EXPLORE Performance in Mathematics and Science: Why are Middle School Students Unprepared for Success in Mathematics and Science?

Janet L. Tassell[a], Kerrie McDaniel[a], Heather Johnson[b], Antony Norman[a], Roger Pankratz[a], Rico Tyler[a]

Corresponding author: Janet Tassell (janet.tassell@wku.edu)
[a]Western Kentucky University, Bowling Green, Kentucky, 42101, USA
[b]Northwest-Shoals Community College, Muscle Shoals, Alabama, 35661, USA

Keywords: mathematics, science, education, middle school, preparation

International Journal of Innovation in Science and Mathematics Education, *20*(1), 1-23, 2012.

Abstract

The research question for this project was "What factors can be identified that contribute to the disparity between EXPLORE scores and state assessment data?" Three themes were examined to study the research question: (1) leadership and planning, (2) classroom instruction, and (3) the assessment of higher-order thinking skills and problem solving. The 2008-2010 Western Kentucky MSP grant, funded by the USA National Science Foundation (NSF), was an the organized effort to understand the circumstances that impact students' performance on the EXPLORE test and, therefore, college and career preparedness. It was an evidence-based problem-solving process to collect and analyse data directly related to declining math/science scores in participating middle school. The data reveal a challenge: Closing an existing gap between teacher beliefs and what researchers observed in classrooms regarding classroom planning, instruction and assessment.

Introduction

Unfortunate, but unsurprising, information in a variety of reports that countries are severely threatened if with failure to produce more scientists, technology experts, engineers, and mathematicians. In the U.S. some of these concerned entities include the National Center for Education and the Economy (2007), the Council of Chief State School Officers Mathematics and Science Task Force (2006), and the National Academy of Sciences, National Academy of Engineering, and the Institute of Medicine of the National Academies (2005). For example, the National Science Teachers Association (NSTA) Web News Digest (2008) recently reported that the U.S. ranks 32nd out of 90 countries in undergraduate engineering production even though 15 of 20 of the fastest growing occupations require significant science or mathematics preparation. The key to producing more scientists and mathematicians is improving mathematics and science preparation, thus bolstering the number of STEM students in the K-12 pipeline. Evaluation of mathematics and science preparedness is necessary to establish current levels of science and mathematics education and set benchmarks for future students. Key stages of the K-12 pipeline must be identified and monitored to insure the success of students in STEM disciplines.

A recent study by ACT (2008) found 8th grade academic achievement to be the best predictor of college and career readiness for high school graduates, far more important than students' background characteristics, course work taken in high school, or high school grade point averages (GPAs). This finding is supported by the Finkelstein and Fong (2008) study wherein unprepared students lose ground over their four years of high school. In addition, both TIMSS (Mullins, Martin, & Foy, 2005) data and National Assessment of Educational Progress (NAEP) reports show that U.S. students perform reasonably well at the 4th grade, but performances drop measurably by grade 8 (Silver, 1998). Given the significance of middle school performance, understanding the potential interventions to improve middle school mathematics and science education become paramount.

To highlight an example of a specific representative of a typical testing scenario, this study looks at Kentucky. Kentucky has had two tests to measure middle school performance in eighth grade – the Kentucky Core Content Test (KCCT) and ACT EXPLORE. Students did well on the former, but not on the latter (Table 1). The KCCT was the state assessment system and the scores were intended to be a direct indicator of success of the 8th grade transition to high school. The assessment items were based on an extensive set of core content standards for all students including standards for "thinking and problem-solving" and "making connections." However, this assessment was not designed to predict college readiness. In 2006, Kentucky's Council on Postsecondary Education commissioned a 110-member STEM (science, technology, engineering, and mathematics/ED (Education) task force that prompted the Kentucky Department of Education in 2007 to adopt the Educational Planning and Assessment System (EPAS) statewide in response to the need for all students to be prepared for high school and the transitions they make after high school. As a component of EPAS, the EXPLORE assessment was implemented in 2007 and is administered to all 8th graders in September of the school year. The results provide data on the extent to which middle school students are being prepared for rigorous courses in high school and are on track in their development for success in grades 9-12 and ultimately postsecondary preparation. According to ACT regarding the 8th grade assessment, "there is a critical defining point for students in the college and career readiness process—one so important that, if students are not on target for college and career readiness by the time they reach this point, the impact may be nearly irreversible" (2008).

Table 1: Comparison of EXPLORE and KCCT Results by Percent of Students Making the Desired Cut Score

	Science		Math	
	EXPLORE	KCCT	EXPLORE	KCCT
2006-07	10.84	45.96	25.75	52.06
2007-08	9.58	59.57	26.85	56.96
2008-09	10.48	61.72	28.96	60.93
2009-10	14.10	57.02	35.59	62.12
	Benchmark	*Proficient*	*Benchmark*	*Proficient*

Interested in whether these trends hold true in southern Kentucky, the authors established a partnership of school districts and other stake-holders to evaluate standard test scores to see how they performed on the state assessment vs. EXPLORE. In 2007, the average EXPLORE scores for mathematics and science for each of the 18 partner schools discussed in this project did not approach the "benchmark score" which would indicate students are likely "on track" for rigorous high school courses. According to ACT, Benchmark scores are 17 for mathematics and 20 for science, which means the performance of students in partner schools would have to improve on average by more than 20% in science and more than 15% in mathematics to have access to and be successful in college preparatory high school courses. What is even more surprising is that while the state assessment scores were showing most partner schools' performance as relatively strong, EXPLORE scores remained low.

When we examined the meaning of the partner schools' EXPLORE scores, it became obvious that without intervention most 8th grade students in partner schools would continue to fall far below benchmark levels for the ACT that predicts success at the college level (ACT, 2008). In reference to Table 2, EXPLORE data from partner schools strongly suggested that most 8th grade students had mastered what is in column 1 (score range 13-15) but were lacking the understanding to be successful in column 2 (score range 16-19) and column 3 (score range 20-23). All students must address the standards in column 2 and many in column 3 to be on track for college preparation courses in high school. The Partnership's analysis of EXPLORE data showed that partner school students were unable to respond to items that required higher-level thinking (Cognitive complexity, shown in column 3).

Table 2: ACT EPAS Standards in Mathematics and Science for Different Score Ranges (ACT, 2011)

Mathematics EXPLORE		
Score Range 13-15	Score Range 16-19	Score Range 20-23
▪ Calculate the average of a list of positive whole numbers	▪ Calculate the average given number of data values and the sum of data values	▪ Translate from one representation of data to another (e.g., a bar/graph to a circle/graph)
▪ Perform a single computation using information from a table or chart	▪ Perform computations on data from tables and graphs	▪ Calculate the missing data value given the average and all data values
Science EXPLORE		
▪ Select a single piece of data (numerical or non-numerical) from a simple data presentation	▪ Select two or more pieces of data from a simple presentation	▪ Select data from a complex data presentation (e.g., a table or graph with more than three variables)
▪ Identify basic features of a table or graph (e.g., headings, units of measurement, axis labels)	▪ Find information in a brief body of text	▪ Compare or combine data from a simple data presentation
	▪ Determine how the value of one variable changes as the value of another variable changes in a simple data presentation	▪ Translate information into a table, graph or diagram

Higher levels of cognitive complexity are beginning to appear in state and national standards. Treffinger (2008) suggests that providing students with problem solving tools prepares them for the future. Requiring students to solve problems and think deeply gives students the cognitive tools necessary to exist in changing circumstances (Wiggins, 1998). Although teachers desire to promote critical thinking in the classroom and provide assessments that are likewise at a higher level, the truth is that most formative and summative assessment and classroom activities are at lower cognitive demand levels.

The Partnership

In response to the disconnect between EXPLORE and state assessment scores, and perceived needs in mathematics and science education at the middle school level, Western Kentucky University, a state university, its three regional campus centers, four community colleges, and 30 schools in Western Kentucky developed and established an expanded partnership, entitled the Western Kentucky Math Science Partnership (MSP) addressing a documented local, state, and national challenge: ***improving teaching and learning in middle school to insure access of all students to rigorous high school mathematics and science***. The 2008-2010 Western Kentucky MSP grant, funded by the National Science Foundation (NSF), was an effort to understand the circumstances that impact students' performance on the EXPLORE test and, therefore, college and career preparedness. It was an evidence-based problem-solving process to collect and analyse data directly related to declining math/science scores in participating middle school. The research question for the project was "What factors can be identified that contribute to the disparity between EXPLORE scores and the state assessment data?" Three themes were examined to study the research question: (1) leadership and planning, (2) classroom instruction, and (3) the assessment of higher-order thinking skills and problem solving.

Method

Western Kentucky MSP Goals

The project goal was to use sound measurement procedures to interpret the data in a manner that would identify the most critical school curriculum instruction and context factors that cause or fail to increase learning for all students. These context factors would be contributors to the disconnect between EXPLORE scores and state assessment (see Table 1).

Participants

The Western Kentucky MSP project involved two key action role groups and four support groups in a collaborative 2-year process of collecting and analysing middle school data related to the decline of mathematics and science scores from 4th to 8th grade; interpreting the data; identifying research based strategies to address the problem; and planning school, community, curriculum, professional development programs to begin implementation in the fall of 2010.

The two key action groups in this process were the 272 mathematics and science teachers along with middle schools of the partnership, and 25 STEM/ED (science, technology, engineering, math, and education) faculty from the six partner postsecondary institutions.

The 25 STEM/ED faculty were trained as participant facilitators who went to the 30 middle schools and interacted with mathematics and science teachers, and school leaders on data collection (Fall, 2008) data interpretation (Spring, 09), strategy planning (Fall, 2009) and pilot testing implementation design spring 2010.

This process was supported by a four-person Management Team, a three-person university Data Team, ten member STEM/ED/school practitioner Work Groups, and a 10-member Advisory

Board representing key partnership role groups and as math/science experts. An external evaluator evaluated the effectiveness of the partnership.

Data Sources and Instruments

To understand this disconnect, data were collected across 30 area middle schools. The instruments and processes used for data collection provided information in many different areas but focused on 3 themes: (1) leadership and planning, (2) classroom instruction, and (3) the assessment of higher-order thinking skills and problem solving. Initially the data were collected from five sources with the purpose of triangulating the data:

- TIMSS surveys completed by students one time during the school year (IEA, 2003)
- TIMSS surveys completed by teachers one time during the school year (IEA, 2003)
- Interviews with teachers one time during the school year (Harmon, Henderson, & Royster, 2003)
- Interviews with principals one time during the school year (Harmon et al., 2003)
- Classroom observations using the Mathematics Classroom Assessment Instrument and the Science Classroom Assessment Instrument for one class period per teacher in the study (Henderson, 2007)

The interview and classroom observation protocols were used from a mathematics and science school audit program (Leadership by Design) published by the Kentucky Center for Science and Technology (Harmon, Henderson, & Royster, 2003). CLASS was developed at the University of Virginia Center for Advanced Teaching and Learning and has both a conceptual and research base and provided information about the physical setting of the classroom, books and supplies available, and what actually occurred in the classroom. Data from all five sources were organised, aggregated, and analysed by researchers at WKU. All of the survey, interview, and observation data were summarised in mean form. Data collectors joined the researchers to interpret the results and draw conclusions.

Inter-Observer Reliability

Training was provided by the designers of the instruments (Henderson, 2007a; Henderson, 2007b) for the use of the classroom observation instruments for the Mathematics Classroom Assessment Instrument and the Science Classroom Assessment Instrument. The protocol was very specific as observers were to look for the most observed aspect of the classroom rather than marking all of the observed items – forcing the observer to make a decision about the top observation. The next step was to then have discussions with colleagues decisions to code certain aspects where made. Great efforts were made in tightening the reliability prior to releasing the observers to the field for the classroom observations.

Data Collection Procedures

The data collection fell mainly to the STEM/ED Management Team and a few STEM community college faculty members. One full school day was spent in each school by the data collectors interacting with school leaders, and mathematics and science teachers. The purposes for the visits were to build trust and collect data. The STEM/ED data collectors met with the principal, school staff, and visited with teachers in the teacher's lounge or in their rooms during their planning period. After school, data collectors met with all mathematics and science teachers and the school principal or building MSP Start Project curriculum coordinator to discuss the

TIMSS questionnaire and a timeline for completion. School principals or curriculum coordinators completed the school questionnaire. All science teachers and all mathematics teachers collectively completed the subject-specific questionnaire, and each mathematics and science teacher completed an individual questionnaire.

TIMSS questionnaires were available online and in print to enable all students and school faculty to complete them. This questionnaire provided information about the teachers' and students' perceptions about mathematics and science learned and teaching at the school. STEM/ED faculty observed mathematics and science classes to record the quality of interaction between teachers and students. All data collected by the STEM/ED faculty were recorded in a structured format that facilitates analysis by the Data Team. Additionally, student achievement data from Kentucky's CATS and EPAS were collected.

The combination of these instruments and data sources provided a comprehensive picture of each school environment as it relates to mathematics and science. The TIMSS instruments allowed for large-scale, quantitative analysis of issues related to math/science education. For example, the student questionnaire tapped into attitudes and beliefs about math/science, their perception of the instruction they received, and their personal study habits, whereas the teacher questionnaire taps into their attitudes and beliefs, instructional practices, professional preparation, and professional development. Henderson's Classroom Observation Instrument allowed for more in-depth investigation of what actually happens in mathematics and science classrooms. The state assessment and EXPLORE data provided the "ultimate" measure of each school's preparation of its students in mathematics and science.

Data Analysis and Confirmation

Two STEM/ED faculties met with each of the 13 four-school clusters of teachers to present the data analysis and facilitate teacher's interaction with each other and data results. An attempt was made to look for relationships and explanations of the results not obvious to the researchers and reach a consensus on the teacher's collective interpretation of the data and its meaning for actions to improve student learning. STEM/ED faculty facilitated the development of a group interpretation report by meeting with the teacher participants and exchanged their interpretations via the final report. Each group had the opportunity to question and/or challenge their colleagues. The final consensus report became an input to the final report to guide strategy development.

Each school was also provided three items: 1) a more detailed report of all the data collected, 2) the results of the school's student responses to the TIMMS survey compared to all 30 participating schools, and 3) a survey to complete regarding their opinions of the data collected.

Standard quantitative and qualitative data analysis procedures were used to address all evaluation questions (have we listed these or are these the actual questions they were asked). Emerging themes were identified and data were reviewed for classification into categories. In the case of document and data reviews, careful comparisons were conducted by trained staff to uncover substantive changes and to infer the impetus behind such changes. All deliverables, and the related data collection and analysis procedures, adhered to *The Program Evaluation Standards* developed by the Joint Committee on Standards for Educational Evaluation (1994).

In late spring/early fall of 2009, teachers in the 30 partner middle schools were provided the summary of findings and conclusions of the project researchers relative to the three focus themes: (1) leadership and planning, (2) classroom instruction, and (3) the assessment of higher-order thinking skills and problem solving. Teachers were then given a survey that asked about their perceptions of the accuracy of each of the researcher's summary conclusion statements. After teachers' response to researchers' findings, teachers indicated how they thought long-range planning; better instruction for higher-level thinking, and better classroom assessments would affect student outcomes.

Results

Four main factors have been identified that contribute to the disparity between EXPLORE and CATS scores and these revolve around the 3 focus themes (leadership and planning, classroom instruction, and the assessment) identified in the project. These contributing factors include:

1) Higher-order thinking in planning
2) Higher-order thinking in instruction
3) Higher-order thinking in assessments
4) Disconnect between teacher discernment of practices and actual teacher practices

Based generally on teacher responses on the TIMSS and in interviews, teachers believe a substantial part of instruction should focus on higher level thinking. They also believe that both mathematics and science applications to students' real-life experiences are very important to the learning process.

The needs assessment revealed four emerging categories that fit with the focus themes as well: (1) long-range planning, (2) instruction in thinking and problem solving, and (3) assessments requiring higher order learning required for high performance on EXPLORE.

The results will combine the themes and factors with the four emerging categories:

1) Higher-Order Thinking in Long-Range Planning (Improvement Plans/Goals for Science and Math)
2) Higher-Order Thinking and Problem Solving Supporting Instruction
3) Higher-Order Thinking In Assessments Required for High Performance on EXPLORE (Teacher vs. Student TIMSS Results vs. Classroom Observations)
4) Disconnect Between Teacher Perspicacity of Practices and Actual Teacher Practices

Emerging Category 1: Higher-Order Thinking in Long-Range Planning

Assessment using the CLASS observation instruments (Harmon, et al, 2003) and TIMSS teacher and student surveys indicated that less than 10 percent of the partner schools visited had long-range plans beyond the next testing period (Table 3). There were also no long-range professional development plans specifically for mathematics and science teachers to increase knowledge, skills and networking in an ever-changing and technology-demanding field of study although most teachers believe it would be beneficial. Few principals had been involved in professional development that addresses qualities and innovations for mathematics and science programs and many lacked an understanding of the financial and professional development needs in the sciences. Thus, with no long-range (two- to five-year) goals, there was no combined effort by teachers, administrators, and the community to support improvements in mathematics and science. As a result of this short-sighted planning, higher cognitive skills are not being considered in planning for future science curriculum, materials, or personnel. The lack of higher thinking skills in long-range planning filters down into the day-to-day science and mathematics classroom.

Table 3: Middle School Teacher Questionnaire on Content-Centered Professional Development and Ideals Regarding Teaching Mathematics and Science

Middle School Teacher Questionnaire - Overall Results

3a.Which of the following subjects do you teach?	N
Math	46
Science	32
Both	7

Mathematics Results

3b1. In the past two years, have you participated in professional development in any of the following?	Yes	No
Mathematics content	76%	24%
Mathematics pedagogy/instruction	58%	42%
Mathematics curriculum	83%	17%
Integrating information technology into math	69%	31%
Improving students' critical thinking or problem solving	61%	39%
Mathematics assessment	64%	36%

3b2. To what extent do you agree or disagree with each of the following statements?	Agree	Disagree
More than one representation should be used in teaching a math topic.	100%	0%
Math should be learned as set of algorithms or rules that cover all possibilities	64%	36%
Solving math problems often involves hypothesising, estimating, testing, and modifying findings	100%	0%
Learning math mainly involves memorising	28%	72%
Few new discoveries in math are being made	33%	67%
Modeling real-world problems is essential to teaching math	100%	0%

Science Results

3c1. In the past two years, have you participated in professional development in any of the following?	Yes	No
Science content	82%	18%
Science pedagogy/instruction	65%	35%
Science curriculum	83%	17%
Integrating information technology into science	53%	47%
Improving students' critical thinking or problem solving	76%	24%
Science assessment	63%	37%

3c2. To what extent do you agree or disagree with each of the following statements?	Agree	Disagree
More than one representation should be used in teaching a science topic.	100%	0%
Solving science problems often involves hypothesising, estimating, testing, and modifying findings	100%	0%
Learning science mainly involves memorising.	11%	89%
There are many ways to conduct scientific investigation.	100%	0%
Getting the correct answer is the most important outcome of a student's scientific experiment.	10%	90%
Scientific theories are subject to change.	100%	0%
Science is taught primarily to give students the skills and knowledge to explain natural phenomena.	67%	33%
Modeling natural phenomena is essential to teaching science.	92%	8%
Most scientific discoveries have no practical value.	3%	97%

Interestingly, the teachers generally seemed to agree with the findings of the researchers (Table 4). One hundred percent thought that it was a “somewhat” to “very accurate” statement that few schools had specific goals for mathematics and science programs, and 100% agreed (somewhat to very) that most instruction was teacher led and factual. All teachers felt that their administrators were not involved in the professional development that was content specific for mathematics and science. There was lesser agreement as “very accurate” (25%) that students were not involved in investigations at higher levels and that questioning techniques were at a higher level. However, 100% still agreed at some degree to this situation.

Table 4: Teachers’ Responses to Researchers’ Conclusions

Finding/Conclusion	Very Accurate	Accurate	Somewhat Accurate	Not Accurate
Few schools have specific goals for mathematics or science programs beyond the present testing period	43%	33%	24%	0%
Few schools have professional development plans specifically related to mathematics and/or science instruction	72%	29%	9%	0%
Most principals have not been involved in professional development programs that address mathematics and science programs	15%	28%	23%	0%
Only a few mathematics and science teachers always find one or more applications within the content covered	15%	44%	41%	0%
In just a few science classrooms, students were engaged in investigations at higher levels	24%	56%	20%	0%
Student-centered learning (student presentations, small group discussion, students solving problems) was rarely observed	43%	41%	9%	0%
Classroom instruction is mostly teacher-led and centers on mastery of factual information or reviewing content	47%	43%	10%	0%
Questioning techniques that led to divergent or critical thinking was observed in just a few of the classrooms	25%	52%	23%	0%

Emerging Category 2: Higher-Order Thinking and Problem Solving Supporting Instruction

The results from the TIMSS for Teachers indicate that nearly all mathematics and science teachers agree (1) that solving mathematics and science problems involves hypothesising, estimating, testing, and modifying findings and (2) that modeling real-world problems is essential to teach mathematics and science (Table 3 and Table 4). However, classroom observation data collected using the Classroom Observation/Assessment Instruments showed 75 percent of instruction time was used for lecture, teacher demonstration, and teacher-led discussion in the mathematics classroom and about 60% of the time in the science classroom is occupied with teacher-led activities (Figure 1 and Figure 2). Less than 10% of mathematics and science classrooms used investigations and problem-solving techniques.

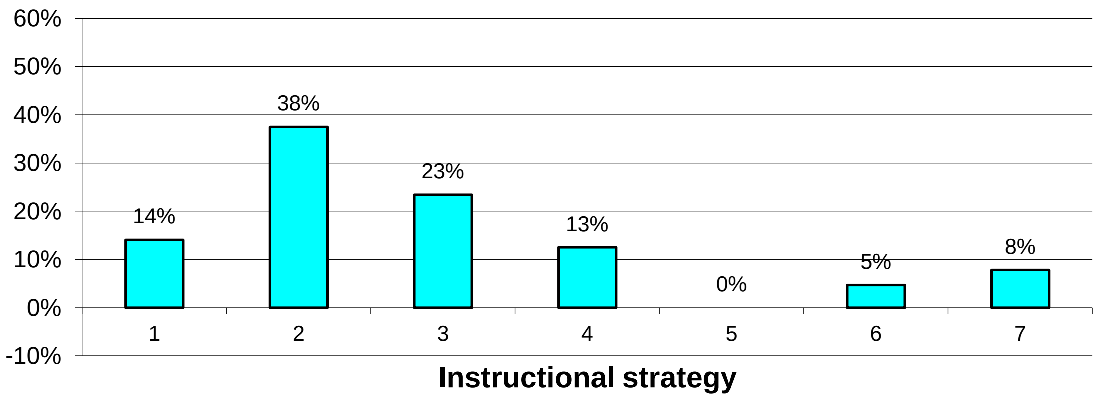


Figure 1: Primary/predominant instructional strategies of mathematics teachers observed in the classroom. Note that 1 = Lecture, 2 = Teacher Demonstration, 3 = Teacher Led Discussion, 4 = Individual Assistance, 5 = Student Presentation, 6 = Small Group Discussion, 7 = Student Solving Problems.

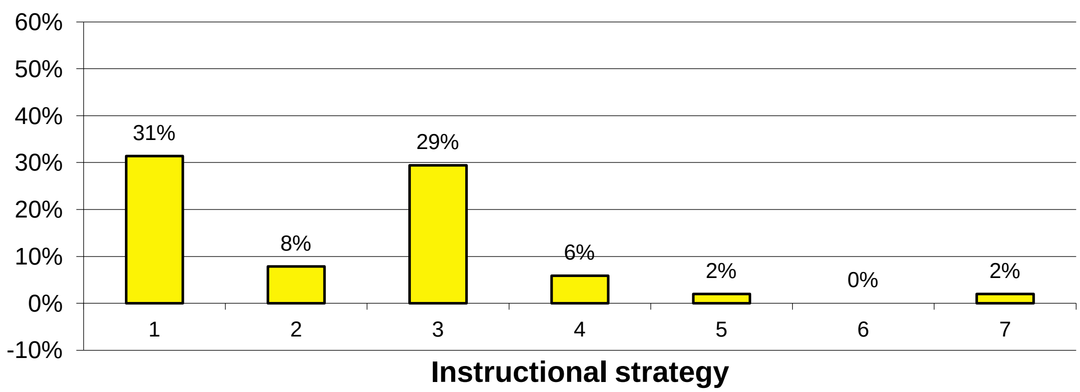


Figure 2: Primary/predominant instructional strategies of science teachers observed in the classroom. Note that 1 = Lecture, 2 = Teacher Demonstration, 3 = Teacher Led Discussion, 4 = Individual Assistance, 5 = Student Presentation, 6 = Small Group Discussion, 7 = Student Investigation.

Figures 3 and 4 show the predominate activities in which students were engaged in during the mathematics and science classroom. Listening to presentations and completing skills-based worksheets were the main activities observed in the mathematics classroom (Figure 3) while the science classroom participated in discussions and listened to lectures (Figure 4). Only 3% of the mathematics classrooms participated in investigations while only 10% of science classrooms were active in inquiry. The use of technology to analyse data and share results was rarely observed in either the mathematics or science classroom and higher order thinking skills were less than 10% in both instances. Activities where students were learning scientific skills such as investigating what involves analysis of data, designing experiments, formulating hypotheses, interpreting data, forming conclusions, or evaluating solutions were rarely observed. This is in stark contrast to what the teachers believe to be how science should be taught (Table 3 and Table 4).

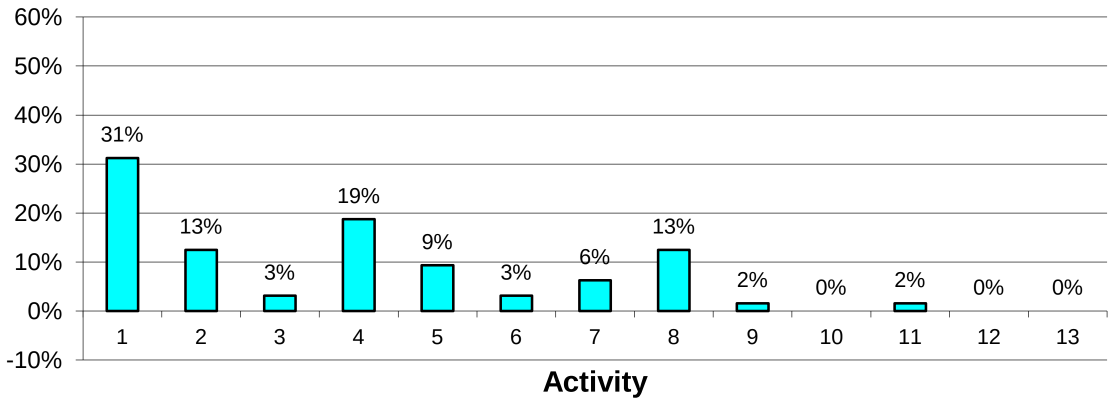


Figure 3: Predominant activity of mathematics students in the mathematics classroom. Note that 1 = Listening to Presentation, 2 = Participation in Discussion, 3 = Conducting Mathematics Investigation, 4 = Completing a Skills Worksheet, 5 = Working on assignment including higher problem-solving, 6 = Using hands-on materials to verify, 7 = Applying mathematics to realistic problems, 8 = Assignment/questions from text/other resources, 9 = Taking test, 10 = Sharing solutions or strategies, 11 = Using computer software, 12 = Using the Internet for research, 13 = Using computer for inputting/analysing data.

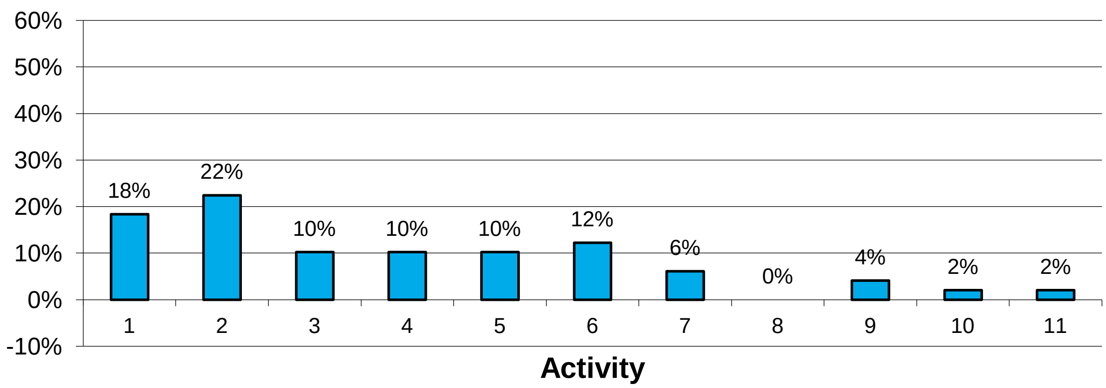


Figure 4: Predominant activity of science students observed by researchers. Note that 1= Listening to Presentation, 2 = Participation in Discussion, 3 = Conducting investigation, 4 = Conducting student or teacher initiated experiment, 5 = Print-based activity: Reading, answering questions, 6 = Working on written assignment, 7 = Taking a test, 8 = Using computer software, 9 = Using the internet for research, 10 = Using computer for inputting analysing data, 11 = Making a presentation or listening to student presentation.

In addition to recording the activities of the students, the researchers examined the skills that were being developed in the students by the teachers in the mathematics and science classroom. The skills that were honed the most in the mathematics classroom were recognising and observing (64%), reciting and recalling (70%), and computing and calculating (70%) (Figure 5). The skills that were not being developed in the mathematics classroom included interpreting data (12%), hands-on investigation (10%), determining problem-solving strategies (12%), and discussion where data is interpreted (4%). In the science classroom, the skills that are most pronounced are observation skills at a low percentage of 19% (Figure 6). All other skills that are traditionally accepted as standard science and scientific method such as predicting, inferring, using technology to analyse data, etc. fell below 15%. Few cognitively-complex science and mathematics tasks were observed in the classroom.

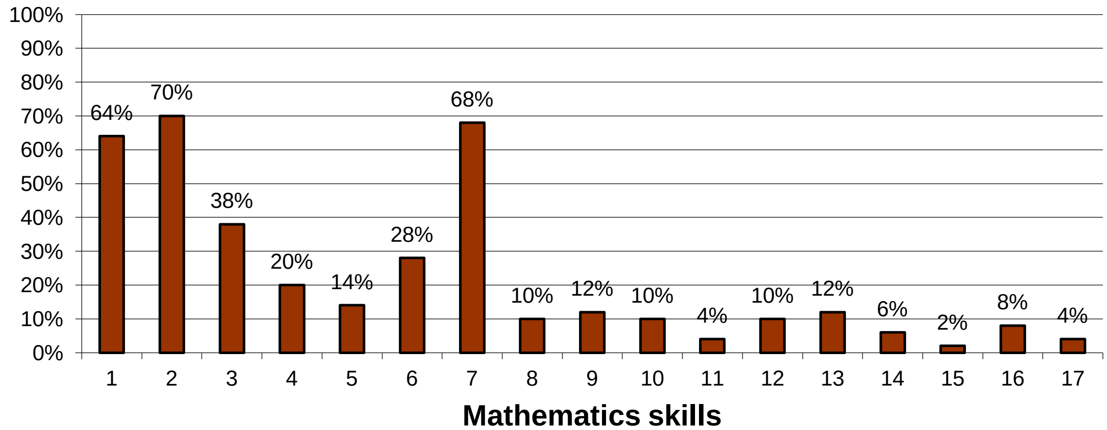


Figure 5: Mathematics skills being developed by mathematics teachers.
Note that 1 = Recognising/Observing, 2 = Reciting/recalling facts, 3 = Classifying, 4 = Measuring/estimating, 5 = Coordinate Graphing, 6 = Constructing charts/graphs, 7 = Computing/calculating, 8 = Collecting/recording data, 9 = Interpreting/analysing data/statistics, 10 = Investigating (Hands-on, Tech), 11 = Applying Theorems/principles, 12 = Evaluating the relevancy of data, 13 = Determining problem solving strategy, 14 = Creating/formulating pattern or equation, 15 = Evaluating logical consistency, 16 = Justifying/verifying solutions, 17 = Interpretive discussion.

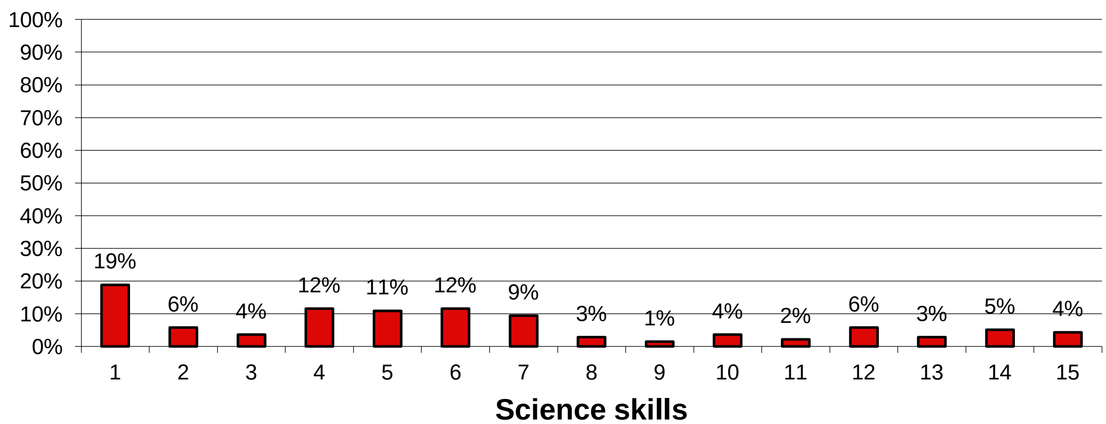


Figure 6: Observed science skills being developed by science teachers.
Note that 1 = Observing, 2 = Measuring, 3 = Classifying, 4 = Inferring, 5 = Predicting, 6 = Communicating, 7 = Investigating (Basic Level), 8 = Investigating (Involves Analysis of Data), 9 = Designing Experiments, 10 = Formulating Hypothesis, 11 = Conducting Experiment, 12 = Collecting/Interpreting Data, 13 = Forming Conclusions, 14 = Evaluating, 15 = Interpretive Discussion.

Table 4: Aggregate of Teacher Response to the TIMMS Survey Showing the Recognition of Mathematics and Science Teachers for Inquiry Style Learning with Higher Order Thinking Skills

Aggregate of Teacher Responses to the TIMMS Survey					
Math			Science		
To what extent do you agree or disagree with each of the following statements?	Agree	Disagree	To what extent do you agree or disagree with each of the following statements?	Agree	Disagree
More than one representation should be used in teaching a math topic.	100%	0%	More than one representation should be used in teaching a science topic.	100%	0%
Math should be learned as set of algorithms or rules that cover all possibilities	64%	36%	Solving science problems often involves hypothesising, estimating, testing, and modifying findings	100%	0%
Solving math problems often involves hypothesising, testing, estimating, and modifying findings	100%	0%	There are many ways to conduct scientific investigation.	100%	0%
Learning math mainly involves memorising	28%	72%	Learning science mainly involves memorising.	11%	89%
Few new discoveries in math are being made.	33%	67%	Scientific theories are subject to change.	100%	0%
Modeling real-world problems is essential to teaching math.	100%	0%	Getting the correct answer is the most important outcome of a student's scientific experiment.	10%	90%
			Science is taught primarily to give students the skills and knowledge to explain natural phenomena.	67%	33%
			Modeling natural phenomena is essential to teaching science.	92%	8%
			Most scientific discoveries have no practical value.	3%	97%

While the observed science skills in the classroom show higher incidence of rote-memorisation, recall, and computation than more advanced skills (Figure 5 and Figure 6), it is not a surprise that the numbers of students engaged in problem-solving activities are low as well (Figure 7). Both mathematics and science classrooms show less than 20% participating in higher-order problem-solving activities.

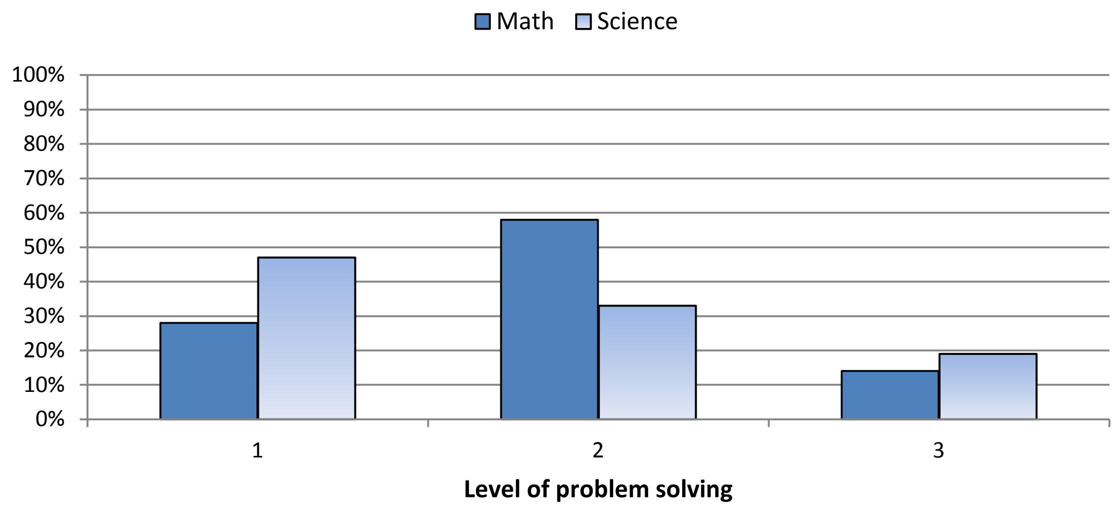


Figure 7: Levels of problem-solving in observed mathematics and science classrooms. Note that 1 = Students not involved in investigation or any type of problem-solving, 2 = Students are involved in investigation, lower level

In addition to observing teaching style in the classroom, student activities and skills that were being developed in the classroom, the researchers looked at questioning strategies in the classroom to see if these displayed higher order levels. Although teachers did engage in questioning techniques, few showed evidence of higher order questioning in both mathematics and science classrooms (Figure 8). About 90% of teachers' classroom questions were classified as convergent, factual, or conceptual recall. Less than 10% of teachers' questions were classified as divergent. Only fifteen percent of questions by mathematics and science teachers were classified as stimulating high-level thinking in students.

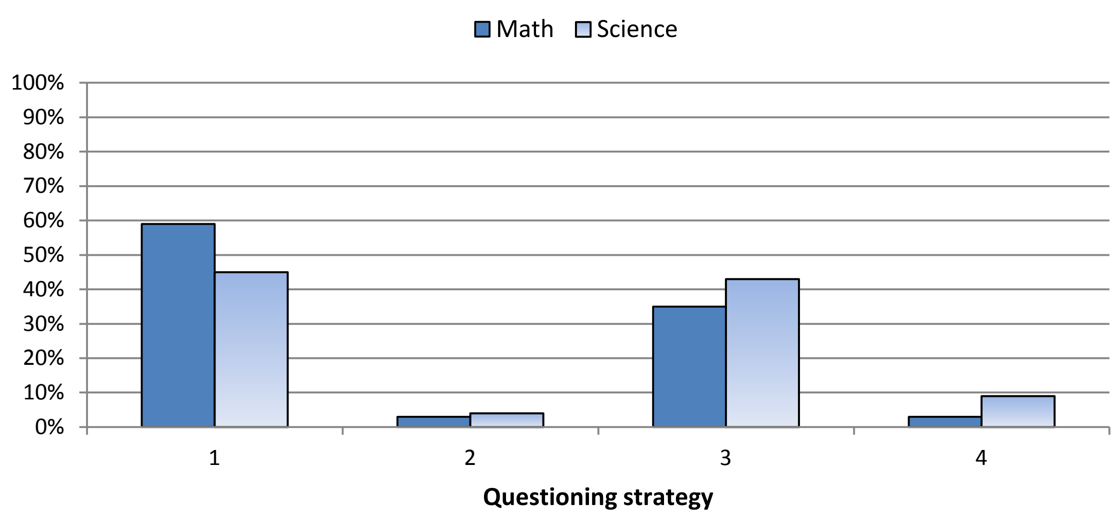


Figure 8: Questioning strategies in Observed Mathematics and Science classrooms. Note that 1 = Questions convergent, 2 = Questions divergent, 3 = Balance between recall and cognitive questioning, 4 = No questions asked.

The results of the project highlight a lack of cognitively-complex activities, instructional strategies, and planning in the classroom that lead to higher-level thinking. This void could be a notable contributor to the low scores observed on the EXPLORE test.

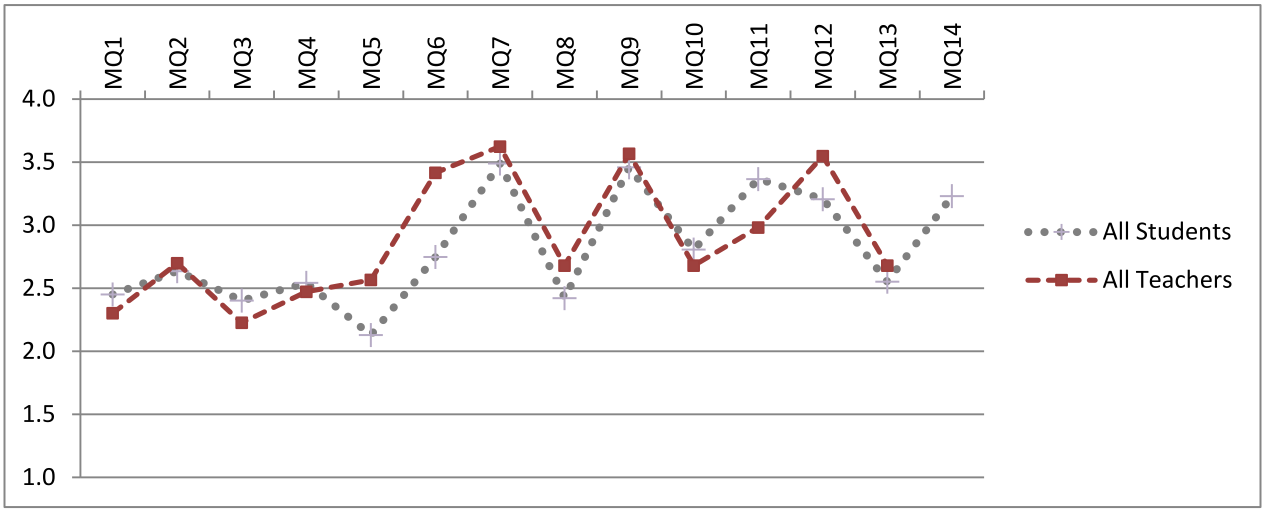
When teachers were asked to rate a list of ideas for improving instruction related to higher-level teaching and learning, the top five were: (1) professional development on questioning strategies, (2) observing exemplary models of inquiry teaching and learning, (3) professional development on an in-depth understanding of mathematics and science content, (4) availability of a mathematics or science specialist, and (5) professional development on effective strategies for teaching and learning.

Emerging Category 3: Higher Order Thinking In Assessments Required for High Performance on EXPLORE

On the TIMSS questionnaire, teachers report their tests “almost always” give assessment questions that involve application of mathematics procedures and scientific principles and “sometimes” include test questions that involve searching for patterns and relationships. However, from classroom tests collected and classroom observations, researchers found very few open-response questions or items that required higher-level thinking (see Figure 3 and Figure 4). In talking with teachers about their classroom assessments in the after-school meetings, it became evident that their understanding of assessments that required higher-level learning at higher cognitive levels was different from that of the observers and different from the standards on the EXPLORE test. For example, many teachers used verbal and written assessment questions at the Remember or Understand level of Bloom’s Cognitive Taxonomy (Bloom, 1954), yet believed that they were reaching the Apply and Analyse level tapping into higher-level cognitive complexity.

Emerging Category 4: Disconnect Between Teacher Perspicacity of Practices and Actual Teacher Practices

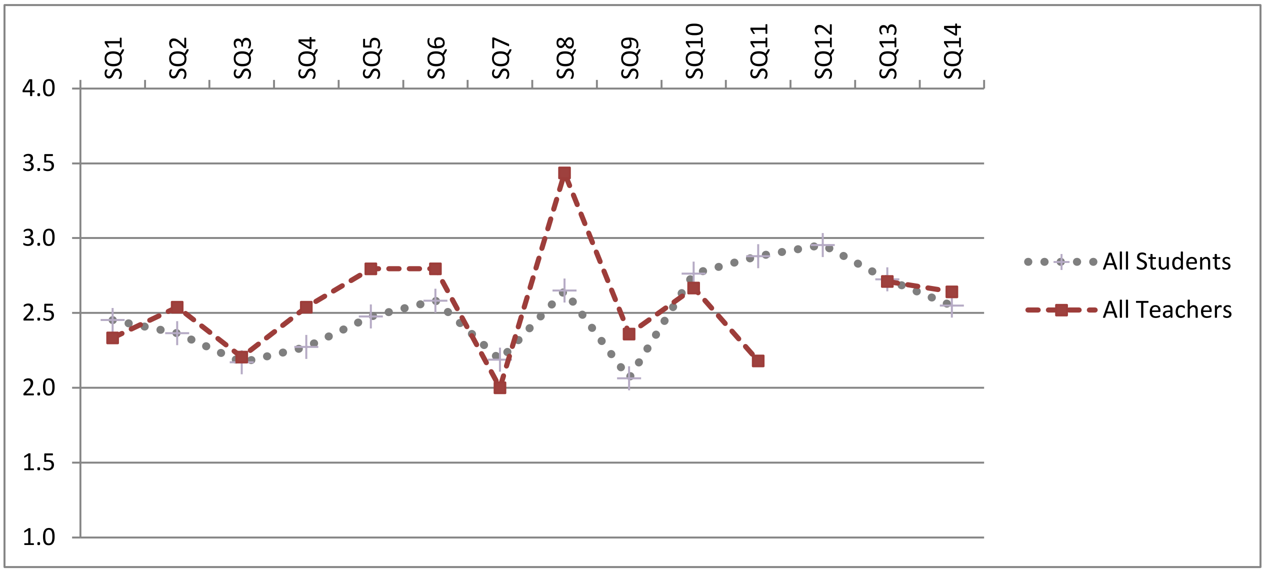
One of the most interesting observations gleaned from this study is believed anecdotally to be true but did not have strong evidence to support the assumptions. This data is the disconnect between what teachers think they are doing and their actual practices. Figure 9 and Figure 10 show that teachers say that they are relating the learning to daily lives at a rate higher than the students are. Teachers in mathematics answer that they are giving small group work opportunities more than what the students perceive. Students in mathematics perceive they are working on problems alone more than what the mathematics teachers reported. In science (Figure 10), students answered that they are listening to lecture-style presentations at a greater rate than the teacher reported.



Scale: 4 - Almost every class, 3 - About half the classes each week, 2 - Some classes each week, 1 - Never

MQ1	We practice adding, subtracting, multiplying, and dividing without using a calculator.
MQ2	We work on fractions and decimals.
MQ3	We interpret data in tables, charts, or graphs.
MQ4	We write equations and functions to represent relationships.
MQ5	We work together in small groups.
MQ6	We relate what we are learning in math to our daily lives.
MQ7	We explain our answers.
MQ8	We decide on our own procedures for solving complex problems.
MQ9	We review our homework.
MQ10	We listen to the teacher give a lecture-style presentation.
MQ11	We work problems on our own.
MQ12	We begin our homework in class.
MQ13	We have a quis or test.
MQ14	We use calculators.

Figure 9: Student and teacher response data for TIMSS survey questions in mathematics.



Scale: 4 - Almost every class, 3 - About half the classes each week, 2 - Some classes each week, 1 - Never

SQ1	We watch the teacher demonstrate an experiment or investigation.
SQ2	We formulate hypotheses or predictions to be tested.
SQ3	We design or plan an experiment or investigation.
SQ4	We conduct an experiment or investigation.
SQ5	We work in small groups on an experiment or investigation.
SQ6	We write explanations about what was observed and why it happened.
SQ7	We study the impact of technology on society.
SQ8	We relate what we are learning in science to our daily lives.
SQ9	We present our work to the class.
SQ10	We review our homework.
SQ11	We listen to the teacher give a lecture-style presentation.
SQ12	We work problems on our own.
SQ13	We begin our homework in class.
SQ14	We have a quis or test.

Figure 10: Student and teacher response data for TIMSS survey questions in science.

Discussion

To further discuss the findings from the study, the research question should serve as the guide: What factors can be identified that contribute to the disparity between EXPLORE scores and state assessment data? This will be discussed through the combined lens of the original theme with the emerging categories. Wrapping up the discussion, will be a "future directions" section where the discussion will continue with implications for the future in practice and for research.

Theme/Category 1: Leadership and Planning
The analysis shows that without long-range (2- to 5-year) goals for mathematics and science programs, there is no combined effort by teachers, administrators, and the community to support improvements. In addition, without long-range priorities, there is little guidance for purchasing and professional development priorities. Even with very limited budgetary resources, priority needs can be presented as the administration and community's part of accountability for improved instruction.

Findings include:

- Few schools have specific goals for mathematics or science programs beyond the present testing period.
- Few schools have professional development plans specifically related to mathematics and/or science instruction.
- Most principals have not been involved in professional development programs that address quality and innovations for mathematics and science programs.
- Few schools' plans have clearly identified goals and strategies for improving instruction, especially addressing gap groups.

Evidence suggests planning for the improvement of mathematics and science programs is primarily short-term and related directly to state assessment scores. There appears to be little or no effort to align budgets and professional development with long-range goals specific to mathematics and/or science. Implementing professional development that is meaningful and content-specific would help teachers be more confident of the subject they teach and able to more effectively plan and actualise higher-order thinking in the classroom.

Theme/Category 2: Instruction
The analysis shows the gap between teachers' perceptions of what happens in their classrooms differs from what students report and from what researchers observed. This could mean a lack of resources, understanding, and time to teach at higher cognitive levels of complexity. In addition, as with assessment, this gap calls for more focused professional development. However, for instruction, the data also suggest the need for an examination of programs and available materials and equipment essential for improvement.

Findings include:

- Only a few mathematics and science teachers always find one or more real-world applications within the content covered.
- In just a few of observed mathematics classrooms, students were applying mathematics to realistic problems.
- In just a few science classrooms, students were engaged in investigations at higher levels, with most completing pre-planned activities requiring little or no analysis.

- In observed mathematics classrooms, only a few provided evidence of student-centered learning (student presentation, small group discussion, students solving problems). In most of the classrooms, students worked in whole group or individually on the same assigned task.
- In observed science classrooms, only a few provided evidence of student-centered learning (student presentation, small group discussion, student investigation).
- Classroom observations indicate instruction is teacher-led and centers on mastery of factual information or reviewing content.
- Teachers and students have disparate views on whether presented math/science content or activities relate to their daily lives as well as the frequency of teacher lecture in science vs. inquiry-based instruction.

Depth of understanding is important if a student is going to gain the skills necessary to think critically. As early as 1954, Bloom stated, "a student who understood a concept could apply it." Too often teachers depend upon mere rote memorisation or other shallow learning skills to "get through" a large body of information. Gardner (1991) echoed this need to grasp concepts in such a way that they can be used to solve problems. Students who are busy completing worksheets and answering questions are not necessarily mentally "on task". It is possible to do lower cognitive skills with little mental effort, which translates into a lack of understanding.

Theme/Category 3: Assessment
The analysis shows, as with instruction, that the identified gap between what teachers believe relevant to their assessment of higher-level thinking tasks and what was observed in their classrooms suggests the need for 1) a clearer common understanding of higher-level task design, 2) more resources for assessing higher-level tasks, and 3) professional development related to assessment for higher-level thinking tasks.

Findings include:

- Teachers report their tests involve application of concepts in mathematics and use of hypotheses and conclusions in science.
- In few mathematics classrooms, observed teachers utilised questioning techniques that led to higher level or divergent thinking; the use of critical thinking in general was observed in just a few of classrooms.
- Similarly, in just a few science classrooms, observed teachers utilised questioning techniques to stimulate higher level or divergent thinking; the use of critical thinking in general was observed in very few classrooms.

A gap exists between what teachers believe, what they report, and what was observed relative to classroom assessment. Gaps also exist between what students and teachers report is assessed in the classroom. Many teachers believe that they are administering assessments that require higher order thinking skills when in fact the questions require simple responses. Few assessments are given that evaluate process skills or inquiry-based learning. Producing assessments that are outside of the typical true/false, multiple choice type question requires training, practice, and time. If higher-order skills are to be implemented in the learning process, then they must also be included in the assessments.

Category 4: Disconnect between teacher perspicacity of practices and actual teacher practices

A gap exists between what teachers believe about what should happen in math/science classrooms, what they perceive happens in their classrooms, and what researchers observed within their classrooms. In addition, there are disparities between what teachers and their students say is happening in the mathematics and science classroom. This is especially found in an anecdotal follow-up visit with the teachers when going over the results. Time after time, the teachers would say that any negatively-perceived data "did represent the region" but "did not represent their own individual school. If teachers do not recognise that they are deficient in producing planning, lessons and assessments that are rich in higher order thinking skills, then they are unlikely to seek professional development opportunities that lead to upper level cognitive complexity.

Limitations

Schools were involved in multiple improvement initiatives. It is difficult to know the impact that this study had on the schools. Furthermore, some of the data came to the researchers as unidentified information. This made it impossible to add a portion of data collected to the analyses.

Future Directions

The data reveal a challenge: Closing an existing gap between teacher beliefs and what researchers observed in classrooms regarding classroom planning, instruction and assessment. The state assessment and EXPLORE tests cover a broad range of topics. Teachers say they are pushed to keep up with curriculum maps and pacing guides that address materials over which their students will be tested on the state assessment and do not have a lot of time to spend on higher-level thinking. To complicate matters and give this research a sense of urgency, the Common Core Standards for mathematics, and soon to come science, will create another level of complexity to the change needed.

The plan for implementation of the study presented in this paper was completed in March of 2010. To utilise the findings, two plans have been proposed for continued development and research – one with NSF support and another plan to implement new programs and strategies with local and state support. The philosophy of the project planners is that outside funding and support should be for development and training for implementation of new programs and strategies that can be sustained with ongoing operational budgets.

This grant initiative's Task Force recommendations included implementing rigorous professional development for K-16 teachers, improving teacher preparation and encouraging people with STEM/ED degrees to enter teaching, revolutionising how STEM/ED subjects are taught, learned, and assessed, and engaging business, industry, and civic leaders to improve STEM/ED education.

To assist with improving leadership and planning, the work by the Regional Education Laboratory (REL) Southwest (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007) and Blank and de las Alas (2010) should be considered. Each conducted a meta-analysis of studies that linked professional development with student achievement. The longest time span of the 25 studies reported was 100 hours of teacher contact over ten months (McCutchen, Abbott, Green, Beretvas, Cox, Potter, Quiroga, & Gray, 2002). In addition, assessments aligned with the focus of the PD are needed to achieve larger effect sises (Saxe, Gearhart, & Nasir, 2001; Jaglelski, 1991) (e.g., if the focus of PD is higher level learning, then student assessments should measure student learning at higher levels). A third finding was the importance of mentoring. Snippe (1992) and Jaglelski (1991) found that when PD program offers mentoring, teachers are more likely to become engaged in a new practice that provides better instruction for their students.

To improve instruction, the use of Master Teachers as mentors is being explored who will help teachers implement 5E lesson plans. The 5E Instructional Model of instruction is an effective instructional model for mathematics and science. Lawson (1995) completed a review of more than 50 research studies conducted on the Science Curriculum Improvement Study Learning Cycle (three phases), the 5E Lesson Plan, and some variations of the learning cycle. The results were significant in favour of the program using the 5E model. Using these types of models could increase the cognitive complexity of instructional learning in the classroom.

Another strategy to improve higher-order thinking skills in instruction is the use of Lesson Studies. This is a cycle of instructional improvement in which teachers work together to develop goals to (1) improve student learning; (2) collaboratively plan a "research lesson" to bring to life these goals; (3) observe the lesson being taught in a real classroom setting while gathering evidence on student learning; (4) reflect on evidence to improve the lessons; and (5) revise, teach, and improve the lesson again in additional classrooms (Lewis, 2002). In Japan, Lesson Study may be sponsored by schools or professional organisations (Lewis, Perry, & Murata, 2006; Lewis & Tsuchida, 1998; Takahashi, 2003). However, the sole purpose of Lesson Study effort is building the capacity of teachers to improve their instruction that can lead to higher order thinking skills (Fernandes & Yoshida, 2004; Lewis, 2002; Lewis & Tsuchida, 1997; Yoshida, 1999).

These findings about the need to improve assessments are not surprising as Guskey (2003) found that most teachers use pre-written test questions from their text or teaching resources without much thought put into the level of cognitive complexity. The change must come in the form of making assessment a process that extends through the planning and learning process and involves higher-order thinking skills. Wiggins (1998) has suggested that assessment should not just be produced after instruction, but should be an integral part of instruction. Until critical thinking is implemented in all levels of instruction, including assessment, then the learning gap will continue.

Conclusion

The data from the EXPLORE assessment, coupled with students' responses to questions about learning opportunities, clearly show that middle school students need interventions of more and better instruction in advanced mathematics and science concepts and processes. Middle school students need more opportunity to develop their capacity for higher-level thinking and more exposure to research-based learning strategies that prepare them for rigorous mathematics and science courses in high school – and then college. Cognitive rigor must be included in all aspects of curriculum and instruction including planning, instruction, and assessment. By increasing critical thinking at all levels and implementing more complex cognitive activity, the gap between students who are prepared for high school and those who are prepared for college, as measured by state assessments and EXPLORE tests, will be minimised.

On the international level, mathematics and science education needs to be turning an eye toward how teacher perceptions may be getting in the way of forward movement in breaking barriers in improving education. If teachers perceive themselves as inquiry-based teachers, for example, yet are simply low-level cognitive instructors, there is a disconnect in perceived versus implemented delivery. Only students will suffer from this mismatch – no matter where in the world this is happening. It is time to recognize this challenge, turn it around, and change a disconnect to a seamless and aligned learning experience for students.

References

ACT. (2008). *The forgotten middle: Ensuring that all students are on target for college and career readiness before high school*. Iowa City, ACT.

ACT. (2011). *Explaining what college readiness scores mean*. Retrieved October 9, 2011, from http://www.act.org/standard/index.html.

Blank, R., & de las Ales, N. (2010). *Effects of teacher professional development on gains in student achievement: How meta analysis provides scientific evidence useful to education leaders.* Evanston, IL: Society for Research on Educational Effectiveness.

Bloom, B. (1954). *Taxonomy of educational objectives: Book 1-Cognitive domain.* White Plains, NY: Longman.

Council of Chief State School Officers Mathematics and Science Education Task Force. (2006). *Mathematics and Science Education Task Force Report and Recommendations*. Washington DC: Council of Chief State School Officers. Retrieved October 1, 2011, from http://www.ccsso.org/content/pdfs/Math%20Science%20Recom%20FINAL%20lowres.pdf.

Fernandes, C., & Yoshida, M. (2004). *Lesson Study: A case of a Japanese approach to improving instruction through school-based teacher development*. Mahwah, NJ: Lawrence Erlbaum.

Finkelstein, N. D., & Fong, A. B. (2008). *Course-taking patterns and preparation for postsecondary.* (Issues & Answers Report, REL 2008-No. 035). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education and Regional Assistance, Regional Education Laboratory West.

Gardner, H. (1991). *The unschooled mind: How children think and how schools should teach*. New York: Basic Books.

Guskey, T. (2003). How classroom assessments improve learning. *Education Leadership, 60*(5), 6-11.

Harmon, H.L., Henderson, S. A., & Royster, W. C. (2003). A research agenda for improving science and mathematics education in rural schools. *Journal of Research in Rural Education, 18*(1), 52-58.

Henderson, S. (2007). Mathematics classroom assessment instrument. Learning By Design, Briarwood Enterprises. Richmond, KY.

Henderson, S. (2007). Science classroom assessment instrument. Learning By Design, Briarwood Enterprises. Richmond, KY.

International Association for the Evaluation of Educational Achievement (IEA). (2003). *International report on achievement in the mathematics cognitive domains: Findings from a developmental project*. TIMSS & PIRLS International Study, Lynch School of Education, Boston College.

Jaglelski, D. A. (1991). *An analysis of student achievement in mathematics as a result ofdirect and indirect staff development efforts focused on the problem-solving standard of the National Council of Teachers of Mathematics*. (Doctoral Dissertation, Loyola University of Chicago, 1991). (UMI No. 9119821).

Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs* (2nd ed.). Newbury Park, CA: Sage.

Lawson, A. E. (1995). *Science teaching and the development of thinking*. Belmont, CA: Wadsworth Publishing Company.

Lewis, C. (2002). *Lesson study: A handbook of teacher-led instructional change.* Philadelphia: Research for Better Schools.

Lewis, C., Perry, R., & Murata, A. (2006). How should research contribute to instructional improvement? The case of lesson study. *Educational Researcher, 35*(3), 3-14.

Lewis, C., & Tsuchida, I. (1997). Planned educational change in Japan: The shift to student-centered elementary science. *Journal of Education Policy*, *12*(5), 313-331.

Lewis, C., & Tsuchida, I. (1998). A lesson is like a swiftly flowing river: Research lessons and the improvement of Japanese education. *American Educator* (Winter), 14-17, 50-52.

McCutchen, D., Abbott, R. D., Green, L. B., Beretvas, S. N., Cox, S., Potter, N. S., Quiroga, T., & Gray, A. L. (2002). Beginning literacy: Links among teacher knowledge, teacher practice, and student learning. *Journal of Learning Disabilities*, *35*(1), 69-86.

Mullis, I.V.S., Martin, M. O., & Foy, P. (2005). International report on achievement in the mathematics cognitive domains: Findings from a developmental project. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine of the National Academies. (2005). *Rising above the gathering storm: Energising and employing America to a brighter economic future*. Washington DC: National Academies Press. Retrieved August 31, 2011, from http://www.nap.edu/execsumm_pdf/11463.pdf

National Center on Education and the Economy. (2007). *Tough choices, tough times: The report of the new commission on the skills of the American workforce*. Washington, DC: National Center on Education and the Economy. Retrieved August 31, 2011, from http://skillscommission.org.

National Science Teachers Association. (2008). *The condition of education 2008*. NSTA Web News Digest. Retrieved September 7, 2011, from http://www.nsta.org/publications/news/story.aspx?id=54947.

Saxe, G., Gearhart, M., Nasir, N. S. (2006). Enhancing students' understanding of mathematics: A study of three contrasting approaches to professional support. *Journal of Mathematics Teacher Education*, *4*, 55-79.

Silver, E. A. (1998). *Improving mathematics in middle school: Lessons from TIMSS and related research*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Snippe, J. (1992, July). Paper presented at the annual meeting of the American Educational Research Association, San Francisco. Unpublished manuscript.

Takahashi, A. (2003, July). *Lesson study overview: Three major types of lesson study*. Paper Presented at the Lesson Study Immersion Program, Tokyo, Japan.

Treffinger, D. (2008). Preparing creative and critical thinkers. *Educational Leadership*, *65*.

Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey-Bass.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007-No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved September 1, 2011, from http://ies.ed.gov/ncee/edlabs.

Yoshida, M. (1999). *Lesson Study: A case study of a Japanese approach to improving instruction through school-based teacher development.* Doctoral dissertation, University of Chicago.