to share their ideas by communicating with classmates and they are able to demonstrate their understanding through projects, why is it that when it comes to the summative assessment, they are expected to keep their mouths quiet and their creativity hidden and complete a paper and pencil test? (Hurren & Reutledge, 2008, p. 20).

Many students falter at the summative assessment portion and as stated previously, the summative assessment is the part that "counts." Jackson and Davis (2000), supporters of the middle level education, encourage educators to vary their assessment techniques to ensure that all of the students' interests, abilities, and intelligences are being targeted. Some of the suggested assessment practices include portfolios, oral conferences, interviews, and performance tasks and projects (Jackson & Davis, 2000).

> Yet, the trend in today's educational world is the movement towards standardized testing. Standardized or high-stakes testing is but a part of the middle school experience. The [standardized] tests are only a small part of the assessment process; whereas classroom assessment is a very large portion of the assessment and a portion that is geared towards what is happening in the classroom and the needs and learning abilities of the students now. (Chappuis & Stiggins, 2008, p. 15).

Even though classroom assessment is the largest form of assessment used, the emphasis on standardized testing is becoming more and more for state officials, administrators, teachers, and parents and unfortunately the traditional nature of state testing doesn't correspond to the needs of middle school students (Jackson & Davis, 2000). Even though authentic assessment matches more so with the needs of the students, it is undeniably difficult to assess at the state level. Since authentic assessments are typically "open-ended and do not yield a single, correct answer or solution process, assessors cannot score student work using a simple answer key" (McTighe & O'Connor, 2005, p. 13). In addition to the difficult process of scoring authentic assessments at the state level and even though "authentic assessment is less costly to develop," it is "two to three times more expensive to administer and score" (Robertson & Valentine, 2000, p. 2).

Although standardized testing may continue to have its place, it must also continually be examined and none of the research that has been conducted up until now definitively supports the positive correlation between one specific form of assessment and positive cognitive gains. By and large, school officials, administrators, and teachers' main concerns lie in the academic performance of their students. They are constantly seeking new ways to ensure that children are learning and retaining the information presented to them. Yet, the question remains: does state testing truly measure students' cognitive understanding?

Robertson and Valentine (2000) argue "standardized tests typically do not test higher-order thinking skills" (p. 2); therefore deep, cognitive understanding cannot be measured on such a test. It is clear that standardized tests do not take into account individual learning styles, skill levels, and learner interests; yet, assessment in the classroom does in fact account for all of those things. Proponents of middle level education believe "learner-centered classrooms and alternative

assessments go hand in hand" (Robertson & Valentine, 2000, p. 1) more so than that of the traditional standardized means of assessments.

Another disadvantage to the traditional standardized testing process is that it takes so long to receive feedback. According to an interview (2008) with an official from the State Department of Public Instruction's Administrative Assistant in the Standards and Achievement Unit, more often than not, school districts do not receive the results of state testing, which are conducted in October, until the end of February or beginning of March. That is well into the year, long after educators are able to use the results to enhance the learning of their students. Teachers are instructed to give specific and immediate feedback to students as soon as possible so that they are able to make the necessary adjustments to their learning. So then why is it that teachers settle for anything less than immediate feedback themselves?

Several organizations have tried to tackle that specific problem—delayed feedback. The Northwest Evaluation Association is a non-profit organization that provides products and services to measure and promote academic student growth and school improvement. These include accurate assessments (state-aligned computerized adaptive assessments that provide accurate, useful information about student achievement and growth) and timely reporting (NWEA, 2008). With timely reporting, administrators and educators have the ability to adapt the curriculum to meet the immediate needs of the students. They are able to use the data to target the students' weaknesses and build upon their strengths, unlike when they receive state testing results so late into the school year.

In a study conducted by Cronin, Dahlin, Adkins, and Kingsbury (2005), the researchers examined and analyzed the fact that several "state tests that measure academic progress are creating a false impression of success in the early grades" (p. 168). The "false impression" was highlighted in the data linked from the 2004 and 2005 North Dakota State Assessment (NDSA) and the Northwest Evaluation Association's Measure of Academic Progress assessment. The researchers found that NDSA expectations are not "smoothly calibrated across grades" so students who "are proficient in third grade are not necessarily on track to be proficient by the eighth grade" (Cronin et al., 2005, p. 169). With such research and the "false impressions" given by the assessments themselves, many have begun to question the validity and accuracy of standardized state testing. If state testing isn't "smoothly calibrated across grades" (Cronin et al., 2005, p. 169), how is it possible for educators to ensure that they are meeting the standards and benchmarks for a specific grade level that will lead to student mastery in future grade levels?

Due to the inconsistent and inconclusive nature of the previously conducted research, the purpose of this research study was to conduct an investigation to examine the research question—"To what extent will the use of MAP assessment data affect students' language usage skills?" This study will attempt to determine whether or not the use of the immediate data from the NWEA's MAP assessments, which are aligned with the state standards and benchmarks, improves cognitive understanding of language usage skills for seventh grade students in an English classroom setting.

Methodology

This research was conducted within a normal educational environment, utilizing an action research approach. The process of action research is a problem solving process. In this particular study, the first identifiable problem was students' varying language usage skills. After identifying the problem, the next step was to research possible strategies or solutions. None of the researched data proved sufficient enough on its own; therefore, it was necessary to investigate further options, hence the action research.

Data Collection

Observational and reflective notes were kept throughout the six-week process. The observational data was typed, analyzed and coded using qualitative research methods. The notes analyzed assessments conducted within the classroom. These assessments ranged from informal diagnostic assessments to summative assessments to the NWEA's MAP assessment data. The notes also analyzed how the assessment led the planning and implementation of the classroom instruction and whether or not the students were engaged in the content.

Participants

The research was conducted at an urban school in the Midwest. It was conducted in a seventh grade functional English classroom with the permission of the administrator. The study consisted of ten participants between the ages of 12 and 13 years old.

Of the participants being observed, six students had learning disabilities with Individualized Education Plans (IEP), one student was on a 504 plan and three did not receive any specialized services from the special education department.

Procedures

The study began with the analysis of the computer-based NWEA's MAP assessment data from the fall Language Usage assessment. The assessments were reviewed for each specific Rasch UnIT (RIT) level. The RIT scores showed the achievement level for language usage according to the Rasch Scale. The scale "is based on the same modern test theory that aligns student achievement levels with item difficulties on the same scale" (NWEA, 2004, p. 1-1). Like a ruler measures a child's height over time, the Rasch Scale measures a child's academic growth over time.

After analyzing the Rasch scores and the rest of the NWEA's MAP assessment data, I developed an individualized instructional plan for each of my students in language usage areas. Based on their RIT scores and on previously observed student weaknesses and strengths, I identified the specific areas on which to focus in order to create a solid foundation of language usage skills. The language usages areas stressed in this study were 1) writing and research process, 2) spelling, punctuation, capitalization, 3) sentence structure and literary elements, 4) prewrite, draft, revise, edit, and 5) parts of speech. Each of the five areas was addressed throughout the course of the six weeks.

After analyzing the ten students' RIT range scores, the students were placed into three groups. Of the ten students, three students fell within one RIT range (180-200) and were labeled Group 1. The students in Group 1 had a variety of percentile ranges. The student percentile range is defined as the percentage of students across the country whose RIT score is less than or equal to the student being tested. Group 1's fall percentile scores were one, four, and seven. So of all the students tested around the country, my students in Group 1 test above or equal to only one, four and seven percent of the students.

In Group 2, there were five students who fell within the next higher range (201-210). Their percentile ranges varied with their nine point RIT range. The percentile ranges for Group 2 were 15, 15, 17, 23, and 31. Even though the percentiles are considerably higher than one percent, they are still significantly low according to the national averages.

The last two students fell within the next RIT range (211-230). Their percentile scores were both 43 percent, near average for seventh grade students. Since several of the ten students were similar in their RIT range scores and their previously teacher-observed abilities in the classroom, I chunked them into groups and planned three separate instructional lessons according to their individual needs or their RIT ranges. NWEA's DesCartes material provided the information I needed to plan the instruction. The information provided me "with a tool to translate student scores so that [I knew] when to move a student, or students, beyond the conventional curriculum at a particular grade level and when to develop skills that may have been presented earlier" (NWEA, 2004, p. 1-1). As I used the DesCartes, it helped "guide instruction based on the reports [and]….enhanced [my] ability to provide targeted instruction for individual students or groups of students" (NWEA, 2004, p. 1-2).

Figure 1 illustrates a basic outline of lesson plans for one week. The figure breaks down the different lessons and instruction being used by all three groups. The target skill addressed in the was adverbs. As shown on the figure, within one class period, the students were receiving differentiated instruction based on their abilities and the skills and concepts the NWEA data suggested I develop. Two of the five days, whole group instruction was a focus; and for the remainder of the week, the students worked on different tasks and assignments. The tasks varied from simple sentence building with blocks, to role-playing, to basic worksheet packets. Each student was taught with strategies that enabled him or her to experience success with adverbs.

Not only was this particular week with adverbs differentiated by readiness or ability, it was also differentiated by interest, as the students were allowed to write about anything of interest to them. It was further differentiated by intelligences. In three days of instruction, the lessons were tailored to the bodily-kinesthetic learners who were able to use their hands to build sentences and act out scenes. It reached the spatial learners who were able to map out their writing and it reached their linguistic intelligence, as they were able to creatively express themselves through their writing.

Similar interventions such as the differentiated lessons for adverbs took place for six weeks. After the sixth week of interventions, the students again took the MAP assessment.

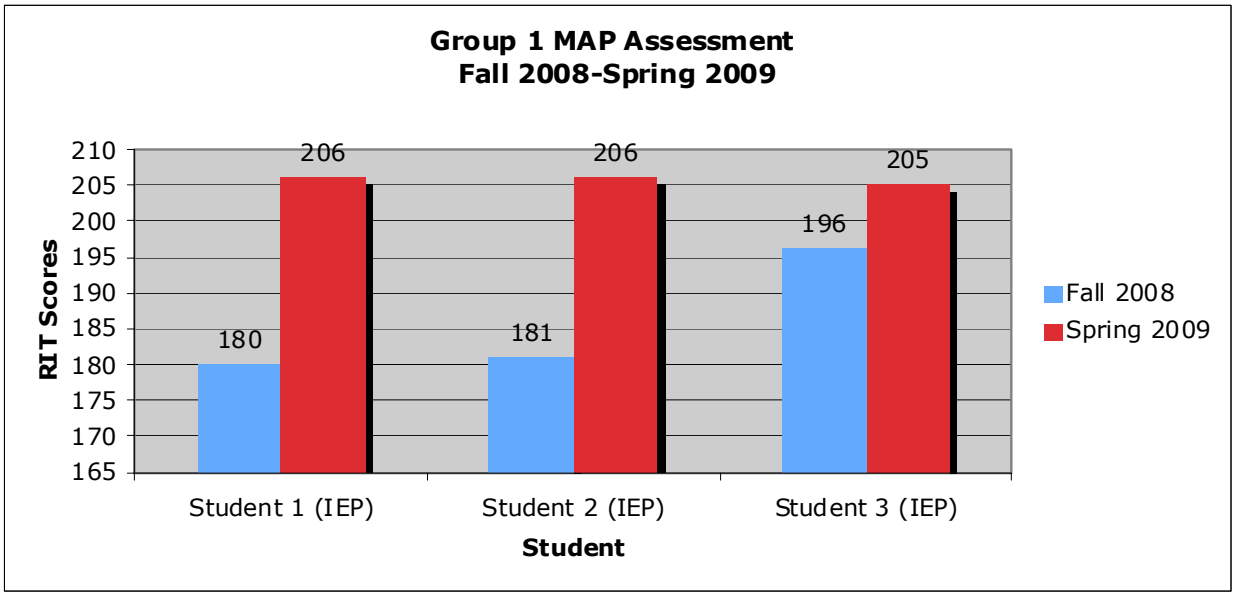
Figure 1

	Monday	Tuesday	Wednesday	Thursday	Friday
Group 1	School House Rock: Adverbs Note taking	8 parts of speech review Paper plate adverb review. Reflection	Work with parts of speech blocks. Create 10 sentences using at least one adverb and one adjective in each sentence.	Create parts of speech blocks: 10 adjectives 10 adverbs 10 verbs 10 nouns (Make sure to color coordinate the parts of speech)	Create 2 paragraphs about one of the sentences. Must: highlight adjectives in pink, adverbs in blue, subjects in yellow, verb phrases in green. It must also contain at least 1 simile.
Group 2	School House Rock: Adverbs Note taking	8 parts of speech review Paper plate adverb review. Reflection	Role Playing: Action Adverbs Students read and highlight adverbs in passage. Student complete dramatic acting of passage (stressing adverbs)	Create a 3-paragraph passage with action and adverbs. Must contain 10 adverbs. Highlight adjectives in pink, adverbs in blue, subjects in yellow, verb phrases in green.	Worksheet Packet on Adverbs: due at the end of the hour
Group 3	School House Rock: Adverbs Note taking	8 parts of speech review Paper plate adverb review. Reflection	Worksheet packet on adverbial phrases.	Correct Worksheet Packet Brainstorm using an event map. Share ideas and receive feedback.	Create a 5-paragraph story using the event map. Must include 10 adjectives, 10 adverbs, and 2 similes and circle and color adjectives in red, adverbs in blue; underline subjects in orange and verb phrases in purple.

Results

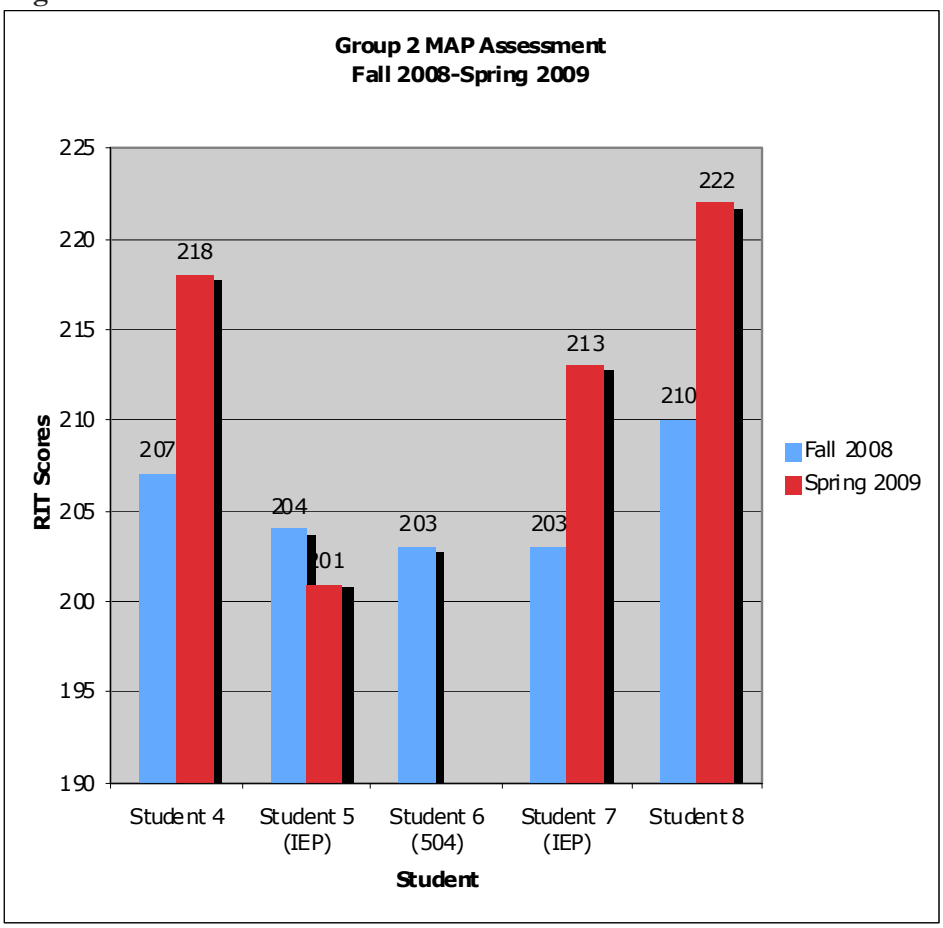
After the students retook the NWEA's MAP assessment in the spring, I analyzed the results of their assessment and compared them to the fall assessment RIT scores. Fall and spring scores were analyzed to see if the intervention of individualized instruction affected their spring assessment data in a positive way. The results of the study varied according to each student. Of the ten students, seven students increased their scores from the fall assessment to the spring assessment, with some increasing their scores by significant amounts. Two students received lower scores when retested in the spring and one student was absent the day of the spring assessment.

Figure 1



The following charts are broken down into the three separate student groups. Each chart breaks down the students' fall assessment score and their spring assessment score according to the RIT level.

Figure 2

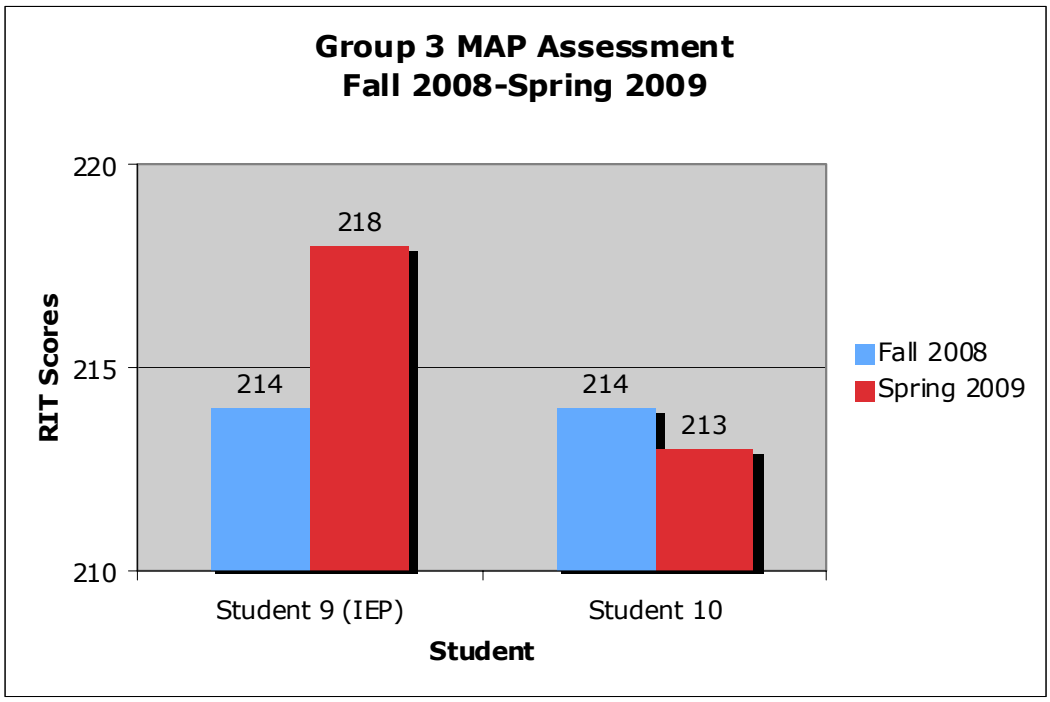


Members of group one improved significantly more so than the other groups. All members of this group improved their RIT scores by three ranges. Their percentile ranges also increased from one to 16 (up 15 percent), four to 16 (up 12 percent), and seven to 14 (up 7 percent). Although all members are still below grade level, notable improvements were made to their RIT scores and subsequently to their percentile ranges.

One student was absent for the test and her results were not available as of the press date. One student's RIT range stayed the same: two students improved by one RIT range, while the other student improved by two RIT ranges. Notably, most student percentile ranges increases as well, from 15 to 32 (up 17 percent), 17 to 9 (down 8 percent), 23 to 47 (up 24 percent), and 31 to 61 percent (up 30 percent).

Members of group three showed the least amount of progress of the groups. Both students remained in the same RIT range. Also, both students were in the 43rd percentile for their fall assessment and when

Figure 3



tested again in the spring, one student went up 3 percentage points, to the 47th percentile, while the other dropped 11 percent, to the 32nd percentile.

Conclusions

I have concluded that the intervention of individualized instruction based on use of MAP assessment data proved to be a worthwhile investment for the students. While adapting and tailoring the instruction to individual's needs and the state standards, student scores increase significantly, many times doubling or tripling the typical student growth numbers provided by the NWEA. Excluding all quantitative data and looking solely at my qualitative notes, while using individualized one-on-one instruction, student comprehension and engagement improved.

Now some will argue that students are bound to increase their scores from the fall assessment to the spring assessment without any interventions. However, I would disagree. The growth I saw in this particular study when interventions were used was anything but typical. For example, the typical growth, according to the NWEA's data, for Student 1 was 11 points. However when tested after the interventions, Student 1 surpassed the 11 points of typical growth and increased his/her score to a total of 25 points. That's over double the normal student growth. Student 4's typical growth should have been 4 points according to NWEA's data, but Student 4 scored well over the typical growth and nearly tripled the growth with a total of 11 points. Student 7 doubled the typical growth scoring 10 points and Student 8 tripled his typical growth by scoring 12 points from the fall to the spring assessment. I am confident these results would not have occurred without the interventions.

Not only do the results of the spring assessment show positive gains on the standardized assessment, the cognitive understanding and the strides I witnessed in the classroom are proof enough to me that using the data from the NWEA was beneficial to my students.

Both the students and the teachers adapted to the change of environment and quickly established a set of expectations that successfully supported the students and their learning. Not once did the students complain that one group was working on a game while they were completing a worksheet. The students bought into differentiation, I feel, because it was presented in a way that showed them that everyone was getting what they needed to become successful learners. And eventually, the teachers acted merely as facilitators while the students demonstrated peer teaching and learning techniques.

Further Research and Limitations

There are several recommendations that would further enhance the study. The first recommendation would be to continue the study for longer than six weeks. With more time to implement and analyze interventions, the results of the study and more importantly, student learning, could have varied dramatically. Another advantage to more time would be more opportunities to establish behavioral and academic expectations.

Formal student surveys, interviews, and goal setting workshops would have been beneficial tools for reflection. Those tools would have enabled me to adapt my instruction according to the students' suggestions.

Discussion

Several themes emerged from the six weeks of qualitative data kept. The following two themes appeared most frequently throughout the data:

Teaching Methods: Several forms of instruction were used (i.e. whole group discussion, role playing, manipulation, conferencing, peer-teaching, cooperative learning). The peer oriented methods, manipulation, and role-playing appeared to be the most engaging to the students.

Engagement/satisfaction: Several students needed many reminders to stay on task. Because there were several different lesson plans occurring at once, the students experienced a great deal of wait time, which at times led to behavioral issues. This improved over time as the students and teachers became more familiar with class expectations and the atmosphere of classroom differentiation. It also improved as the students increased their accountability for their actions. Teacher and student satisfaction came later on when students took accountability, worked collaboratively, and constructed/facilitated their own learning.

While conducting this study, I discovered, even through the difficulties that arose, that differentiation is in fact possible. It was a manageable process that required time and energy on my part. I sought out strategies from the Internet, fellow language arts teachers, and professors at the university. It required innovative thinking and planning. I believe if we expect our students to think creatively, we must in turn teach them in a way that enables them to do so.

Don't get me wrong, for someone who has never differentiated his or her instruction, the process will seem overwhelming at first. But it is so very important just to start: start somewhere. Start by giving the students choice in what they want to research. Start by giving them options on how they would like to present their material: be it standard written paper, through a video, in the form of a song, etc. By simply starting, you will be surprised at how quickly the process snowballs and how easily your students will come up with ideas that facilitate differentiated instruction and differentiated learning.

For me, by no means was it possible to differentiate every day in every class, especially as a first year teacher with little to no resources. Eventually, as resources start to accumulate and my knowledge of middle schoolers increases, I believe that differentiation will no longer be a strategy for creating lessons, but instead it will become a philosophy—a philosophy for educating students according to what they need. It will in essence be the "way of life" in my classroom. By following the philosophy of differentiation, I can guarantee that each and every student will receive the instruction he/she needs to be successful in my classroom. I can push them to learn in new ways and enhance the ways in which they learn best. By differentiating, I can create students who construct meaning from the world around them.

From this research, I have concluded that for this particular study the use of MAP assessment data to differentiate instruction did in fact affect students' language usage skills. Now I put the challenge out to you. Can you use the latest standardized tests in a meaningful way to help differentiate your instruction to meet the needs of your students?

References

Bergland, H. (2008) Personal Phone/Email Interview. Conducted December 15, 2008.

Chappuis, S. & Stiggins, R. (2008). Finding balance: Assessment in the middle school classroom. *Middle Ground*, 12(2). 12-15.

Cronin, J., Dahlin, M., Adkins, D., and Kingsbury, G. (2005). *The proficiency illusion*: North Dakota. 163-169. Retrieved November 21, 2008 from www.edexcellence.net/detail/news.cfm

Hurren, B. and Rutledge, M. (2008). Conversation piece: Varying assessment with team testing. *Middle Ground*, 12(2). 16-20.

Jackson, A. and Davis, G. (2000). *Turning points 2000: Educating adolescents in the 21st century*. New York: Teachers College Press.

McTighe, J. & O'Connor, K. (2005). Seven practices for effective learning. *Educational Leadership*, 31(4). 11-17.

Northwest Evaluation Association (NWEA) (2004). *Essentials*. Retrieved June 21, 2009 from https://reposrts.nwea.org/help/Essentials.pdf

Northwest Evaluation Association (NWEA) (2008). *About us: Assessment system*. Retrieved December 14, 2008 from http://www.nwea.org/about/index.asp

Robertson, T.S., & Valentine, J.W. (2000). *Research summary: What are appropriate assessment practices for middle school students?* 1-6. Retrieved November 21, 2008 from http://www.nmsa.org/Reseach/ResearchSummaries/Summary16/tabid/271/Default.aspx

Author Contact Information

Danette Asperheim
PO Box 704
Larimore, ND 58251
danette.asperheim@und.edu