EDITORIAL

In this thirteenth issue of the Colloquium Journal I am very pleased to introduce our new authors.

The feature article by Ellen Boiselle is about learning difficulties in mathematics. Ellen asserts that standardized achievement tests alone reveal little information as to the nature and quality of performance or the reasons why a child experiences difficulty. Her research explored whether systematic observation of specific behaviors – or qualitative features of performance such as Pace, Precision, Management of Complexity, and Self-Efficacy – which accompany and give rise to performance on arithmetic tasks might be used in conjunction with standardized tests of arithmetic achievement to better describe a child’s areas of vulnerability in mathematics. She has found that difficulties with qualitative features are associated with low scores on the Numerical Operations and correlated with low scores on neuropsychological measures of cognitive function. Only the qualitative feature of Self-Efficacy is associated with teacher report of child performance. These results suggest that each qualitative feature may be considered a behavioral manifestation of underlying neuro-cognitive processes as they relate to the particular demands of arithmetic.

Almost all teaching and learning takes place using the medium of language. Chemistry students experience difficulty with words that have one meaning in chemistry (scientific meanings) and another meaning in everyday language. While high school chemistry content has its specific vocabulary, rules, concepts, models, and theories, students come into the chemistry classroom with their own personal vocabulary, rules, and concepts about the meanings that are different from the meanings used in the chemistry curriculum. This creates an obstacle for successful learning of chemistry.

Peggy Labrosse’s study analyzed the effects of specific vocabulary instruction on high school chemistry students’ knowledge and understanding. She implemented The Frayer Model, a research-based teaching strategy for specific vocabulary instruction, as a treatment to learn more about vocabulary knowledge and understanding of chemistry content.

In recent decades, reform documents have highlighted the need to improve science education in the United States. These documents insist that teachers must increase their knowledge of the content they teach and improve their instructional methods. National Research Council calls for greater use of the inquiry method in teaching science. Educators have designed professional development programs to improve science teaching, but these programs have shown only limited success. In order to be effective, professional development programs need to be designed with an understanding of how teachers learn.

Wesley Johnson is studying science teachers’ professional learning experiences that enhance their instructional practices and raise their students’ achievement. He is collecting materials about existing professional development programs to determine the characteristics of teachers’ learning experiences which are most likely to lead to increased student achievement.

Qualifying Paper is the concluding component of the Ed.D course work and is considered as groundwork of the future dissertation proposal. The purpose of the Qualifying Paper is to provide an opportunity for advanced Ed.D. students to make evident their academic growth and to reveal their readiness to develop scholarly research. Students are expected to demonstrate their ability to analyze, synthesize, evaluate, and conceptualize the research literature as it applies to their individual interests. They must be able to develop a rational model of their understanding of theoretical constructs and ideas, and demonstrate their ability to work independently. The literature review is not a pro forma part of a research publication and not a respectful acknowledgement of those who have done related work, but rather a record of the researchers’ attempts to answer the questions using the existing literature.

In this issue I would like to extend the practice of publishing some recent Qualifying Papers in the Educational Resources section. First, I want to acknowledge the work of our students and second, to display the examples of writing for those who are facing Qualifying Examinations. As Philosopher Daniel Dennett said that science is “making mistakes in public. ... Making mistakes for all to see, in the hopes of getting the others to help with the corrections.”

Regina Panasuk
GUIDELINES FOR SUBMISSION
The papers submitted for the Journal must discuss psychological and pedagogical issues and trends related to mathematics and science education.

WHEN SUBMITTING A PAPER, PLEASE USE THE FOLLOWING GUIDELINES:
1. Submit an electronic version of the paper, an abstract, approximately 150 words, and a biographical sketch, about 30 words. All pictures and diagrams must be submitted in a separate document.
2. Use double spacing with one-inch margins.
3. For references, tables, and figures follow the style described in the Publication Manual of the American Psychological Association (APA), Fifth Edition.
4. Paper length must not exceed 30 pages, including pictures, tables, figures, and list of references.
5. Paper must be submitted by November 15.
6. Authors will be notified about the status of their papers by January 15.
7. The Colloquium is scheduled in April.

SUGGESTIONS TO THE AUTHORS:
When preparing a research paper include:
a) a rationale and an identification of the research question(s)
b) a conceptual framework or brief statement of relationship to the literature
c) an identification of research methodology
d) a summary of the analytical technique(s)
e) a summary of preliminary findings.

SUBMIT PAPERS AND CORRESPONDENCE TO:
Dr. Regina M. Panasuk
Professor of Mathematics Education
Graduate School of Education
University of Massachusetts Lowell
61 Wilder Street, O’ Leary 5th Floor
Lowell, MA 01854
Phone: (978) 934-4616
Fax: (978) 934-3005
Regina_Panasuk@uml.edu
# TABLE OF CONTENTS

Assessment of Learning Difficulties in Arithmetic: Looking Beyond Standard Scores  
*Ellen Boiselle* …………………………………………………………………………………………………………………………………………………

Analysis of the Effect of Specific Vocabulary Instruction on High School Chemistry Students’ Knowledge and Understanding  
*Peggy LaBrosse* …………………………………………………………………………………………………………………………………………………

The Characteristics of Effective Teacher Professional Learning  
*Wes Johnston* …………………………………………………………………………………………………………………………………………………

## EDUCATIONAL RESOURCES

### Qualifying Papers

Can Participation in a School Science Fair Improve Middle School Students’ Attitudes Toward Science and Interest in Science Careers?  
*Valerie Finnerty*

Involving Scientists in K–12 Education: Influencing Students’ Understanding of the Nature of Science  
*Michael Wadness*

Use of Portable Technology in a Teacher Preparation Program for the Purpose of Enhancing Teachers’ Understanding of Scientific Inquiry  
*Sachiko Tosa*
CONTRIBUTORS

Ellen Curtis Boiselle is a math specialist in the Learning Disabilities Program, Department of Behavioral Neurology, Children’s Hospital, Boston and a Research Associate in the Department of Neurology, Harvard Medical School.

Peggy LaBrosse completed her Doctor of Education in Science at the University of Massachusetts Lowell in 2007. She has been teaching Chemistry for almost 32 years. She is also teaching education courses in the Master of Education Program and science courses to undergraduate students at Franklin Pierce University. Peggy is actively involved in curriculum development and in the programs that support disengaged high school students.

Wesley Johnson currently works as a Physics teacher at Burlington High School in Massachusetts. He has also worked as the Planetarium Director at Nashua High School South, presenting interactive shows to students of all ages. He is currently working on his dissertation proposal at the University of Massachusetts Lowell’s Graduate School of Education.

Valerie Finnerty has been an 8th grade science teacher at Littleton Middle School for three years. Prior to that, she taught 6-8 grade science and math at Oak Meadow Montessori for five years, and prior to that, she taught emotionally disturbed 4th and 5th graders in the Boston Public Schools for 10 years. She is currently working on her dissertation at the University of Massachusetts Lowell’s Graduate School of Education.

Mike Wadness has been teaching physics for almost 10 years at Medford High School. He is working on his proposal.

Sachiko Tosa earned a PhD in Physics from the University of Rochester in 1986. She has been teaching science courses for children at the Museum of Science and mathematics for at the Japanese Language School of Greater Boston. In 2005, her book for children on the history and theory of flight was chosen as an essay-contest book in Japan, and 60,000 copies were sold. She is a doctoral candidate in Science Education and currently working on her proposal.

2006-2007 Academic Year
Mathematics and Science Education Program

DISSEMINATION PROPOSAL STAGE

<table>
<thead>
<tr>
<th>Kate McLaughlin</th>
<th>Sachiko Tosa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alex Ballantyne</td>
<td>Wesley Johnson</td>
</tr>
<tr>
<td>Valerie Finnerty</td>
<td>Neil Hatem</td>
</tr>
<tr>
<td>Matthew Beyranevand</td>
<td>Jaimie Stone</td>
</tr>
<tr>
<td>Mike Wadness</td>
<td></td>
</tr>
</tbody>
</table>

For information about Graduate Programs in Mathematics and Science Education at UML, please contact the appropriate faculty member:

Vera Ossen, Ed.D. | David Lustick, Ed. D.
Vera_Ossen@uml.edu | David_Lustick@uml.edu
The Director of Educator Licensure Programs | Science Education
Anita Greenwood, Ed.D. | Regina Panasuk, Ph.D.
Anita_Greenwood@uml.edu | Regina_Panasuk@uml.edu
The Department Chair | Mathematics Education
Science Education

Application packets may be obtained on line or by contacting the Graduate School of Education at (978) 934-4601
Assessment of Learning Difficulties in Arithmetic: Looking Beyond Standard Scores

Ellen Curtis Boiselle
Children’s Hospital, Boston

ABSTRACT
A challenge confronting special educators and school psychologists involves the assessment of learning difficulties in mathematics. This paper argues that educational professionals can improve their assessment practices by using seven qualitative features of performance that can be observed as a student works on arithmetic tasks. The seven features include: Precision, Pace, Graphomotor Output, Visuospatial Skills, Language Demands in Math, Management of Complexity in Math, and Self-efficacy in Math. This paper defines these features and describes research supporting their utility as diagnostic tools. The central argument of this paper is that a consideration of these seven features can enhance information derived from standardized test scores in arithmetic, identify factors that may interfere with a student’s performance in arithmetic, and provide direction for interventions and accommodations. Guidelines for how educational professionals might incorporate qualitative features of performance into their assessment practice and case studies are also discussed.

INTRODUCTION
A key challenge confronting special educators and school psychologists involves the assessment of learning difficulties in mathematics. A majority of professionals use standardized achievement instruments such as the Woodcock Johnson III (WJIII) and the Wechsler Individual Achievement Test II (WIAT II) to evaluate a student’s competence in mathematics. The scores generated from these measures, however, do not always fully reveal the nature of the student’s difficulties, nor do they yield suggestions for intervention.

The focus of this paper is on seven qualitative features of performance that can be observed as a student works on arithmetic tasks. These features have been described by Marolda and Davidson (1994, 2000) and include: Precision, Pace, Graphomotor Output, Visuospatial Skills, Language Demands in Math, Management of Complexity in Math, and Self-efficacy in Math. This paper builds on the work of Marolda and Davidson by 1) operationalizing these features, 2) delineating their impact on a student’s learning and performance in arithmetic, and 3) describing research supporting their utility as diagnostic tools.

This paper begins with a brief review of current practice in the assessment of learning difficulties in mathematics. Next, each qualitative feature of performance is discussed, and strategies for how educational professionals can incorporate these features into assessment practice are described. In the third section, findings from recent research (Boiselle, 2007) investigating the association of each feature with neuropsychological measures are presented. The fourth and final section of the paper provides two brief case studies that illustrate how qualitative features of performance can be used in conjunction with standardized measures of arithmetic to enhance the quality of assessment. General implications of this work and directions for future research are also discussed.

ASSESSMENT OF LEARNING DIFFICULTIES IN MATHEMATICS: CURRENT CHALLENGES

The instruments most commonly used by psychologists and special educators for assessing a student’s learning difficulties in mathematics are the math subtests from the WJIII (Math Fluency, Calculations, Applied Problems, and Quantitative Reasoning) and the WIAT II (Numerical Operations and Mathematics Reasoning). These measures provide information about how a student’s performance compares with that of a nationally normed group. This information is frequently used to determine whether a student has a math disability and is eligible for special education services. Often, however, individuals who perform within the average range on achievement tests such as the WJIII and WIAT II struggle with math in the classroom setting and perform poorly on state-mandated achievement tests. Therefore, what is an “average” performance on the national level may not be “average” in a particular setting. Moreover, because performance on these standardized batteries is measured solely by the accuracy of response, these measures yield little information as to the nature and quality of the student’s performance or the reasons why the student experiences difficulty. Although the authors of the WJIII and the WIAT II herald the diagnostic utility of these measures, their function as clinical tools is limited.

Criticism of the use of standardized tests to assess an individual’s cognitive and educational profile has a long and distinguished history. Over seventy years ago, developmental theorist Heinz Werner cautioned that
conclusions based entirely upon achievement scores could be quite misleading (Werner, 1937). A central tenet of Werner’s criticism was the observation that individuals can arrive at a solution via distinctly different processes. Werner maintained that these different processes, which are observable at the behavioral level, are manifestations of underlying differences in the central nervous system and therefore relevant in any assessment of an individual’s learning and achievement. Werner argued for a distinction between process and achievement. He defined process as those observable behaviors and thinking strategies that give rise to a given solution. He defined achievement as the accuracy of that solution. Werner advocated for the inclusion of process in assessment, arguing that a focus on process can provide a more accurate portrait of an individual’s learning and performance in a given domain.

This emphasis on process—on how an item is passed or failed in addition to whether it is passed or failed—is familiar to many educators and clinicians today. Current emphases in mathematics education encourage teachers to “move beyond a superficial ‘right or wrong’ analysis” to consider factors that give rise to a given performance (National Council of Teachers of Mathematics [NCTM], 2000, p. 24). Teachers often describe qualitative features of performance in the narrative section of report cards. Educators like Richardson (2003) encourage teachers to record thinking behaviors that students exhibit as they work on tasks. Neuropsychologists routinely incorporate a consideration of process into the assessment of cognitive function (Kaplan, 1990).

A focus on process is also central to the non-standardized, clinical-interview approach to the assessment of learning difficulties in mathematics advocated by Ginsburg (1997) and Marolda and Davidson (1994, 2000). In the clinical interview, clinicians focus on qualitative features that illustrate how the student thinks about, approaches, and solves problems, as well as the accuracy of the answers themselves. Recently, the authors of the WIAT II incorporated a checklist of “qualitative observations” designed to “provide additional information that can guide the development of effective instructional plans” (Psychological Corp., 2005, p. 2). The incorporation of such a checklist signals an important shift in clinical assessment of math competencies, suggesting that a consideration of qualitative features of performance is gaining currency in this field.

Yet despite this growing recognition of the importance of process, standard scores, percentile ranks, and grade equivalents remain the cornerstone of clinical assessment of learning difficulties in mathematics. When a student scores in the average range on a standardized achievement test, it is assumed that learning and performance in mathematics is not problematic and that any reported difficulties reflect affective or behavioral issues. It can be argued, however, that a consideration of “process” is a critical component in the assessment of learning difficulties in mathematics, as it may reveal a more comprehensive picture of a student’s strengths and vulnerabilities than that provided by scores alone and provide information helpful for educational planning.

But how can an educational clinician, school psychologist, or teacher incorporate a consideration of process features into the assessment of students who demonstrate learning difficulties in mathematics, specifically arithmetic? What features should be examined? How are these features evident in arithmetic? And what interventions can be implemented to address the revealed features?

QUALITATIVE FEATURES OF PERFORMANCE IN ARITHMETIC

Marolda and Davidson (1994, 2000) describe qualitative features of mathematics performance which can be used in a diagnostic context. Seven of these features include: Precision, Pace, Graphomotor Output, Visuospatial Skills, Language Demands of Math, Management of Complexity in Math, and Self-efficacy in Math. These features have been an important component of interview-based evaluations of math learning difficulties conducted at the Learning Disabilities Program, Children’s Hospital, Boston for over thirty years (Davidson, 1983; Marolda & Davidson 1994, 2000). In order for these features to be useful to clinicians and educators, however, they need to be operationalized and their utility as diagnostic tools must be examined (Boiselle, 2007). Specific examples of each feature, their impact on learning and performance in arithmetic, and recommendations for instructional practice are all described in the following section.

DEFINITIONS AND EXAMPLES OF QUALITATIVE FEATURES OF PERFORMANCE IN ARITHMETIC AND RELATED RECOMMENDATIONS

Precision

Precision in arithmetic refers to the ability to generate a precise solution to a given arithmetic task. Conversely, difficulties with precision — or imprecision — refers to difficulties arriving at exact solutions. There
are three types of imprecision in arithmetic: imprecision in facts, imprecision in reading operation signs, and imprecision in multi-digit calculations. Students who demonstrate imprecision in facts have difficulty generating accurate solutions to one-digit facts (e.g., $6 + 7 = 14; 5 + 8 = 12$). Students who demonstrate imprecision in reading operation signs apply one operation to one or more arithmetic examples regardless of the operations indicated (e.g., $12 – 8 = 20; 2 \times 3 = 6; 10 ÷ 2 = 12$). Frequently, this type of imprecision is observed on tasks such as the WJIII Math Fluency subtest that require that a student switch between addition, subtraction, and multiplication examples.

Imprecision in multi-digit calculations is evident in inaccurate solutions to multi-digit arithmetic problems (e.g., $85 – 29 = 57; 127 + 546 = 683$). This type of imprecision can be secondary to imprecision with one-digit facts, described above. Alternatively, this type of imprecision can reflect difficulties with complexity. Some students demonstrate accuracy with one-digit facts in isolation but experience difficulty maintaining precision with those same facts when pursuing multi-digit arithmetic. This type of imprecision reflects difficulties with the management of complexity, a qualitative feature described in more detail later in this paper.

In the classroom setting, students who demonstrate imprecision may receive low scores on standardized tests and other evaluation instruments, regardless of their conceptual facility or their understanding of the strategy they are implementing. Students who struggle with precision may be deemed careless, lazy, or cavalier when, in fact, their difficulties stem from inherent aspects of their cognitive profile. These students tend to become frustrated in that they often have an understanding of the process at hand yet perform inconsistently on homework, tests, and quizzes. This frustration may, in turn, occasion self-doubt in the learner regarding mathematical competence and/or dislike of mathematics.

To help students who struggle with precision, teachers may want to score work twice, once for the appropriateness of approach, and once for the accuracy of solution. The double scoring approach may give credit for conceptual understanding at the same time it helps the student to become more aware of the imprecision. Teachers can also provide students with specific strategies to help monitor for imprecision, such as checking the accuracy of multi-digit subtraction calculations by adding the difference to the subtrahend. Students who demonstrate imprecision with operation signs should be encouraged to circle the operation sign and to verbalize the operation (e.g., “I have to subtract”) prior to calculating. Finally, students who struggle with imprecision may benefit from learning specific strategies to help generate accurate solutions to facts (Leutzinger, 2002).

**Pace**

At a general level, pace refers to the speed with which one processes and/or produces information. Regarding academic performance, pace is typically conceptualized as a measure of fluency with a given skill. Difficulties with pace—or slow pace—characterize a performance on a given set of arithmetic tasks that is markedly slower, or dysfluent than that expected given a child's age and grade level. Currently the only standardized measure of arithmetic to incorporate a measure of fluency is the Math Fluency subtest from the WJIII, which assesses a student's efficiency with simple one-digit addition, subtraction, and multiplication facts. A careful review of this subtest, however, reveals that the one-digit facts included represent a comparatively easy subset of facts. Children with learning difficulties in mathematics typically solve these types of problems quickly but struggle with more difficult examples, such as $7 + 6$ and $13 – 5$ (Baroody & Tillikainen, 2003). Given the relative simplicity of the one-digit facts included in the Math Fluency subtest, it is possible that results may provide an over-estimation of a student's efficiency and accuracy generating solutions to one-digit facts.

One manifestation of slow pace in arithmetic is a slow rate of producing solutions to one-digit facts. A second manifestation of slow pace is an overly deliberate pace in pursuing multi-digit calculations. This type of slow pace is sometimes, though not always, secondary to difficulties managing complexity.

Pace in arithmetic is important. Computational fluency is a stated objective of mathematics education in the middle elementary and middle school period (NCTM, 2000). This emphasis on efficiency is consistent with research highlighting the importance of computational fluency as a foundation for success in solving more sophisticated mathematical tasks, including word problems. As Berninger and Richards (2002) note, children who must devote significant cognitive energy to generate solutions to arithmetic problems may not have the necessary resources left over for other “thinking jobs” (p. 108).

Pace in arithmetic is also important in meeting the demands of the classroom. Although students who work at a very slow rate may be able to arrive at accurate solutions, their deliberate pace may hinder their ability to follow along during an instructional session.
These students often have difficulty responding to arithmetic questions posed in class because they need more time to process and produce information than their peers. Timed tests and situations that call for rapid responses are both challenging and frustrating. Students who struggle with pace often have to expend more time and energy to complete arithmetic homework. Sometimes, they are kept inside during recess or after school to complete classwork. These consequences may, in time, give rise to negative feelings about mathematics and result in frustration, avoidance, and/or anxiety.

Educational recommendations for students who work at a slow pace include providing extended time on tests and quizzes and modifying the number of examples a student must complete for in-class assignments and for homework. Teachers should avoid or adjust situations where speed-of-processing is emphasized, such as timed math-fact tests and math games that emphasize speed. Students who work at a slow pace benefit from learning specific strategies which enable efficiency with one-digit facts (Leutzinger, 2002). Finally, in situations where arithmetic demands are secondary to other topics under study, teachers should consider letting the student use a calculator so that the effort and challenge involved in computational demands will not preclude an understanding of the concept being developed.

**Graphomotor output**

Although difficulties with handwriting are often considered detrimental to performance in written expression, they can also hinder a child's performance in arithmetic. Manifestations of this feature include effortful output, illegible written productions, or both. For example, some students produce written output in an arduous and deliberate fashion. Although the written output is legible, the numerals themselves are effortfully produced and sometimes awkward. An example of this is provided in Figure 1.

![Image of student's written output](Image)

**Figure 1. Example of awkward and effortful output.**

Here, a fourth-grade student produced numerals that were legible, but her output was laborious and the numerals somewhat awkward and large for a student of her age.

Difficulties with graphomotor output are also evident in productions that are difficult to read/interpret, such as the example provided by a sixth grade student in Figure 2.

![Image of student's written output with poor legibility](Image)

**Figure 2. Example of graphomotor output of poor legibility.**

These productions are marked by their poor legibility. It is difficult to discern whether these symbols represent the numbers 6, 17 and 23 (as intended) or whether they represent a different set of numbers (e.g., 0, 19 and 22 or 27). Moreover, the faintness with which the writing was produced may be indicative of the student's lack of confidence.

Students who struggle with written output often experience difficulty keeping pace with their classmates and meeting the demands of the classroom context (Levine, Oberklaid, & Meltzer, 1981). They also experience difficulty accurately and efficiently copying information from the board or from textbooks. These students are vulnerable in situations where they must provide a written account of their arithmetic reasoning, as well. The effort devoted to producing symbols may come at the expense of reasoning processes. Students who struggle with graphomotor output are also vulnerable in situations where they must produce numerals in a fast and accurate manner, such as timed math-fact drills. These students may be able to demonstrate efficiency with one-digit facts when allowed to respond orally but perform poorly on assessments that require written output.

Students who have difficulty producing legible output sometimes have difficulty interpreting what they have written. As a result, their solutions are sometimes imprecise, and this may lead to an inconsistent performance on homework, tests, and quizzes. Poor legibility also makes it difficult for teachers to interpret a student's responses.

Recommendations for students who struggle with graphomotor demands include limiting copying demands and providing a scribe when a student must provide a written explanation of the solution to a given problem. Teachers might also consider providing notes of material covered in class so that the student's cognitive energies are directed to the topic of instruction rather than to producing written output. When assessing fact fluency, a teacher should consider asking the student to respond orally rather than in written formats.
so that the burden of written demands does not intrude. Finally, students who struggle with producing legible written output benefit from learning to check over their work to ensure that their written output can be interpreted by others.

**Visuospatial skills**

Visuospatial skills characterize the ability to accurately perceive and manipulate visual stimuli and to coordinate a visual perception with a motor response. Although these skills are commonly thought of in relation to geometry, research has demonstrated their role in learning and performance in arithmetic (Bryant, Bryant, & Hammill, 2000; Rourke, 1993). Some researchers posit that difficulties with visuospatial skills inhibit the effective use of the ‘mental number line’ which supports arithmetic calculations (Landerl, Bevan, & Butterworth, 2004).

Difficulties with visuospatial skills are evident in arithmetic as difficulties accurately aligning digits in multi-digit arithmetic calculations (Bryant et al., 2000), as shown in Figure 3.

![Figure 3. Example of difficulties aligning digits.](image)

As the example in Figure 3 illustrates, visuospatial difficulties can hinder a student's accuracy in arithmetic, regardless of conceptual knowledge or understanding of the strategy being implemented. Students who have difficulty aligning digits often have to erase and re-copy problems, thereby hindering the efficiency of their work. Students with visuospatial difficulties also have difficulty organizing their work on the page. They have difficulty anticipating how much space to allocate for a given problem. As a result, their written work on one calculation runs into that from a previous calculation. Sometimes they “crowd” their written work onto one portion of the page.

Difficulties with visuospatial skills may not be readily recognized on standardized arithmetic measures such as the WJIII or WIAT II. On these tests, each item is presented in a square. The student does not have to copy the problem and does not have to organize written output on the paper. However, these difficulties may be evident in classroom settings or clinical interview formats. Reviewing work samples and observing the student as work on a given set of arithmetic tasks can shed light on whether a student has difficulties with visuospatial skills.

Recommendations relevant to difficulties with visuospatial skills include having the student perform arithmetic using centimeter-square paper, writing one numeral per square. Alternatively, the student can perform calculations on lined paper that it turned at a ninety degree angle so that the lines are vertical. Both these strategies can assist with the alignment of digits. Students who struggle with graphomotor problems also benefit from learning estimation strategies to assess the reasonableness of solutions generated. For example, the student whose work is shown in Figure 3 should be encouraged to generate an estimate before attempting the calculation and to then compare that estimate with the eventual solution. In situations where the estimate and the solution are at odds with one another, the student can check to make sure that digits are aligned correctly.

**Language demands in arithmetic**

The feature of language demands in arithmetic characterizes facility with counting and arithmetic vocabulary, the ability to understand instruction, the ability to explain solution strategies, and the ability to use verbal strategies to help guide calculation. In years past, deficits in language were thought to primarily hinder the development of reading skills, with mathematics skills remaining relatively intact. Research, however, has revealed that deficits in language are associated with deficits in arithmetic achievement (Manor, Shalev, Joseph, & Gross-Tsur, 2001). Also important, the language demands of the mathematics classroom have changed. Current emphasis in mathematics education require that students explain their problem-solving approaches in verbal and written formats. Many state-mandated assessments include open-ended response questions where a student must provide a written explanation of the solution strategy. Communication is now included as a process strand in the NCTM's Principles and Standards for Elementary School Mathematics (2000). Students are expected to engage in conversations about mathematical ideas. Through these conversations, they are supposed to “develop a language for expressing mathematical ideas and an appreciation of the need for precision in that language” (NCTM, 2000, p. 60).
A student’s facility with language can have an impact on learning and performance in arithmetic. For example, students who have difficulty with language demands often struggle with arithmetic terminology. They may be confused by terms such as difference and product, which have precise meanings in arithmetic that are different from their everyday meaning. These students often have difficulty retrieving arithmetic vocabulary, as well.

Difficulties with language demands are also evident in students who struggle to understand instruction and directions. These students are confused by lengthy verbal explanations, particularly those which involve complex syntax or unfamiliar vocabulary. They often need directions simplified and repeated before they understand what is being asked of them. Sometimes these students appear inattentive, lacking in interest, or even defiant when presented with a given question when, in fact, they have not been able to understand the instruction provided.

Students who struggle with language demands have a particularly difficult time explaining their thinking processes in oral and/or written formats. Their explanations are often poorly organized, terse, and confusing for the listener. As a consequence, they have difficulty participating in math discussions. These students also have difficulties using verbal skills (e.g., a verbal script or a mnemonic strategy) to monitor the steps of complex calculations (Bryant et al., 2000).

It should be noted that standardized tests in arithmetic, which are typically limited to paper and pencil formats, offer limited opportunities to assess the quality of language skills. However, educators and clinicians can gain information about a student’s proficiency with this feature by observing how well the student deals with direction and by asking the student to describe approaches to tasks. Difficulties under these demands may reveal underlying issues with language that will affect a student’s availability to instruction and participation in math-related discussions. Information regarding language facility can also be gathered from classroom observations, consultation with the teacher, and review of testing in the speech and language domain.

Recommendations for helping students who struggle with language demands include breaking down complex directions into smaller, more manageable chunks and monitoring textbooks and worksheets for their language content. Teachers should check in with these students frequently to ensure that they understand the information presented. Simplified content and re-explanations should be offered when necessary. Pre-teaching arithmetic vocabulary may also prove helpful. Students with language difficulties often benefit from having the arithmetic situation represented in concrete, visual models (e.g., base-ten blocks, money). The representation and subsequent manipulation of these models can provide a framework to help students organize and guide their efforts on arithmetic tasks. Moreover, the visual model can be retrieved and used as a reference if the formal numerical routine falters. Finally, students who struggle with language demands should be encouraged to communicate solution strategies in non-verbal formats (such as pictures and diagrams) whenever possible.

**Management of complexity in arithmetic**

This feature of management of complexity in arithmetic is defined as the ability to manage and coordinate multiple demands of arithmetic simultaneously. Success in multi-digit arithmetic requires multiple skills, including facility with one-digit facts, accurate and efficient graphomotor output, the ability to recall and implement a sequence of steps, the ability to allocate and shift attention between various sub-components of the arithmetic problem, and the ability to monitor one’s progress. In order to pursue multi-digit arithmetic tasks, students must manage and coordinate all these demands. Students who exhibit difficulties with complexity struggle with this management and coordination.

Difficulties with the management of complexity are evident when a student struggles with a skill in a more complex context but is able to demonstrate proficiency with that same skill in isolation. For example, a student may be able to demonstrate accuracy and efficiency with one-digit facts in isolation but exhibits difficulties with precision and pace when pursuing multi-digit arithmetic. Or, a student may be able to pursue relatively straightforward two-digit subtraction or multiplication problems (e.g., 80 – 57 or 24 x 11) but struggle with the multiple demands inherent on longer, more demanding examples (e.g., 809 – 574 or 247 x 311). These difficulties are often apparent in the student’s written work on arithmetic tasks, which is marked not only by errors but also by multiple erasures and repeated trials.

Another manifestation of difficulty with complexity is when a student fails to maintain the operation throughout a multi-digit calculation (see Figure 4).

Difficulties with the management of complexity are also evident when a student can demonstrate proficiency with a multi-digit calculation skill (e.g., subtraction in circumstances that require regrouping) only to struggle with that same skill in the context of a more complex task (e.g., long division). An example of this is provided in Figure 5.
As the previous examples illustrate, difficulties with the management of complexity can have a significant impact on a student’s learning and performance in arithmetic. Students who struggle with complexity are vulnerable when it comes to demonstrating proficiency in arithmetic situations that require the management of multiple demands. These students often have particular difficulty with multi-digit subtraction where one must regroup across multiple digits, multiplication with a multi-digit multiplier, long division, and fraction operations. Although they may have a generally sound appreciation of the concept and associated algorithm, their performance is often inconsistent and marred by slips in execution. Students who have difficulty with complexity can become frustrated and dismayed by their difficulties. Often, these students perform well in arithmetic during the early elementary years when there is an emphasis on learning one-digit facts. They encounter difficulty, however, during the upper elementary and middle school years, when arithmetic demands become more involved and multi-faceted.

Recommendations for this feature echo many of the suggestions described in the sections on precision and pace, as students who struggle with complexity will have to devote extra time and cognitive energy to compensate for their challenges. More specifically, teachers should consider limiting the number of examples a student must complete, providing extended time allowances, allowing the use of a calculator, and double scoring work — once for approach, once for accuracy of solution. These students also benefit from direct instruction in how to use estimation strategies to assess the reasonableness of their solutions to arithmetic problems. Specific strategies designed to help them monitor and check over their work are helpful as well. Finally, these students benefit from having a “math notebook” that provides examples and directions for how to pursue a given type of arithmetic problem. Such a notebook can be used at home and in class to cue a student if he becomes overwhelmed.

Self-efficacy in arithmetic

Low self-efficacy in arithmetic is characterized by persistent feelings of dismay, frustration, and discouragement about one’s ability to manipulate numbers. Over the course of the pre-school, elementary school, and middle school years, students typically develop feelings and opinions about mathematics. These feelings and opinions arise from a triadic interaction between personal factors (including cognition), experiences (e.g., experiences in the classroom, after-school, home) and environment (e.g., classroom setting,
instructor, curricula) (Bandura, 1986). This interaction affects perceptions of self-efficacy, which Bandura (1997) defines as “beliefs in one’s capabilities to organize and execute the courses of action required to manage prospective situations” (p. 2). Self-efficacy beliefs influence effort, persistence, engagement, and perserverance, and thereby influence academic achievement (Schunk, 1984).

Specific manifestations of low self-efficacy in arithmetic are evident in a student’s emotional responses to arithmetic tasks and in negative self-commentary. Students with low self-efficacy often display excessive dismay, trepidation, anxiety, and/or discouragement when pursuing arithmetic tasks. They frequently respond to a given task by stating “I don’t know” or “I’ll skip this one.” These students tend to deride their arithmetic abilities, making frequent comments such as “I’m stupid with numbers” or “I can’t do math.” They typically feel that their own attempts at learning will not result in positive outcomes and that math is a special subject reserved for a select group of individuals.

The consequences of low self-efficacy in arithmetic include avoidance of situations which require arithmetic and lack of engagement in learning opportunities involving arithmetic. Research suggests that low self-efficacy may also inhibit a student’s thinking processes. Individuals with low self-efficacy beliefs in a given academic area tend to immediately perceive new endeavors as challenging, often before fully appreciating the nature of the task to be solved. As Pajares (1996) notes, this conception fosters stress and a “narrow vision of how best to solve the problem” (p. 543). Conversely, students who believe that they are capable of performing well on academic tasks use more cognitive and metacognitive strategies and persist longer than those who do not (Pintrich & DeGroot, 1990). Low self-efficacy in arithmetic can hinder the individual’s willingness to engage in, practice, and persist on calculation tasks. It can also have an impact on the cognitive resources brought to bear on these tasks.

Teachers can help students who demonstrate low self-efficacy in arithmetic by carefully orchestrating opportunities for success, ensuring that tasks presented are accessible, solvable, and appropriately challenging given the student’s demonstrated level of skills. Students with low self-efficacy also benefit from being able to preview a given topic or set of tasks prior to its introduction in class. This can foster a sense of ownership of the material. Students who struggle in arithmetic often demonstrate competencies in other subtopics of math. Exploring topics in which the student excels and explicitly recognizing those strengths can help to ameliorate the impact of low self-efficacy in arithmetic. Finally, students who are discouraged by their arithmetic competencies sometimes enjoy practicing these skills in game formats. Teachers can use games with dice, dominoes, and chips to teach and provide practice with arithmetic skills. Teachers may also consider non-competitive commercially-available number-related games.

**Relationships between qualitative features**

Although each qualitative feature has been described individually, it is important to consider how they interact with one another. Frequently, students demonstrate difficulty with two or more features. In some cases, these features co-exist alongside one another. In others, one or more features may be secondary to a different feature. For instance, slow pace might be secondary to difficulties with graphomotor output. Imprecision may be secondary to difficulties with complexity. Self-efficacy may be secondary to difficulties with pace and precision. The clinician should consider the full constellation of features, scrutinizing the specific items on which they were observed and considering the interplay between and among features. Only then can a more complete picture emerge.

A few examples will illustrate the interplay between features. Consider the student who demonstrates adequate precision and graphomotor output skills when working on one-digit facts and multi-digit addition and subtraction in circumstances which do not require regrouping. Difficulties with precision and graphomotor output emerge, however, when the student pursues multi-digit addition and subtraction which require regrouping, multiplication with a multi-digit multiplier, and long division. In this situation, difficulties with precision and graphomotor output would be considered secondary to difficulties with complexity, as they are only observed on tasks which require the management of multiple demands.

Next, consider the student who demonstrates imprecision in multi-digit calculations. Careful inspection of work samples reveals that the imprecision is related to difficulties aligning digits and difficulties producing legible numerals. Here, the imprecision would be considered secondary to difficulties with visuospatial skills and graphomotor output.

In other situations, features may exacerbate one other. For example, consider the student who has difficulty with language demands and works at a slow pace. The student also demonstrates low self-efficacy in arithmetic. The student’s difficulties understanding instructions and producing work quickly are likely reasons for
discouragement in arithmetic. This discouragement can, in turn, cause the student to be less engaged and motivated to attend to instruction and to complete work, setting into motion a downward, negative cycle.

As these examples illustrate, clinicians must consider behaviors in relation to one another and to the tasks themselves, with the goal of determining whether the features observed are primary or secondary. This analysis will help to clarify the nature of the student’s difficulties and help in the design of appropriate instructional recommendations.

**INCORPORATING QUALITATIVE FEATURES INTO ASSESSMENT PRACTICE: A BLUEPRINT FOR THE PRACTITIONER**

The keys to incorporating a consideration of the qualitative features of performance into assessment practice are straightforward, including 1) familiarity of the examiner with qualitative aspects of performance and their manifestations in arithmetic, 2) careful and systematic observation of the student while working on tasks, and 3) flexible questioning designed to reveal the student’s understandings and thinking processes. These practices are not exclusive to one type of arithmetic assessment. They can be used by clinicians who administer standardized achievement tests and by teachers during their work with students in the classroom setting. Thus, a consideration of qualitative features of performance can be incorporated in different assessment contexts.

A critical component of incorporating qualitative features into assessment is observation. Currently, many clinicians administer standardized tests of arithmetic by reading the directions and letting the student complete the test. These clinicians often use the time the student takes to complete the test to begin scoring other subtests from the battery. Although understandable, this practice is unfortunate, as the clinician can learn a great deal by carefully observing how the student approaches, pursues, and solves each arithmetic task. More specifically, the interviewer should attend to both the accuracy of response and to the behaviors that give rise to that response, taking detailed notes. After the student has completed the arithmetic tasks (either standardized or clinician/educator-created), the interviewer may probe with additional questions for the purpose of gaining insight into a student’s skills, understandings, strategies, and cognitive and affective behaviors. For example, the clinician / teacher might ask the student to revisit particular items and to explain solution strategy. The interviewer should select examples which the student answered correctly as well as those answered incorrectly in order to fully gauge thinking processes. Questioning the student about successes and errors offers two advantages. First, it reveals strategies that are successful and those that are not. Second, it may prevent the student from becoming self-conscious about whether the performance was good or bad.

The student’s responses and qualitative features observed should be carefully recorded. The clinician / teacher should keep all data produced during the session, including the recording forms from standardized achievement batteries, the student’s answer booklet, and notes taken during the testing session. The clinician / teacher should then review and analyze the information subsequent to the testing session in order to formulate findings, make recommendations, and write a report.

In analyzing a student’s work, it is important that the interviewer distinguish errors arising from difficulties with qualitative features from errors generated by a lack of knowledge or faulty knowledge. Studies of arithmetic errors distinguish systematic errors from “slips” (VanLehn, 1990). Systematic refers to errors generated from consistent application of a faulty method or algorithm. Slips, on the other hand, are unsystematic errors (e.g., 8 – 2 = 5), which can occur frequently but do not reflect a lack of conceptual understanding. Describing the difference between systematic errors and slips, VanLehn (1990) writes: “Some errors are so systematic and stable that one must assume that they are caused by flaws in the students’ procedural knowledge rather than careless mistakes or random reactions to lack of knowledge” (p. 15). For example, consider a student’s performance on the following addition tasks (adapted from Ashlock, 1994):

\[
\begin{array}{ccc}
35 & +24 & 618 \\
+92 & +582 & +782 \\
19 & 719 & 1112
\end{array}
\]

Here, the systematic error involves a reversal of the procedure used in the traditional addition algorithm, with addition and “carrying” performed left to right. Although these solutions are incorrect, these errors do not constitute imprecision or difficulties managing complexity per se. Rather, they reflect a lack of procedural / conceptual knowledge.

Special care should be taken when discerning a student’s competence with language and visuospatial skills. As noted previously, these features are not always easily assessed in the context of an arithmetic test. If the examiner suspects that a student may have difficulties
with these features, the examiner might consult with the student’s teacher, observe the student in the classroom, review work that the student has produced, and review educational and cognitive testing that can shed light on a student’s competencies in these domains.

Another critical component of the assessment process is looking for patterns of qualitative features that emerge across individual tasks. As the interviewer observes and questions, initial hypotheses about difficulties observed can be formulated. These preliminary observations will be informed by performance on subsequent items. On occasion, the interviewer may want to ask the student to complete additional items in order to confirm or dismiss initial clinical impressions. Using careful observation and questioning, the interviewer can identify whether a student demonstrates difficulty with one or more qualitative features. This information can then be incorporated into a written report and used to guide educational planning.

QUALITATIVE FEATURES OF PERFORMANCE: RESEARCH FINDINGS

As noted, the qualitative features delineated above were first described by Marolda and Davidson (1994; 2000) and have been used in a diagnostic context at Children’s Hospital Boston. Boiselle (2007) recently investigated the relevance and utility of these features for the assessment of arithmetic competencies. Specific research questions included: 1) What is the prevalence of these features amongst a population of school-aged children who have been referred for evaluation of academic difficulties? 2) What is the association of these features with traditional, norm-referenced measures of neuropsychological function? Although a comprehensive discussion of this research is beyond the scope of this paper, a summary of methods and findings is provided here to further illustrate the utility of the seven features as diagnostic tools.

METHOD

Participants in the study (n = 52) were a subset of a larger sample of children (n = 188) who had been referred to a hospital outpatient program in the Department of Behavioral Neurology at Children’s Hospital, Boston, for the evaluation of learning issues.1 Participants were between 7 and 11 years, 11 months of age. The principal criterion for inclusion was referral (by parent, pediatrician, school, or some combination therein) for evaluation of suspected learning problem(s). Reasons for referral were heterogeneous. Although some students were referred for specific concerns about achievement in mathematics, reading, and/or writing, others were referred for more general questions including attention difficulties and organizational issues. All participants demonstrated average IQ (Full Scale IQ > 85 on the Wechsler Intelligence Scale for Children III (WISC III)) and had no known emotional or behavioral disorders or neurological syndromes.

As part of a comprehensive learning disabilities evaluation, participants underwent a battery of cognitive, neurological, and academic tests. Arithmetic competencies were assessed via the Numerical Operations subtest of the WIAT I. Qualitative features of performance were assessed via administration of the Mathematics Diagnostic and Prescriptive Inventory (MDPI) (Marolda & Davidson, 1978, 1983). The MDPI is a mathematics evaluation instrument designed to reveal salient qualitative features of performance and to assess achievement, ability, and learning approaches in mathematics. Data regarding whether a participant demonstrated difficulty on each qualitative feature of math performance were gathered via retrospective review of the clinical notes from the MDPI data forms. Each feature was coded as a dichotomous variable, with 1 representing that the participant demonstrated difficulty with the qualitative feature of math performance, and 0 representing that the participant did not demonstrate difficulty with that qualitative feature.

RESEARCH FINDINGS

Prevalence

Analyses concerning prevalence revealed that the vast majority of the fifty two participants (92%, n = 48) were coded as demonstrating difficulty with one or more qualitative features of performance. This suggests that these features are relatively common amongst school-age children who are experiencing difficulties in school. Further investigation was undertaken to determine whether these features were more common amongst children who met criteria for a diagnosis of an arithmetic disability (AD) than in those who were experiencing learning difficulties in school but did not meet criteria for AD. (AD was defined as a performance at or below the 15% on the Numerical Operations subtest on the WIAT I.) Findings (Table 1) reveal that difficulties with precision, pace, management of complexity, and self-

1This larger sample took part in a multidisciplinary NIH-funded study of children with learning difficulties led by Dr. Deborah Waber, a neuropsychologist at Children’s Hospital, Boston.
efficacy were more common amongst participants with AD (n = 19) than in the non-AD population (n = 33).

Also interesting was the finding that non-AD participants (n = 33) demonstrated difficulties with all qualitative features, albeit at lower percentages. This suggests that difficulties with qualitative features of performance may be observed regardless of AD status.

Results related to prevalence are interesting for three reasons. First, the frequency with which these features are observed within the referred population suggests their importance in clinical assessment of arithmetic. Second, results indicate that difficulties with pace, precision, management of complexity, and self-esteem are endemic within the AD population. This suggests that these features can contribute to poor performance in arithmetic. Therefore a consideration of these features may illuminate the source of a student’s difficulty and provide useful information regarding recommendations. Finally, even participants who performed within the average to above-average range on the WIAT I Numerical Operations subtest demonstrated difficulties with various qualitative features of performance. For these students, a consideration of the qualitative features of performance may be critical, as it can reveal why the student may be struggling in the classroom setting despite an average performance on a standardized or teacher-created arithmetic measures.

 Associations with neuropsychological measures of cognitive function

The associations between traditional, norm-referenced cognitive measures and each feature were investigated for the purpose of determining their diagnostic utility as measures of underlying neuropsychological processes. As part of previous investigation with the larger study population (n = 188), Waber and colleagues (Waber, Forbes, Wolff, & Weiler, 2004) identified five distinct latent factors of neuropsychological function. These included:

- Written Language Factor (a composite of performance on various decoding, reading comprehension, and spelling measures);
- Oral Language Factor (a composite of performance on expressive and receptive language abilities and vocabulary knowledge);
- Visuospatial Processing Factor (a composite of measures assessing an individual’s facility with perceptual reasoning, visual-motor integration skills, and visual-perceptual abilities);
- Motor Speed Factor (a composite reflecting a child’s fine motor dexterity); and
- Processing Speed Factor (a composite reflecting the efficiency with which an individual processes and produces information).

The association between each qualitative feature of math performance and these five neuropsychological factors was explored via point biserial correlations.

The association between qualitative features and performance on selected subtests of naming speed, specifically the Rapid Alternating Naming test (RAN) and the Rapid Alternating Stimulus test (RAS), was also explored. The RAN measures the individual’s ability to rapidly identify and name individual letters (RAN Letter) or single-digit numerals (RAN Number) presented in a ten by five array. The RAS requires that the individual quickly name either alternating numbers and letters (Two-set) or alternating numbers, letters, and colors (Three-set). Research indicates that prolonged naming speed on the RAN reflects insufficient speed of processing and/or asynchrony within and among a complex ensemble of attentional, perceptual, conceptual, phonological, semantic, and motoric subprocesses (Wolf, 2007) and is predictive of a broad range of learning problems (Waber, Wolff, Forbes, & Weiler, 2000; Waber et al., 2004). Research concerning the RAS indicates that this measure taps working memory, inhibition, set-switching, selective attention, and monitoring skills.

<p>| Table 1. Prevalence of Difficulties with Qualitative Features in AD and non-AD Participants |
|-----------------------------------------------|-----------------------------------------------|</p>
<table>
<thead>
<tr>
<th><strong>Participants with AD (n = 19)</strong></th>
<th><strong>Participants without AD (n = 33)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>58% (n = 11)</td>
</tr>
<tr>
<td>Pace</td>
<td>42% (n = 8)</td>
</tr>
<tr>
<td>Graphomotor Output</td>
<td>32% (n = 6)</td>
</tr>
<tr>
<td>Visuospatial Skills</td>
<td>42% (n = 8)</td>
</tr>
<tr>
<td>Language Demands</td>
<td>26% (n = 5)</td>
</tr>
<tr>
<td>Management of Complexity</td>
<td>79% (n = 15)</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>58% (n = 11)</td>
</tr>
</tbody>
</table>
(Amtmann, Abbott & Berninger, in press). Associations between each feature and naming speed measures were also explored via point bi-serial correlations.

Analyses revealed that difficulties with each of the qualitative features were associated with poor performance on one or more measures of neuropsychological function (Table 2).

More specifically, the qualitative feature of precision was modestly to moderately correlated with all measures of naming speed, indicating that prolonged naming speed was associated with imprecision. Pace was negatively correlated with the Processing Speed Factor, suggesting that the difficulties with pace are strongly associated with low scores on traditional, norm-referenced measures that assess speed of processing. Pace was also modestly to moderately correlated with all measures of naming speed. This indicates that slow scores on measures of naming speed are associated with slow pace.

The qualitative feature of graphomotor output was moderately associated with the Motor Speed Factor, indicating that difficulties with this feature are associated with difficulties with fine motor skills. The visuospatial skills in math feature was strongly negatively associated with the Visuospatial Factor. This indicates that difficulties with visuospatial skills are correlated with low scores on traditional, norm-referenced measures of visuospatial / perceptual abilities. The qualitative feature of language demands in math was associated with the Oral Language Factor, indicating that difficulties with language demands associated with low scores on the Oral Language Factor.

The management of complexity feature in math was correlated with the Processing Speed Factor, the Written Language Factor, and the Visuospatial Factor, indicating that difficulty with this feature is associated with low scores on all three of these Factors. Management of complexity was also correlated with all measures of naming speed, indicating that difficulties with complex-

<table>
<thead>
<tr>
<th>Difficulties Noted In:</th>
<th>Neuropsychological Factors</th>
<th>Naming Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Written Language</td>
<td>Oral Language</td>
</tr>
<tr>
<td>Precision</td>
<td>-0.01</td>
<td>0.12</td>
</tr>
<tr>
<td>Pace</td>
<td>-0.28*</td>
<td>-0.2</td>
</tr>
<tr>
<td>Graphomotor Output</td>
<td>-0.02</td>
<td>-0.03</td>
</tr>
<tr>
<td>Visuospatial Skills</td>
<td>-0.13</td>
<td>-0.1</td>
</tr>
<tr>
<td>Language Demands In Math</td>
<td>-0.25</td>
<td>-0.33*</td>
</tr>
<tr>
<td>Management of Complexity in Math</td>
<td>-0.35*</td>
<td>-0.05</td>
</tr>
<tr>
<td>Self-efficacy in Math</td>
<td>-0.19</td>
<td>-0.26</td>
</tr>
</tbody>
</table>

* significant at $p < .05$; ** significant at $p < .01$; *** significant at the $p < .001$

Table 2.
Point Biserial Correlations Among Difficulties with Qualitative Features of Performance and Measures of Cognitive Function

The reader should note that for neuropsychological factors of Written Language Factor, Oral Language Factor, Visuospatial Factor, Motor Speed Factor, and Processing Speed Factor, high scores denote a strong performance, and low scores denote a poor performance. Thus a negative correlation between a feature and a Factor indicates that difficulties with a feature were associated with low scores on a neuropsychological Factor. However, for naming speed variables (e.g., Rapid Automatized Naming and Rapid Alternating Stimulus Naming), high scores denote poor performance. Here, a positive correlation indicates that difficulties with a feature were associated with prolonged naming speed. This information will help the reader to interpret the results in Table 2.
ity are associated with slow performance on measures of naming speed. Finally, the qualitative feature of self-efficacy in math was modestly correlated with the Processing Speed Factor, indicating that low self-efficacy is associated with low scores on measures of processing speed. Self-efficacy was also modestly to moderately associated with all measures of naming speed.

A comprehensive discussion of these findings is provided in Boiselle (2007). For the purposes of this paper, the associations between qualitative features and traditional norm-referenced measures of cognitive function are important for two related reasons. First, they support the contention that these features represent a behavioral manifestation of underlying cognitive processes. Thus, any observed difficulty with a given feature may reflect underlying cognitive difficulties rather than carelessness, indifference, or lack of investment, as is sometimes supposed. Second, the fact that each of the seven qualitative features is correlated with norm-referenced neuropsychological measures supports their utility as descriptors of what Werner (1937) referred to as process. Although more research is needed to firmly establish the validity of these features, these initial results suggest that they can be used by educational professionals in their assessment of arithmetic.

**DISCUSSION OF FINDINGS AND FUTURE DIRECTIONS**

Considered together, the findings from this research indicate that these seven qualitative features have relevance and utility as diagnostic measures. These findings highlight the prevalence of these features within a population of children who are experiencing difficulty in school. They also indicate that difficulties with certain features may be present despite an average or above-average performance on standardized measures of arithmetic. Thus, a consideration of qualitative features can provide important information for understanding a student’s competencies in arithmetic that might not otherwise be revealed. Finally, these findings indicate that each of these features is correlated with traditional measures of cognitive function and therefore has diagnostic value in an assessment context.

This research sets the stage for future investigations concerning qualitative features of performance. There is a need for more research investigating the prevalence and impact of qualitative features of performance amongst a larger population of elementary school students with and without learning difficulties in mathematics. Adopting a longitudinal perspective, in which qualitative features are assessed at different time points over the elementary school years might prove illuminating. Future research endeavors should also move beyond coding of qualitative features via a dichotomous dimension to include a Likert scale where qualitative features are coded according to degrees of severity.

Another critical component of future research in this area lies in exploring the relation between different constellations of qualitative features of performance and the educational context in which they occur. Children do not learn in a vacuum, nor are the programs utilized and instructional approaches offered uniform. It is therefore possible that certain features may not have a significant impact on learning and performance in a given educational context, while others will. Future research must examine how factors such as classroom context, mathematics program, size of class, and teacher experience mediate the impact of difficulties with one or more qualitative features of performance.

A third direction for future research involves examining the impact of qualitative features of performance on different subtopics of mathematics. The investigation focused solely on arithmetic. Future endeavors must explore how qualitative features affect learning and performance in algebraic thinking, geometry, measurement, data analysis and probability, and various types of problem solving.

**INCORPORATING QUALITATIVE FEATURES OF PERFORMANCE IN ASSESSMENT OF ARITHMETIC: EXAMPLES AND IMPLICATIONS**

**Case Studies**

Having delineated the seven qualitative features and described research supporting their utility in assessment, I now discuss two case studies that illustrate how a consideration of qualitative features can enhance assessment of arithmetic competencies.

**Case study 1**

Student A is a beginning fourth-grade male. His parents report that although he receives Bs and the occasional C in mathematics, he has to spend more time than his peers to complete assigned work and states that he does not like math. His teacher notes that although he works diligently, he seems to be frustrated in math class and sometimes needs extra time to complete assignments.

On the WIAT II Numerical Operations subtest, Student A’s performance yields a standard score of 101 (grade equivalent 4.2, 51¾%ile). A review of the answer sheet reveals that Student A was able to generate solutions to one-digit addition, subtraction, multiplication,
and division facts. He was also able to arrive at correct solutions to a number of the multi-digit addition and subtraction examples, including 41 + 14, 68 - 43, 37 + 54, 120 - 15, and 80 - 56. Thus, it would seem that Student A is performing well within the average range for students his age and should not experience difficulty in math class.

But how does this performance fit with the observations made by Student A’s parents and teacher? Consideration of the qualitative features observed while Student A worked on this test provides a more complete picture. First, Student A produced written output in an awkward and effortful manner. Although he was able to arrive at correct solutions for various test items, he exerted extra effort to meet the graphomotor demands. As a result, he worked at a more deliberate pace than is typical of children his age. Second, Student A was not able to arrive at accurate solutions for some of the more complicated multi-digit addition and subtraction items including 698 + 426, 705 - 489, and $5.31 - 2.47$. On these items, he had difficulty recalling solutions to the one digit facts therein (e.g., 8 + 6 = 13; 15 - 9 = 7; 12 - 8 = 7), despite having demonstrated proficiency with facts in isolation. Moreover, when pursuing 705 – 489, Student A needed multiple trials and extended time in order to successfully regroup across the zero. It seemed that the effort involved in coordinating the various calculation, graphomotor, and sequencing demands was challenging for him. Student A also had difficulty maintaining the subtraction operation throughout $5.31 - 2.47$. On this item, he added the dollars after the correct regrouping for subtraction, arriving at a final solution of $6.84.

Taken as a whole, Student A’s performance suggests that he struggles with graphomotor output and with the management of complexity. These features may be contributing to the difficulties noted by his parents and teacher vis-à-vis completing homework. Student A’s difficulty managing multiple demands suggests that he may be vulnerable in the future, particularly when he learns skills for multiplication with a multi-digit multiplier, long division, and fraction calculations. Thus, although Student A does not qualify for special education services based on his standard score on the WIAT II, he would benefit from close monitoring by his teachers as well as accommodations and strategy-instruction to ameliorate the impact of his difficulties with graphomotor output and complexity.

Case study 2

Student B is a mid-year third-grade female. Her teacher reports that she demonstrates a good understanding of math concepts but performs poorly on tests and quizzes. Student B shares that she enjoys math but is not good at it. On the WJIII Calculation subtest, Student B receives a standard score of 81 (grade equivalent, 2.1; 10th%-ile). On the Math Fluency subtest, Student B receives a standard score of 84 (grade equivalent 2.3, 15th%-ile). Based on this performance, Student B would likely be eligible for special education services in mathematics and would qualify for a diagnosis of an arithmetic disorder.

A consideration of qualitative features, however, reveals a different picture. Student B’s work is marked by imprecision in facts, calculations, and operation signs. More specifically, many of her solutions to one-digit facts were incorrect (e.g., $7 + 5 = 13; 7 - 2 = 4; 8 + 2 = 3$). This imprecision with facts hindered her performance in multi-digit calculations as well ($476 + 61 = 2,611 = 3,138$). Student B also exhibited imprecision with operation signs. On both the Calculation and Math Fluency subtests, she would add when the operation called for subtraction ($5-2=7; 17 - 9 = 26$) and subtract when the operation called for addition ($7 + 1 = 6; 2 + 1 = 1$). She seemed to have particular difficulty switching between operations. For example, when solving one-digit facts on the Math Fluency subtest, Student B was confronted with two addition items, followed by one subtraction, one addition, one subtraction, one addition and two subtractions. On these items, however, she added each time. Student B actually completed forty items on the Math Fluency test. Due to her inattention to operation signs, however, she only received credit for thirty items.

Close inspection of Student B’s work reveals grade-appropriate conceptual understandings and procedural abilities. She is able to pursue multi-digit addition and subtraction calculations (in circumstances with and without regrouping) and can explain her solution strategies. She is also able to use her conceptual understanding of multiplication as repeated addition to attempt items that involve multiplying a two digit number by a single digit (e.g., $13 \times 7, 14 \times 6$). Thus it appears that Student B does not have conceptual or procedural deficits per se. Rather, her skills and understandings in arithmetic are undermined by imprecision. It would therefore be a mistake to consider Student B as having a learning disability in arithmetic. She does not need remedial instruction in arithmetic skills and concepts. Rather, she would benefit from a short-term, directed effort directed at 1) developing effective strategies with which to generate accurate solutions to one digit facts and 2) learning monitoring strategies to address her difficulties with precision. She would also
benefit from accommodations, such as scoring work twice, once for correctness of approach and once for accuracy of solution.

**Advantages of Incorporating Qualitative Features of Performance into Assessment of Math Learning Difficulties**

As these case studies illustrate, a consideration of qualitative features may enhance assessment of learning difficulties in arithmetic in five important ways. First, in circumstances where a standardized test of arithmetic is used, qualitative features of performance may more accurately reflect a student’s level of success/vulnerability in the classroom than standard scores, per se. By describing the various qualitative features of performance that accompany and give rise to solutions, the clinician or teacher may be better able to describe the nature and causes of difficulties in arithmetic.

Second, qualitative features of performance can help clinicians and teachers to offer educational recommendations directly tailored to specific learning needs. To illustrate, consider two children, both of whom perform at the 15th %-ile on a standardized arithmetic measure. A consideration of qualitative features of performance reveals that the first child’s low scores are related to features of imprecision (e.g., \(28 + 16 = 45\); \(32 + 5 = 38\)). In contrast, the second child’s performance is not marked by specific qualitative features of performance. Rather, errors reveal a lack of underlying place value concepts (e.g., \(28 + 16 = 314\); \(32 + 5 = 10\)). No doubt each student would benefit from tailored instruction, but it would be a mistake to assume that they have similar mathematical profiles and would be well-served by the same types of interventions. Yet based on current assessment practice, both children would be coded as AD and might receive the same educational services. A consideration of “process,” however, would likely yield a different result, both in terms of describing the distinct math learning vulnerabilities of each student and in delineating appropriate educational services.

Third, a consideration of qualitative features may enable more effective prediction of the child’s future learning and point to ways in which difficulties can be prevented or circumvented. For example, a child who demonstrates difficulties with the management of complexity may be particularly vulnerable when it comes to more demanding arithmetic topics, such as long division and fraction operations. For this child, previewing opportunities, mnemonic strategies, and work with concrete manipulatives may ensure mastery of the material and help to avert frustration and failure.

Fourth, qualitative features of performance might be usefully employed by teachers when they write report cards documenting a student’s learning and performance. Teachers might even consider incorporating these features in their report-card templates. Such a practice might be helpful for two reasons. First, it could document whether features remain constant or diminish over time. Second, and perhaps more importantly, it would allow teachers to pass on information about a student from year to year, enabling subsequent teachers a “window” into the student’s potential vulnerabilities in arithmetic and fostering continuity in the student’s mathematical instruction.

Finally, a consideration of qualitative features of performance might occasion a productive partnership between teachers and clinicians regarding assessment. Too often, these professionals work in relative isolation from one another, with the teacher conducting assessment via classroom measures of performance and the clinician assessing a student via formal standardized assessments. The qualitative features of performance might provide a common framework and vocabulary through which teachers and clinicians can work in partnership to enhance assessment practice.

**References**


Applying neuropsychological analyses to standardized measures of arithmetic achievement. Dissertation Abstracts International Section A: Humanities and Social Sciences, 68 (3-A). (UMI No. AAI3258281)


Analysis of the Effect of Specific Vocabulary Instruction on High School Chemistry Students’ Knowledge and Understanding

Peggy LaBrosse
Hollis Brookline High School

ABSTRACT
This study analyzed the effects of specific vocabulary instruction on high school chemistry students’ knowledge and understanding. The Frayer Model, a research-based teaching strategy for specific vocabulary instruction, was used as a treatment.

Vocabulary knowledge was examined using multiple-choice pre- and posttests administered to all student participants. Student understanding of chemistry content was examined using chemistry content understanding pre- and posttests administered to all student participants.

A slight gain in vocabulary knowledge scores was found in the treatment group compared to the control group, but it was not statistically significant. Chemistry content understanding scores were also compared between both the treatment and control groups. The treatment group did not do as well on chemistry content understanding tests as the control group. However, analysis of the confounding variables, which included the amount of contact time, student attendance, and duration of the research, lead the researcher to conclude that the use of the Frayer Model for specific terms in content across the curriculum is worth further study.

INTRODUCTION
The purpose of this research was to explore the effect of specific vocabulary instruction on high school chemistry students’ knowledge of academic language and their subsequent understanding of chemistry. Academic language includes vocabulary that is used beyond social conversations. Such vocabulary is needed to communicate effectively and to comprehend various texts in different content area classes (Zwiers, 2004). While high school chemistry content has its specific vocabulary, rules, concepts, models, and theories, students come into the chemistry classroom with prior knowledge, including their personal vocabulary, rules, and concepts about the meanings that are different from the meanings used in the chemistry curriculum. This creates an obstacle for successful learning of chemistry.

For further discussion, word is used to name the set of words commonly used by students in their everyday language; term is used to denote words that are specific to chemistry content. Terms can be viewed as a subset of the set of words; words become terms when a specific content meaning is applied. In some cases, student prior knowledge of the meaning for particular words is consistent with the formal meanings in the chemistry content. However, in many cases, student prior knowledge of the meaning of the words differs from that suggested by the curriculum. For example, the word period has at least two common meanings: a mark at the end of a sentence to indicate a statement is complete, and a time span as in a class period. In chemistry, period becomes a specific term indicating a row on the Periodic Table of the elements that helps to predict specific characteristics of those elements. Student prior knowledge about what the word period means differs from the meaning of the term period in the chemistry context. This prior knowledge interferes with the acquisition of the new use of the term (Meyerson, Ford, Jones, & Ward, 1991; O’Farrell, 2000; Wellington & Osborne, 2001).

RATIONALE FOR THE RESEARCH
Researchers (Cassels & Johnstone, 1983; Gabel, 1999; Johnstone, 1991) have been studying the difficulty students have in learning science, and particularly chemistry. Gabel (1999) suggests that, “while research into misconceptions (also known as alternative conceptions) and problem-solving have dominated the field for the past 25 years, we are no closer to a solution that would improve the teaching and learning of chemistry” (p. 549). She relates the difficulty in learning chemistry to the use of language. She refers to student difficulty both with words that have more than one meaning in English and with words with one meaning in chemistry and another meaning in everyday language. Two examples to further demonstrate the latter concern are: the everyday word strong (e.g., coffee is strong) has its counterpart in chemistry as concentrated and the everyday word melts (e.g., candy melts in a mouth) has its counterpart in chemistry as dissolves. In an international study, Dori, Gabel, Bunce, Barnea and Hamieri, (1996) researched the issues related to language difficulty. They asked teachers to differentiate pure substances from other materials. The choices provided were: concrete, Pepsi, milk, sugar, and all of the above. In the U.S. sample,
only 48% of the teachers surveyed gave the correct answer that sugar was the pure substance. About 40% of the sample indicated that all of the materials were pure substances. In Israel, the test was translated into Hebrew. There was no equivalent word for substance, so substance was translated as pure matter, and about 83% of them chose the correct answer. Other languages also have words with scientific meanings that differ from the meanings in everyday usage. This creates a problem which is not necessarily related to the complexity of chemistry, and it can be solved by carefully defining terms, coming to a consensual meaning, and selecting appropriate vocabulary (Dori et al., 1996).

Almost all teaching and learning takes place using the medium of language. Learning science, in many ways, can be viewed as learning a new language. Terms such as energy, work, and power have precise meanings in science and sometimes exact definitions, but very different meanings in everyday life. Teaching science involves dealing with familiar words like energy and giving them new meanings in new contexts. Equally, many of the “naming” words of our lives — such as element, conductor, cell, field, circuit, and compound — have been commandeered by science. This is made even more complex because many of these words are used in science as metaphors — for example, a field in science is not a physical field used for planting corn — to describe abstract ideas. Teaching science also involves introducing new words in familiar contexts (e.g., tibia, fibula), as well as in unfamiliar contexts (e.g., molarity, stoichiometry, allele, and enzyme).

Ausubel (1968) and Novak (1993) suggested that meaningful learning occurs only if new information is linked to the existing relevant concepts. Teachers need to help pupils make connections between new and existing vocabulary and to facilitate students’ understanding of new meanings of words necessary to understand chemistry.

Most chemistry teachers agree that science, particularly chemistry, is perceived by students as a difficult subject. The reasons why chemistry seems so hard to students are still not clear, although many suggestions have been offered, e.g., abstract nature (three levels: macroscopic, submicroscopic, and symbolic), seeming irrelevance to everyday life, language difficulty, and insufficient calculation skills to support laboratory work. There is a growing need to take positive steps to improve science learning conditions (Byrne, Johnstone, & Pope, 1994).

If students find difficulties linking ideas together, as suggested by the results of the research into logical con-nectives (Byrne et al., 1994), then mental processing and subsequent absorption during a lesson is likely to be diminished. Cognitive Load Theory (Paas, Renkl, & Sweller, 2003) puts forward the concept that students have difficulty learning when their working memory is overloaded. Unfamiliar academic language along with the abstract subject matter increases the load.

Zwiers (2004) states that not knowing even a few of the terms and phrases of academic language can sabotage a student’s comprehension of the subject. In Language and Literacy in Science Education, Wellington and Osborne (2001) note that one of the major difficulties in learning science is learning the language of science. They base their discourse regarding students learning the language of science on three premises:

1. Learning the language of science is a major part (if not the major part) of science education. Every science lesson is a language lesson.
2. Language is a major barrier (if not the major barrier) to most pupils in learning science.
3. There are many practical strategies which can help to overcome these barriers. (p. 2)

This research focuses on the second premise, i.e., that language (in this study, academic language) is a major barrier to most pupils in learning science.

Knowing the terms—the academic language—does not mean that students can attach meanings to them or that chemistry concepts are understood or the acquisition of knowledge has taken place (Cassels & Johnstone, 1983; Childs & O’Farrell, 2003; Johnstone, 1991). A student can formally recite a definition for a term without actually having understood the meaning of the term and its connection to other terms and to related concepts. Tovani (2000) contends that this can be the result of “fake reading” when the student is just formally trying to find answers to questions. Tovani further asserts that when terms are not meaningful to students it is still possible for them to answer questions including the terms and appear to understand based on their ability to craft appropriate answers from the available text. Meyerson et al. (1991) found that when students are probed for further understanding of terms in science they will often revert to the meanings from prior experiences. An example of this is the term mass which is used in science to denote the weight of an object. During interviews with students, Meyerson et al. found that students related mass to the experience of attending church services on Sundays rather than to its scientific meaning. Meyerson et al. concluded that the students did not actually acquire new knowledge about the term.
mass, which apparently has dual meaning. The researchers asserted that terms with dual meanings actually interfere with knowledge acquisition; the new meaning is often rejected since the primary meaning does not necessarily fit to the new information in the new context.

According to Hayes-Jacobs (1997) and Wellington and Osborne (2001), all teachers are teachers of language, and part of the responsibility of teachers is to teach the language of the content to ensure meaningful learning has occurred. In her seminar presentation for Curriculum Mapping, Hayes-Jacobs (2005) expresses how well teaching language in context is modeled in the foreign language classroom, where students see, say, write, recite, and use words in the context of the language. She points out that teachers of subjects such as mathematics and science do not emphasize using context when teaching subject matter language, whereas teachers of foreign languages do.

Mid-Continent Regional Education Library (MCREL) and the Association for Supervision and Curriculum Development (ASCD) have taken a step towards providing teachers with research-based tools for literacy instruction across content areas (Billmeyer & Barton, 2002). The Frayer Model, one of these tools, is a graphic organizer which students use to create meaningful definitions for terms in context (Billmeyer & Barton, 2002; Frayer, Frederick, & Klausmeier, 1969).

The Frayer Model is a graphic organizer consisting of a diagram with five sections, arranged as a center section surrounded by four quadrants. The center section contains the term to be defined. Working in small groups, students fill in the four quadrants — listing essential and non-essential attributes, examples from their own lives, and non-examples — to define the term. Typically, textbooks and other resources are used instead of teacher-led vocabulary instruction (where the teacher writes the term on the board, defines it, asks if there are any questions, and there is little or no student input or interaction). The model can be used as part of classroom instruction with little or no additional time or specialized training (Billmeyer & Barton, 2002). The major feature of the instructional approach using the Frayer Model is that the students are encouraged to develop meanings of the chemistry terms on their own instead of being told the meanings by the teacher. The students work in groups to share and negotiate their meanings.

Monroe and Pendergrass (1997) found the Frayer Model to be effective for teaching mathematics vocabulary. Among elementary school students in Shawnee Mission Elementary School, 56% of students showed an increase in vocabulary scores of 50% or more by using the Frayer Model (Enge, 2005). At a Northern New England Co-Mentorship Network (NNECN) Conference in Boothbay Harbor, Maine M. Baker, B. Fortier, and M. Mailhot (personal communication, June 29, 2006) described the use of the Frayer Model as part of a literacy initiative funded by the Maine Board of Education and earmarked for No Child Left Behind.

In response to the No Child Left Behind Act (NCLB), many science teachers are now being asked to teach reading and writing as English teachers do, and to make science accessible to all students so that they can meet national standards in science (Padilla, 2005). While discussing how English-Language Learners (ELL) study science, Padilla asserted that the National Science Education standards (American Association for the Advancement of Science [AAAS], 1993) call for “science … for all students.” (p. xi) He claims that to meet this requirement, teachers must become more proficient at teaching science literacy. If literacy instruction is appropriate for ELLs to understand the language of science, would it not also be appropriate for native English language-speaking students when they are learning chemistry? This research used the Frayer Model as a treatment to enhance the instruction of academic language in an attempt to help students manage the cognitive load.

**RESEARCH QUESTIONS**

Based on the above, it would seem that studying students’ knowledge of academic language is important to explore the interrelationship between students’ knowledge of academic language and understanding of chemistry. This research examined the following questions:

1. What effect does using the Frayer Model have on high school students’ knowledge of academic language used in chemistry?
2. What effect does using the Frayer Model have on high school students’ understanding of chemistry?
3. What is the relationship between high school chemistry students’ knowledge of academic language and understanding of chemistry as a result of the Frayer Model application?

**STUDY DISCUSSION**

Since the research inquiry is oriented toward the analysis of effects, it seemed more appropriate to use
quantitative methods to answer the questions. To study these questions, 95 eleventh-grade students in a New England high school were divided into two groups. One group — the control group (CG) — received traditional vocabulary instruction. This instruction included the faculty participant writing the term on the board, providing a definition, asking for student examples, providing additional examples, and soliciting student questions. The students were expected to copy the definitions and examples from the board into their notes. The other group — the treatment group (TG) — used the Frayer Model for vocabulary instruction as a part of their classroom experience. All other chemistry teaching methods were similar for both groups, based on the prescribed curriculum. Before and after instruction of the specific terminology, all students received the same test questions to assess their actual academic language knowledge and chemistry content understanding. The data were analyzed to answer whether there was any effect due to the Frayer Model application.

RESEARCH DESIGN

SAMPLE

The school

This research was conducted on the campus of one of southeastern New Hampshire’s high schools. The school is a four-year, academically oriented high school accredited by the New Hampshire Department of Education and the New England Association of Schools and Colleges (NEASC) with an enrollment of 862, a current senior class of 183, and a faculty of 76. Eight-six percent of the 2005 graduating class are pursuing further education. Graduates of the class of 2005 attend 78 different schools and colleges — 77% at four-year colleges, 9% at two-year colleges, 5% employed, 1% military, and 8% other programs. Flexible curriculum planning for all is achieved by allowing students to select courses from the entire program. The degree of difficulty ranges from AP/weighted honors courses to special needs/remedial courses. Students are bussed to other area high schools under an area program for vocational/technical courses.

Student participants

Four intact Chemistry class sections with a total enrollment of 95 students were involved during one semester of a regular academic year. Two class sections participated in the Frayer Model activity experimental treatment for vocabulary instruction and two class sections received traditional vocabulary instruction. Student participants ranged in age from 15 to 17 years of age. Class sections were 50% female and 50% male.

Students enroll in Chemistry based on teacher, guidance counselor, and parent advisement, two assessments, their mathematics grades, and career plans. Enrollment is by choice — students are not placed using an objective test; free will and a comprehensive process are employed. Students and parents take the selection process seriously and less than 10% of the students change levels. This suggests that students are placed in the level that is most appropriate for their ability. When a level change is called for it is usually completed within the first 12 school days; less than 1% of the students change levels later in the year. These are often the result of unique circumstances, such as health or family issues, and require administrative oversight.

The homogeneity of student population provided by this particular high school reduces variables normally introduced by factors such as students receiving free and reduced lunches, second language students, widely varying socioeconomic status, and other such differences. Students in these categories comprise less than 1% of the high school and did not take the Chemistry classes studied.

The class sections participating in this research met four times a week (two 51-minute meetings and two 81-minute meetings), for a total of four and a half hours of instruction per week.

FACULTY PARTICIPANT

Conscious about the limitation, I played two roles, the researcher and the teacher. I taught all four sections of Chemistry participated in the study. According to the state standards I am Highly Qualified to teach high school chemistry. I have been teaching high school science for 24 years and am in my seventh year of service at this school. Four volunteer observers (an administrator, a science teacher, and two graduate students of education) made notes to assess fidelity of instruction and bias during the vocabulary instruction in the CG and the TG. While being the teacher and researcher admittedly introduces limitations to the study, there were certain benefits that outweighed potential bias introduced by such practice. For example, when adjustments needed to be made, I was aware of them and noted them in a composition notebook. When classes were interrupted by unforeseen circumstances I was able to make certain adjustments to minimize the impact on the research. The observers were able to ascertain that there were no inconsistencies in the presentation of the terms in the TG and that the terms were treated the same in the CG.
The volunteer observer notes suggest fidelity of instruction was adhered to within each type of class and between the class sections.

**Volunteer Observers**

Four volunteers observed the classes on the days of vocabulary instruction. One administrator and one science teacher from the school where the study took place volunteered to observe. In addition, two graduate students from a nearby Masters in Education program volunteered to observe. They were not informed of any details of the study. I requested the observers make note of the vocabulary instruction using the Observer Rubric provided. It was disclosed to the volunteers that their notes would be used as part of a research study.

The rubric has a space for the term, date, period, and observer name. The data requested consist of a “yes” or “no” answer to the following four questions:

1. Was the term introduced?
2. Was the Frayer Model used?
3. Was the instruction student centered?
4. Was the instruction teacher centered?

In addition, each question has a space next to it for comments and observations. Below the questions a space was provided for the observers to make “Any Observations Noted.” The rubrics were collected and stored in a secure location with all of the artifacts of this study.

There were seven treatment days over the research period. At least one observer was able to attend and make notes on six of the seven days, however on several occasions the observer could not stay for all four classes, subsequently, 64% of the classes were observed during the vocabulary instruction.

**Instruments**

To collect data several instruments were used. These instruments included chemistry course selection pretests, the Frayer Model activity, terms for student knowledge of academic language, the vocabulary knowledge test used as a pre- and posttest, and the pre- and posttests for chemistry content understanding. It was important to use the instruments in a particular order, as follows:

a. Assessments for Chemistry Course Selection — pretest;

b. Vocabulary Knowledge Test (based on the 14 terms deemed troublesome and having everyday meanings as words and specific science meanings as terms) — pretest;

c. Chemistry Content Understanding Tests — pretest;

d. Frayer Model activity — in-class activity, TG only;

e. Vocabulary Knowledge Test — posttest; and

f. Chemistry Content Understanding Tests — posttest.

**Pretests and Posttests**

Four pretests were administered to all student participants and scored to obtain a baseline. These tests consisted of the assessments for chemistry course selection (“Elements Quiz” and “Math Pre-Quiz Evaluation for Chemistry”), the vocabulary knowledge test, and the chemistry content understanding test. The scores from the assessments for chemistry course selection were used to determine the distribution information. Mean, median, and standard deviation were used to analyze variance between and within classes. The scores on the vocabulary knowledge test were used to measure the dependent variable student knowledge of academic language before the Frayer Model activity was implemented with the TG. The vocabulary knowledge test was administered a second time as a posttest after using the Frayer Model activity to determine if there was any significant difference in student knowledge of academic language between the TG and CG. The scores on the chemistry content understanding test were used to measure the dependent variable student understanding of chemistry before the Frayer Model activity was implemented with the TG. The chemistry content understanding tests were also administered as a posttest after using the Frayer Model activity to determine if there was any significant difference in understanding of chemistry between the TG and CG.

**Frayer Model**

The Frayer Model (Billmeyer & Barton, 2002; Frayer et al., 1969), the independent variable in this research, was used as an experimental activity with the TG only. The effect of this treatment was measured by comparing pretest and posttest scores (dependent variables – effect scores) between the TG and the CG.

**Academic Language Terms Tested**

The terms that were the focus of this investigation appear in Table 1 below. Each of these terms has been found to be troublesome to high school students, foundational to national benchmarks, and connected to misconceptions.

These terms were selected because they have certain
meanings in everyday language and specific meanings in chemistry. Cassels and Johnstone (1985), Gardner (1971), Johnstone and Selepeng (2001), and Marshall and Gilmour (1990) have shown these terms to be troublesome for many high school students. These terms are connected to the understanding of concepts surrounding energy transformation and the structure of matter in Benchmarks for Science Literacy (AAAS, 1993). These terms are also related to conservation of energy and the increase of disorder in National Science Education Standards (National Research Council, 1996).

**METHODOLOGY**

The design for this research included two intact groups, a TG and a CG. The participants in the TG and CG received the same pre- and posttests. The two convenience samples were statistically similar (p < .05), making them more effective for comparison (Campbell & Stanley, 1963). The TG received instruction using the Frayer Model activity as a treatment. The total number of student participants in the TG was 48 and the total number of student participants in the CG was 47. Thus, according to Campbell and Stanley (1963) the more similarity between the CG and the TG the more effective this design is, especially in comparing it to a *Pretest-Postest Control Group* experimental design (Campbell & Stanley, 1963, p. 48). It seemed warranted to employ a quasiexperimental *Nonequivalent Control Group Design* (Campbell & Stanley, 1963, p. 13). Table 2 summarizes the treatment conditions applied to each group.

**TIMELINE**

The data collection began early in the first semester and was completed at the end of the first semester. All student participants were informed of this research on the first day of classes. I explained the nature of the research by reading the Informed Consent Form to the student participants and then handing out Informed Consent Forms to all student participants in each class.

The assessments for chemistry course selection were administered to the TG and CG as pretests, to provide data regarding the homogeneity of each class sec-

**TREATMENT CONDITIONS**

The treatment condition consisted of using the Frayer Model activity with the TG. When the 14 academic language terms were first introduced in the curriculum, the Frayer Model activity was used to teach their meanings to the TG. The CG received traditional teacher-led vocabulary instruction as described previously. All of the terms were introduced as they occurred in the curriculum during the semester of the study.

**FRAYER MODEL PROTOCOL**

When each of the 14 academic language terms was first introduced in the curriculum, the TG received instruction as follows:

1. The student participants worked in small groups (four student participants per group). These groups were pre-determined—they were the existing four-member laboratory groups—and remained the same throughout the research.
2. Each group of four student participants received one copy of the Frayer Model with one of the 14 academic language terms to be studied printed in the middle section. Each treatment day, two terms were studied simultaneously. Each group received one copy of the Frayer Model for each term.
3. The student participants had 7 minutes to fill in all four sections on the Frayer Model using their textbooks and other resources, such as dictionaries. The student participants were encouraged to actively use their prior knowledge.

| Table 1 |
| Terms studied in this research |
| conduct | insulate | radiation |
| conservation | liquid | solid |
| energy | mass | temperature |
| gas | matter | volume |
| heat | property |

| Table 2 |
| Summary of Treatment Conditions |

| Treatment Group (TG, \( n_{total} = 48 \)) | Pretest | Pretest | Treatment | Posttest |
| Chemistry | \( n_1 = 24 \) | \( n_2 = 24 \) | \( O_1 \) | \( O_2 \) | X | \( O_2 \) |

Control Group (CG, \( n_{total} = 47 \))

| Chemistry | \( n_1 = 23 \) | \( n_2 = 24 \) | \( O_1 \) | \( O_2 \) | \( O_2 \) |

The “X” indicates that the test or treatment was administered to that particular section.
4. I drew a blank Frayer Model on the board for each term studied. The students were encouraged to share their ideas by independently recording their ideas on the blank Frayer Models on the board. This was included in the 7-minute group activity time.

5. Then I drew a close to the group work by calling attention to the completed Frayer Models on the board and enlisting student participants to help correct any misconceptions or incorrect information. This took approximately 7 or 8 more minutes. As a result, the whole process took approximately 15 minutes. The corrected models were copied by student volunteers and submitted to the faculty participant. Students were neither encouraged nor discouraged from copying the Frayer Models for their own notes.

6. The group Frayer Models were collected and stored in a secure location.

**Classroom Conditions, Fidelity of Instruction**

For fidelity of instruction, an administrator, a staff member, and two graduate students agreed to observe my classes and report on consistency of the use of the Frayer Model with the TG as well as vocabulary instruction with the CG. The 14 terms specifically selected for this research were taught as vocabulary with expository instruction to the CG. The amount of time spent in expository instruction with the CG was comparable to the amount of time spent using the Frayer Model with the TG. The volunteer observers provided me with feedback and their comments were collected as part of the data. I used the comments to make adjustments and ensure as much consistency as possible within each of the groups (TG and CG).

**Analysis of Data**

During the study student scores on the elements quiz, math assessment, vocabulary knowledge pretest, vocabulary knowledge posttest, chemistry content understanding pretest, chemistry content understanding posttest, vocabulary knowledge effect score (the difference between the vocabulary knowledge posttest and pretest scores), and chemistry content understanding effect score (the difference between the chemistry content understanding posttest and pretest scores) were collected.

**Mean, Median, Standard Deviation, and Normality**

The assessments for chemistry course selection were administered to the CG and TG as pretests, to provide data regarding the heterogeneity of each class section. The mean, median, and standard deviation were calculated to determine if the intact class sections had similar distributions on the variables of interest. These tests were administered early in the semester as originally planned. Any student participant who was absent was asked to take the tests as soon as possible after returning to school. The date of the test was noted on each instrument.

The elements quizzes were scored and the data were entered by the research assistant. The elements quiz contained 25 questions and the responses were either correct or incorrect. The scores on the elements quiz reflected the number of correct responses out of a possible 25. The math assessments were also scored and the data were entered by the research assistant. The math assessment contained 11 questions and the responses were either correct or incorrect. The scores on the math assessment reflected the number of correct responses out of a possible 11.

The vocabulary knowledge pre- and posttests were scored and the data were entered by the research assistant. The vocabulary knowledge test contained 14 questions with a possible total of 70 points. The pretests were administered early in the semester during which the study was conducted and the posttests were administered at the end of the same semester.

The chemistry content understanding pre- and posttests were scored and the data were entered by the research assistant. The chemistry content understanding test consisted of four probes, each containing one question that included both an objective portion and an open-response portion. The objective portion was scored as correct or incorrect. The open-response portion asked students to provide a rule or explanation for their objective answers. As a result, scoring the open-response portion required analysis. The scores on the chemistry content understanding test reflected the number of correct objective responses and the analysis of the open responses, with a possible total of 20 points. These pretests were administered early in the first semester during which the study was conducted and the posttests were administered at the end of the same semester.

Table 3 summarizes the results of data analysis for the various test instruments in terms of mean, median, and standard deviation. For the elements quiz and math assessment, the CG mean was slightly higher than the TG mean. For the elements quiz, the CG mean was 20.15 and for the TG the mean was 19.75. For the math assessment, the CG mean was 2.72 and the TG mean was 1.79. The mean vocabulary knowledge pretest
Table 3
Medians, Means, and Standard Deviations for Instruments as a Function of Vocabulary Instruction

<table>
<thead>
<tr>
<th>Tests</th>
<th>Control Group</th>
<th>Treatment Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Median</td>
</tr>
<tr>
<td>Elements Quiz</td>
<td>47</td>
<td>21.00</td>
</tr>
<tr>
<td>Math Assessment</td>
<td>47</td>
<td>2.00</td>
</tr>
<tr>
<td>Vocabulary Knowledge Pretest</td>
<td>47</td>
<td>37.00</td>
</tr>
<tr>
<td>Vocabulary Knowledge Posttest</td>
<td>47</td>
<td>39.00</td>
</tr>
<tr>
<td>Vocabulary Knowledge Effects Score</td>
<td>47</td>
<td>3.0000</td>
</tr>
<tr>
<td>Chemistry Content Understanding Pretest</td>
<td>47</td>
<td>11.00</td>
</tr>
<tr>
<td>Chemistry Content Understanding Posttest</td>
<td>47</td>
<td>12.00</td>
</tr>
<tr>
<td>Chemistry Content Understanding Effects Score</td>
<td>47</td>
<td>3.0000</td>
</tr>
</tbody>
</table>

Scores were slightly higher in the CG, at 37.15, than in the TG, at 37.04. The mean vocabulary knowledge posttest scores were slightly higher for the TG. The mean for the CG was 39.62; the mean for the TG was 40.27. The mean vocabulary knowledge effect score was slightly higher for the TG. The CG mean effect score was 2.4681; the TG mean effect score was 3.2292. The mean for chemistry content understanding pretest was 9.83 for the CG and slightly higher, at 11.35, for the TG. The mean for the chemistry content understanding posttest for the CG was 12.45 and 11.52 for the TG. The chemistry content understanding effect score was 2.6170 for the CG and 0.1667 for the TG. There was a slight decrease in the content understanding scores in the TG compared to the CG.

The Shapiro-Wilk test of Normality results are shown in Table 4. This test was used to discover whether the vocabulary knowledge test scores and the chemistry content understanding test scores were roughly normally distributed. Based on this information, tests were then chosen for further analysis of the data, to discover if any of these results were significant, and to respond to the research hypotheses.

The results for the Shapiro-Wilk tests of Normality indicate that the roughly normally distributed scores include the vocabulary knowledge pretest and vocabulary knowledge posttest. According to the Shapiro-Wilk test of Normality, the elements quiz, math assessment, and chemistry content understanding pre- and posttests were not roughly normally distributed.

Table 4
Normality Test Results for Instruments Used for this Research

<table>
<thead>
<tr>
<th>Test</th>
<th>Shapiro-Wilk Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elements Quiz</td>
<td>.000</td>
</tr>
<tr>
<td>Math Assessment</td>
<td>.000</td>
</tr>
<tr>
<td>Vocabulary Knowledge Pretest</td>
<td>.169</td>
</tr>
<tr>
<td>Vocabulary Knowledge Posttest</td>
<td>.594</td>
</tr>
<tr>
<td>Chemistry Content Understanding Pretest</td>
<td>.003</td>
</tr>
<tr>
<td>Chemistry Content Understanding Posttest</td>
<td>.000</td>
</tr>
</tbody>
</table>
All student participants (total of 95, 47 in the CG and 48 in the TG) completed each test in class. The test scores did not affect individual student grades and the student participants were aware of that. The tests were collected by a student volunteer in each class, and the research assistant scored each instrument and recorded the data. The raw data were not available to me until the full completion of the treatment and the data collection.

The first research question was examined using student scores on vocabulary knowledge pre- and posttests and vocabulary effects scores found by taking the difference between the vocabulary knowledge pre- and posttest scores. These scores were used to determine the distribution within and between class sections on all premeasures and the effect of the Frayer Model vocabulary instruction on high school chemistry students’ vocabulary knowledge. The vocabulary knowledge pretests, vocabulary knowledge posttests and vocabulary effect scores were normally distributed. These data indicate the classes were similar in initial vocabulary knowledge based on the vocabulary knowledge pretest. However, there was no significant difference in vocabulary knowledge effect scores between the CG and TG, indicating that the Frayer Model vocabulary instruction did not have an effect on high school chemistry students’ vocabulary knowledge.

The second research question was examined using student scores on the chemistry content understanding pre- and posttests and chemistry content understanding effects scores found by taking the difference between the chemistry content understanding pre- and posttest scores. These scores were used to determine the effect of the Frayer Model vocabulary instruction on high school chemistry students’ content understanding. Analysis of the data showed a greater increase in chemistry content understanding effects scores in the CG over the TG, indicating that the Frayer Model vocabulary instruction did not have an effect on high school chemistry students’ content understanding.

The third research question was examined using student effect scores on the vocabulary knowledge tests and chemistry content understanding tests. The student effect scores were calculated by taking the difference between the post- and pretest scores for each category of test. These scores were used to determine the effect of the Frayer Model vocabulary instruction on high school students’ knowledge of academic language and understanding of chemistry. Analysis of the data showed no significance comparing the effects scores on the vocabulary knowledge test with the effects scores on the chemistry content understanding test, indicating that the Frayer Model application did not have an effect on high school students’ knowledge of academic language and understanding of chemistry.

**Test Reliability and Validity**

The vocabulary knowledge test created for this study by the researcher was found to be both valid and reliable. The content understanding test used was created by Page Keeley (Keeley, Eberle, & Farrin, 2005) and found both valid and reliable as well.

The test-retest reliability revealed that both the vocabulary knowledge and the content understanding tests were reliable, and the scores of the vocabulary knowledge test fell under a normal distribution. That the tests were found reliable means that they can be used again in future research.

The validity of the vocabulary knowledge test is supported by the normally distributed means across the whole student sample. I collaborated with Reid (2006) while constructing the test. He suggested some formatting and found the test valid for the methodology for this study. Taber (2006) was sent a copy of the test and an explanation of its intended use for this research and he replied that it was a difficult and valid test. A teacher at the school where the research took place analyzed the test and found it valid given the terms and content covered by the curriculum.

The content understanding test probes were piloted in over a thousand participants through *Northern New England Co-Mentorship Network*. They were found valid to test for student misconceptions.

**Summary of Volunteer Observers’ Observations**

Four volunteer observers accomplished 18 observations of the 28 classes where specific vocabulary instruction related to this study was conducted. The observers were asked to view any period they could attend and were given the Observer Rubric when they arrived to the class. I asked the observers to fill out the rubric with what they observed and to add any notes they wished to add on the lines provided.

The analysis of the responses on the Observer Rubric is as follows:

- The term was introduced 100% of the time
- For the CG (periods 1 and 5) the Frayer Model was not used 100% of the time
- For the TG (periods 3 and 7) the Frayer Model was used 100% of the time
- For the CG the instruction was documented as teacher-centered 100% of the time
- For the TG the instruction was documented as student-centered 100% of the time
The CG received traditional teacher-led vocabulary instruction; the Frayer Model protocol was followed in the TG. The voluntary observer protocol reports contained no information to suggest the protocols were not applied as stated.

**SUMMARY OF DATA ANALYSIS**

Although researchers (Gabel, 1999; Marzano, 2004, 2006; Stahl, 1999) have found a relationship between knowledge of vocabulary and content understanding, this study does not support previous research findings. Educators also believe that vocabulary instruction is related to vocabulary knowledge and some research (Marzano, 2006) supports this belief; however, this study does not.

The volunteer observer data support that fidelity of instruction was evident. They noted that the vocabulary instruction in the CG was “teacher centered with little or no student involvement except for copying what the teacher wrote on the board” and that “students were reluctant to offer their own ideas for definitions.” During the specific vocabulary instruction with the TG, a volunteer observer noted that students were “actively engaged in the learning process,” “the group discussions are lively and the students are using the vocabulary with each other to negotiate meaning and share ideas,” and “the students seem to be having fun writing their ideas on the board.” These types of comments were consistent for the TG vocabulary instruction. The artifacts produced in each of the TG classes contain both words and symbols to represent the terms. This was a product of student discussion and participation; the only guidance provided by the faculty participant was that of correcting incorrect information.

In this research, the quantitative variables were scores on the elements quiz, math assessment, vocabulary knowledge pre- and posttest, chemistry content understanding pre- and posttest, as well as vocabulary knowledge effects scores and chemistry content understanding effects scores. An analysis of the confounding variables, which include the amount of contact time, student attendance, and duration of the research is discussed in detail in the Limitation section.

**IMPLICATIONS FOR EDUCATION**

Educators are aware of the value of vocabulary knowledge and vocabulary instruction. I expected to find significant gains in both vocabulary knowledge and content understanding as a result of using specific (Frayer Model) vocabulary instruction. I also expected to find a relationship between vocabulary knowledge and content understanding. The results of this research show that students receiving specific vocabulary instruction using the Frayer Model did not do as well on chemistry content understanding tests as those students not receiving specific vocabulary instruction. However, the student participants receiving the specific vocabulary instruction did slightly better on the vocabulary knowledge tests than those not receiving the specific instruction, but the result was not statistically significant.

Research (Frayer et al., 1969; Herron, 1996; Marzano, 2006; Taber, 2002) on the use of specific vocabulary instruction supports its use for increasing both vocabulary knowledge and content understanding in students. Research (Draper, 2003; Gabel, 1999; Herron, 1996; Marzano, 2006; Mintzes, Wandersee, & Novak, 1998; O’Farrell, 2000) also supports that vocabulary knowledge is related to content understanding. It may be the case that the chemistry content understanding test used for testing students in this study was not an effective test. The test that Marzano (2006) used in his research on vocabulary knowledge and content understanding contained less complex understanding items and he found a significant increase in student understanding related to the specific vocabulary instruction. It may be possible that a content understanding test including less complex items along the continuum of Bloom’s Taxonomy (Bloom, 1976; Committee of College and University Examiners, 1956) would better determine student understanding.

The results of Marzano’s study (2006) show that in the 9th grade the TG did significantly better on the total for multiple-choice responses and the science multiple-choice responses. In grades 1-7 (no results are stated for grade 8), Marzano (2006) found that all of the TG did better on both the science written test and the science multiple choice test. Perhaps the terms used in this research were not appropriately difficult for eleventh grade, given their prior knowledge, at least for the student participants in this study.

Research (Ellis & Worthington, 1994) states that “Time planning instructional activities, time should be considered as an important instructional principle” (p. 15). In this study, time was impacted by enough interruptions to the schedule that the TG met for substantially less contact time. Student participants in the TG also had less contact time due to the higher number of absences recorded for that group compared to the CG. Time also was a factor in that the time period for this study was only two and half months compared to seven months for a similar study by Marzano (2006).
Researchers Fisher, Berliner, Filby, Marlikee, Cahen, and Denshaw (1980) identified three aspects of time that directly impact student learning: 1) the maximum amount of time that is allotted for that activity, 2) the degree to which students are engaged during that allocated time, and 3) the degree to which the students engage in the activity at high rates of success. Research from the Beginning Teacher Evaluation (BTE) (Denham & Lieberman, 1980) supports that these three aspects of time impact on student outcomes directly and positively. Another aspect of time beyond allocated time is opportunity to learn. This relates to the amount of content coverage actually provided for students, whereas as allocated time is the time assigned (Ellis & Worthington, 1994). Researchers (Englert, 1983; Powell, 1979; Wyne & Stuck, 1982) found that variations in student achievement can be explained, at least in part, by variations in content coverage. In this study both allocated time and opportunity to learn were less in the TG. The results of this study—that the group with less allocated time and less opportunity to learn would have less gain in content understanding—are expected based on the research cited above.

In this study the terms were not evenly spread out over the course of the study period. The terms were introduced as they emerged in the curriculum and the last few weeks of the study were laden with the introduction of terms and specific vocabulary instruction in the TG.

Research (Enge, 2005; Marzano, 2006; Monroe & Pendergrass, 1997) indicates that specific vocabulary instruction can influence vocabulary knowledge and content understanding, and that vocabulary knowledge is related to content understanding. The model used for this research has been shown to be effective in improving student vocabulary knowledge and content understanding (Enge, 2005; Marzano, 2006; Monroe & Pendergrass, 1997). The results of this research indicate that time, both allocated time and opportunity to learn, may be related to student achievement. The results also indicate that average to high-achieving eleventh grade students may not benefit from specific vocabulary instruction using terms they have seen before. Finally, the results of this research indicate that the test used for assessing student participants’ ability to evaluate misconceptions may not have been appropriate to test student understanding related to vocabulary knowledge.

**UNINTENDED RESULTS**

I find it important to report on some issues that emerged during the study.

The vocabulary knowledge test created for this research was found to be statistically both valid and reliable. As a result, it can be used for further research.

Student participants in the TG still had a gain in content understanding, even though the contact time was 20% less than in the CG over the ten-week period.

Data regarding time lost were provided to the faculty council and administration of the school in the research study. The data helped inform a discussion about whether to make changes to the schedule for the next year.

The Frayer Model was altered by the students over the research period. Originally, the boxes on the model contained the prompts “Essential Attributes” and “Non-essential Attributes.” Early in the study, the students erased those prompts from the board and wrote “must have/must be” in place of “essential attributes, and “may have/may be” in place of “Non-essential Attributes.” These prompts were found to be more meaningful to the students. As a result, I will adopt these prompts when using the model in the future.

Although the data did not show significant results from the use of the Frayer Model, there was a slight increase in the mean vocabulary effect score in the TG as compared to the CG. The student participants in the host school were informed of the methodology after the research period and have become more involved in their own note making. As a result, they have been showing their teachers in other subjects how to use the Frayer Model for taking notes, making comparisons, and clarifying terms. One student from the 5th period has renamed the Frayer Model “LaBox.”

**LIMITATIONS**

I recognize several limitations to this study. First of all, I was the faculty participant for all four classes in this study. While it enabled me to control course content, treatment implementation, and instructional techniques, I recognize that it could have introduced certain bias into the study.

Due to the small sample size (n = 95) and the homogeneity of the student participants, the results may not be generalizable to other populations. I also could not control interruptions to contact time. The number of student days absent due to sports and other extracurricular activities was unexpected and added to the amount of instruction time lost. The TG was more heavily affected by the time loss.

Also, the research was conducted in a small school setting, allowing the possibility of contamination by student participants from the TG sharing the specific
vocabulary instruction with student participants in the CG. At the conclusion of the research period, I questioned the student participants to determine whether they were aware of when the treatment was being administered or which instructional strategies were related to the research study. As far as I was able to determine, the student participants were not aware of any of the research investigation.

The time period of 10 weeks, or two and a half months, may not have been long enough for significant effects to be measured in either vocabulary knowledge or chemistry content understanding.

**AREAS FOR FUTURE RESEARCH**

Based on the findings of this study, I would suggest the following areas for further research:

1. Effective methods for vocabulary instruction in high school,
2. The nature of content tests used for assessing student understanding,
3. The effect of direct instruction related to chemistry misconceptions, combined with vocabulary instruction using a context-based model (such as the Frayer Model),
4. The effect of specific vocabulary instruction on vocabulary knowledge and content understanding with students younger than the eleventh graders in this study,
5. The amount of time high school students need to process new terms and make connections to content understanding.

**SUMMARY**

Specific vocabulary instruction has been advocated for many reasons. Vocabulary knowledge has a strong positive relationship with intelligence (Davis, 1944; Spearitt, 1994; Thorndyke & Lorge, 1944), one’s ability to comprehend new information (Chall, 1958; Harrison, 1980), and one’s level of income (Sticht, Hofstetter, & Hofstetter, 1997). Direct vocabulary instruction is related to student content understanding (Enge, 2005; Frayer et al., 1969; Marshall & Gilmour, 1990) and student achievement. This study does not corroborate these findings. Based on the discussion of the results, the findings of this research are inconclusive. The purpose of this research was to determine the effects of specific vocabulary instruction on vocabulary knowledge and content understanding in high school chemistry students and to determine if there was a relationship between vocabulary knowledge and content understanding. I found only a slight increase in vocabulary knowledge, but it was not statistically significant. I also found that the CG improved in content understanding over the TG and there was no relationship found between vocabulary knowledge and content understanding.

More research, over longer duration, needs to be conducted with high school students. The research could explore such variables as minimizing disruption to contact time, addressing the effects of student prior learning, including direct instruction related to alternative conceptions, and allowing students to keep the vocabulary artifacts for review.

**References**


Billmeyer, R., & Barton, M. L. (2002). *Teaching reading in the content areas: If not me then who?* (2nd ed.). Aurora, CO: MCREL.


The Characteristics of Effective Professional Learning Experiences for Science Teachers

Wesley Johnson
Burlington, MA, High School

ABSTRACT

Science teachers need high-quality professional learning experiences to enhance their instructional practices and raise their students’ achievement. Existing professional development programs need improvement. This paper reviews the research literature on the effects of professional development programs for science teachers. It attempts to determine the characteristics of teachers’ learning experiences which are most likely to lead to increased student achievement. A model of effective teacher learning is presented in order to explain why these characteristics are necessary.

INTRODUCTION

In recent decades, reform documents have highlighted the need to improve science education in the United States (American Association for the Advancement of Science [AAAS], 1993; National Commission on Teaching & America’s Future [NCTAF], 1996; No Child Left Behind [NCLB], 2002; National Research Council [NRC], 1996). These documents insist that teachers must increase their knowledge of the content they teach (NCTAF, 1996) and improve their instructional methods. Specifically, the NRC (1996) calls for greater use of the inquiry method in teaching science.

Educators have designed professional development programs to improve science teaching, but these programs have shown only limited success. Convincing teachers to change their classroom practices is notoriously difficult. In order to be effective, professional development programs need to be designed with an understanding of how teachers learn.

Researchers have become increasingly interested in the relationship between teachers’ knowledge and their practice (Evans, 2002). From this work, a broader conception of professional learning has emerged. High-quality professional learning experiences empower teachers and enable them to continuously improve their own practice. These experiences do more than just tell teachers what or how to teach.

Science teachers are not alone. Researchers agree that professional development is vitally important to any education reform (Birman, Desmone, Porter, & Garet, 2000; Cohen & Hill, 2000; Elmore & Burney, 1999; Guskey & Huberman, 1995; Loucks-Horsley, Love, Stiles, Mundry, & Hewson, 2003; NCTAF, 1996). But only a few professional development programs have been shown to contribute to increased student achievement. This situation presents educators with a dilemma—professional development is essential, yet effective programs have been hard to find. The challenge is to determine which programs work and why.

Research on a broad range of professional development programs has revealed several characteristics common to effective programs. These characteristics, viewed with an understanding of how teachers learn, can serve as guides to the development of effective professional learning experiences for science teachers.

Why has effective professional development been so hard to find? The answer to this question is complex. First of all, teachers do not respond to learning experiences in a straightforward way. Second, the effects of professional development programs on student achievement are difficult to measure. Both of these difficulties must be dealt with in order to design the professional development experiences science teachers need.

PROFESSIONAL DEVELOPMENT PROGRAMS FOR SCIENCE TEACHERS

The reform documents call for improvements in teacher knowledge, teacher practice, and student achievement. Specifically, the NRC’s National Science Education Standards (1996) call for professional learning experiences that help teachers integrate their knowledge of science content, scientific inquiry, and pedagogy. Simply giving teachers more hands-on activities to perform with their students is not enough.

Research studies on the effects of professional development programs for science teachers have shown some positive results. These studies have also had shortcomings. Research shows that programs designed around lessons plans can lead to their use (Basista & Mathews, 2002; Radford, 1998; Rigden, 1999; Supovitz & Turner, 2000), but these studies do not provide evidence of significant changes in teachers’ pedagogical approaches. Other studies show that increasing teacher knowledge of student learning can improve teacher practice (Fennema et al., 1996; Holstein, Carmeli & Shore, 2004), but
these studies do not show evidence of increased student achievement. Many professional development studies rely on survey responses for their data. Such studies can offer useful descriptions of a program. They cannot, however, analyze the effects of a professional development program on teaching and learning. Better research is needed to help educators design the professional learning experiences science teachers need. Educators also need a better understanding of how teachers learn.

**A SELF-SUSTAINING, GENERATIVE CHANGE MODEL OF PROFESSIONAL LEARNING**

Educators recognize the need for a better theoretical foundation from which to understand teacher learning (Evans, 2002; Guskey & Huberman, 1995; Hyde & Pink, 1992; Moore & Barab, 2002). According to Ball and Cohen (1999), our current “theories of professional learning have been implicit and undeveloped” (p. 5). Any new theory must be able to explain how a teacher's learning affects his or her practice (Ball & Cohen, 1999; Pink & Hyde, 1992).

Cochran-Smith and Lytle (1999) differentiate among three conceptions of teacher knowledge: knowledge-for-practice, knowledge-in-practice, and knowledge-of-practice. Knowledge-for-practice is explicitly stated knowledge of subject matter, pedagogy, and student learning. It is the type of knowledge most often addressed in both university courses and professional development programs—the kind of knowledge teachers acquire in one setting and then use in another. Knowledge-in-practice is very different. It is the practical knowledge embedded in the craft of expert teachers. Unlike knowledge-for-practice, knowledge-in-practice cannot be separated from teachers’ actions and reactions in specific situations.

Cochran-Smith and Lytle's (1999) third conception of teacher knowledge, knowledge-of-practice, rejects the separation between a teacher's formal knowledge and his or her knowledge in action. Knowledge-of-practice is created when teachers connect events in their classroom with larger issues, both educational and political. This requires teachers to critically examine their practices and assumptions, turning their classrooms into sites of inquiry. Research based on knowledge-of-practice is not only concerned with determining what good teaching looks like, but also with “understanding, articulating, and ultimately altering practice and social relationships in order to bring about fundamental change” (p. 279).

Franke, Carpenter, Fennema and Behrend (1998) also focus on the link between teacher learning and practice in their self-sustaining, generative change model of professional learning. This model defines high-quality professional learning experiences as those which enable teachers to continuously improve their instruction. These experiences provide teachers not only with procedures to follow but also with an understanding of the principles behind them. This gives teachers principles through which to evaluate their practices, enabling them to sustain reform initiatives and generate their own improvements. Unless teachers actively engage in their own professional growth, they are unlikely to undergo the substantial changes necessary to improve student achievement.

Other educators have used similar thinking when describing effective professional development programs. They suggest designing programs to support long-term growth (Guskey & Huberman, 1995; Hyde & Pink, 1992; Miller, 1992; Richardson, 2003) and tailoring these programs to teachers' strengths instead of their weaknesses (Elmore & Burney, 1999; Pink & Hyde, 1992; Loucks-Horsley et al., 2003). Sykes (1999) describes professional development not as a discrete activity but as a fundamental disposition towards improved practice. It is an attitude that will eventually lead to higher student achievement. Loucks-Horsley et al. (2003) found that students do better in schools where teachers are also learning. However, the process through which a teacher’s learning affects his or her practices is not simple.

**CHARACTERISTICS OF TEACHER KNOWLEDGE**

Reformers have found that many teachers resist efforts to change how they teach. This is a natural result of how teachers learn. Throughout their careers, teachers amass a unique, idiosyncratic collection of instructional ideas and practices, what Van Driel, Beijaard and Verloop (2001) refer to as their “personal practical knowledge.” Changing a teacher's thinking is a difficult and incremental process in which new ideas and practices are built off of older ones (Ball & Cohen, 1999; Borko & Putman, 1995; Cohen & Hill, 2000; Darling-Hammond & McLaughlin, 1999; Davis, 2003; Sykes, 1999). Reformers cannot expect teachers to immediately adopt every new idea or practice presented to them. Ideas must be internalized, tried out, and reflected upon before they are likely to be consciously incorporated into a teacher's thinking (Kelly, 2006).

Depending on how it is used, a teacher's practical knowledge can be either a source of valuable experiences and ideas or an obstacle to change. Lortie (1975/2002) observed that the way in which teachers
learn about their profession does not often “lay the basis for informed assessment of teaching technique or encourage the development of analytic orientations toward the work” (p. 67). As a result, many teachers repeat the same “common-sense” instructional decisions over and over without ever giving them serious thought.

The haphazard development of teachers’ ideas and practices can have other consequences as well. Two independent studies examined teachers’ thinking over three years and reached the same conclusion: teachers’ stated beliefs are often inconsistent with their teaching practices (Fennema et al., 1996; Rodrigues, Zozakiewicz & Yerrick, 2005). The same conclusion was reached by Kennedy’s (1999) review of the Teacher Educator and Learning to Teach (TEL T) survey which invited pre-service teachers to respond to hypothetical classroom situations. How can teachers be expected to align their practices with reformers’ ideas if they do not even align their practices with their own thinking? A special kind of professional learning experience is required to help teachers coordinate their beliefs and practices.

**Reflective Practice**

Reflective practice forces teachers to examine their instructional choices and articulate the reasoning behind them. For Davis (2003), this is a necessary first step to changing a teacher’s practices. Reflective practice involves teachers consciously and carefully observing, questioning, and experimenting with their lessons. By paying attention to student feedback, teachers can find ways to improve their lessons. Significantly, the ideas behind these improvements come from the teachers themselves. Therefore, improvements can continue even after the professional development program ends. Van Driel et al. (2001) suggest several practical ways for teachers to become more reflective. These include performing curriculum work, taking part in collaborative action research, visiting with peer coaches, and discussing case studies.

Multiple studies have shown that mathematics teachers who reflect on their practices improve both their mathematical content knowledge and their instructional practices (Akerson, 2004; Barnett, 1998; Daehler & Shinohara, 2001; Goodnough, 2001; Schifter, 1996). A similar conclusion was reached by Huffman, Thomas and Lawrenz’s (2003) study of over 200 teachers’ reactions to different forms of professional development. Teachers claimed that reflective practice was highly effective, improving both their content knowledge and instruction.

**Teacher Learning and Professional Development**

A professional development program cannot be accurately evaluated based on the content of the program itself. Instead, a program must be evaluated based on the extent to which it promotes lasting improvements in how participants teach. Researchers should look for evidence of these enduring changes, not just evidence that teachers are using one or two new activities.

Templin and Bombaugh (2005) used a “praxiological paradigm” to examine changes in teachers’ behavior. They evaluated the curriculum goals of professional development programs in terms of their ability to encourage teachers to participate in reflective practice. Lustick and Sykes (2006) took another look at how participation in a professional development program affects teachers. They identified three different levels at which preparing for National Board Certification changed participants’ practice. In their study, half of the teachers dynamically changed their practice while the other half either only adjusted their practice in order to complete the program or deferred any changes to a later time.

In summary, teacher knowledge is a growing commodity, not only a prerequisite for good teaching (Davis, 2003; Van Driel et al., 2001). High-quality professional learning experiences prompt teachers to reflect upon and improve their own practice. To be effective, professional development programs must do more than just improve teachers’ knowledge and skills. If the ability to generate and sustain change is not passed on to the teachers themselves, these improvements will not endure. This view of teacher learning helps explain what makes some professional development programs effective.

**Research on Effective Professional Development Programs**

Educators agree that, in order to be called effective, professional development must ultimately raise student achievement (Ball & Cohen, 1999; Guskey, 1998; Loucks-Horsley et al., 2003; Pink & Hyde, 1992; Rigden, 1999). Educators also agree that most professional development activities are poor (Ball & Cohen, 1999; Birman et al., 2000; Borko, 2004; Bransford, Brown & Cocking, 1999; Cohen & Hill, 2000; Desimone, Porter, Garet, Yoon & Birman, 2002; Garet, Porter, Desimone, Birman & Yoon, 2001; Hawley & Valli, 1999; Loucks-Horsley et al., 2003). Still, a consensus has begun to emerge from the literature concerning the characteristics of effective professional development programs.
THE INADEQUACY OF CURRENT PROFESSIONAL DEVELOPMENT PROGRAMS

Several different reasons have been advanced to explain why most professional development programs are so poor. Garet et al. (2001) noted that high-quality professional development is difficult to create and expensive to implement. Birman et al. (2000) add that most existing programs are too short and are disconnected from other reform initiatives. Yager (2005) observed that most professional development experiences, even long-running summer workshops, do not address how to implement the material they cover. Workshop designers have also been criticized for failing to address the many complex reasons behind why teachers do what they do (Franke et al., 1998). They often tell teachers what to do without explaining the reasons why to do it (Yager, 2000).

Van Driel et al. (2001) link ineffective professional development activities to what they call the traditional science education reform cycle. In this cycle, educational leaders adopt a new idea and offer teachers workshops on how to implement it. The teachers return to their classrooms, do not alter their practices, the reform fails, and the next one is tried. To change this cycle, reformers must give teachers more control over their own learning. After all, the reform’s fate is ultimately in their hands.

DIFFICULTIES IN CONDUCTING RESEARCH ON EFFECTIVE PROFESSIONAL DEVELOPMENT PROGRAMS

There is little evidence, albeit an increasing amount, linking professional development with student achievement (Desimone et al., 2002; Guskey & Sparks, 1996; Hiebert, 1999; Killion, 1998; Loucks-Horsley, 1999; Loucks-Horsley et al., 2003; Sykes, 1999). Such research is extremely difficult due to the large number of complicating factors between the two variables (Huffman et al., 2003). Reformers may have clear goals for a professional development program, but these goals may not be reflected in the participants’ actually learning, their desire or ability to implement curriculum changes, or their students’ responses.

Rodrigues et al. (2005) observe that most research on professional development programs provides either “cheerful narratives” of all the great elements of one program or “deficit narratives” of all the problems with another (p. 360). These descriptions do not carefully and critically examine why something worked or failed. Akerson (2004) questions whether these issues really concern most professional development planners.

These educators are typically more interested in planning follow-up activities than they are in studying the results of their programs.

Hyde and Pink (1992) further criticize most professional development research for being non-objective and short-sighted. Many researchers play a role in both designing and evaluating professional development programs—leading to a conflict of interest. Also, measuring the effects of a professional development activity can take years (Garet et al., 2001; Goldston & Bland, 2002). For example, the dramatic increases in student achievement found through the Science Education for Equity Reform project (Project SEER) did not materialize until the program’s fourth year (Rigden, 1999). Many professional development studies are completed in less time.

CONDUCTING RESEARCH ON PROFESSIONAL DEVELOPMENT PROGRAMS

Researchers have developed a number of strategies to deal with the difficulties in professional development research. Guskey (1998, 2002) recommends incorporating the evaluation of a professional development program into its planning, design and implementation. He also proposes a five level system for professional development research. The first level involves the participants’ reactions, the second is the amount of content they learned, the third is the degree to which this learning is supported by their schools, the fourth is the extent to which they use the new knowledge, and the fifth level is the effect on student achievement. Success at one level is necessary for success at the next, and each successive level is more difficult to measure. So, before trying to link professional development directly with student achievement, Guskey (1998, 2002) suggests examining each of the four lower levels first.

Changes in teachers’ thinking are difficult to measure. Hyde and Pink (1992) recommend performing qualitative and analytical research to uncover these changes. Such studies can take many forms including case discussions, lesson studies, interviews, action research, and other research methods tied closely to teachers’ daily actions (Bernauer, 2002; Bransford et al., 1999; Connelly, Clandinin & He, 1997; Loucks-Horsley et al., 2003). Quality research should, in the end, help educators determine which types of professional development activities positively affect student learning.

In reviewing the literature on effective professional development, Loucks-Horsley (1999) observed that those studies which included evidence of increased student achievement generally reached the same conclusion as those that did not. Also, Fennema et al. (1996)
found a correlation between the amount of change in teachers’ practices and the increase in student achievement. These two conclusions indicate that a consensus on the qualities shared by effective teacher learning experiences is possible.

**CHARACTERISTICS OF EFFECTIVE PROFESSIONAL LEARNING EXPERIENCES**

Many researchers have commented on the characteristics of effective professional learning experiences, often specifically identifying a “consensus” (Davis, 2003; Desimone et al., 2002; Desimone, Smith & Phillips, 2007; Elmore & Burney, 1999; Garet et al., 2001; Hawley & Valli, 1999; Justi & van Driel, 2005; Loucks-Horsley, 1999; Lusin & Sykes, 2006; Supovitz & Turner, 2000; Sykes, 1999; Wilson & Berne, 1999; Yager, 2000). What they call a consensus is, in fact, a rather broad collection of characteristics. This is partly because these researchers approach the topic from different viewpoints and with different purposes. Still, among the diversity, there emerges a set of common elements which could be described as a consensus of the consensuses. In general, effective professional learning experiences have been found to be:

- Centered on the curriculum being taught
- Directly connected to teachers’ practices
- Focused on student learning
- Long-term or ongoing
- Coherent with other goals in the school community
- Designed to allow for teacher collaboration

Garet et al. (2001) reached similar conclusions from their analysis of over 1,000 teacher surveys on self-reported changes in knowledge, skills, and practices. They concluded that, “professional development that focuses on academic subject matter (content), gives teachers opportunities for ‘hands-on’ work (active learning), and is integrated into the daily life of the school (coherence), is more likely to produce enhanced knowledge and skills” (p. 935).

Some other commonalities exist within the research literature on effective professional development. These include having a comfortable balance between new information and opportunities for discussion (Freeman, Marx & Cimellaro, 2004; Melber & Cox-Petersen, 2005; Posnanski, 2002), a presentation style which practices what it preaches (Borko & Putnam, 1995; Loucks-Horsley et al., 2003; Rigden, 1999), and an emphasis on quality learning over quantity (Appleton, 2003; Bransford et al., 1999).

Conspicuously absent from the consensus are recommendations for which topics to address. Narrowly focused, quick-fix workshops have been widely derided by the literature. For example, Cohen and Hill’s (2000) study of California teachers found that special topic workshops do not raise student test scores. Also, the consensus does not mention what form professional development activities should take. In two studies which surveyed teachers on their classroom practices, Garet et al. (2001) and Lee, Hart, Cuevas and Enders (2004) both concluded that the content of a professional development program is more important than its form.

The question remains, however, why do these characteristics make professional development effective? What do these elements provide that enables teachers to continuously transform their own practices? Answers to these questions can be found by examining each characteristic through a self-sustaining, generative view of effective teacher learning.

**CURRICULUM-BASED PROFESSIONAL LEARNING EXPERIENCES**

Many researchers have concluded that professional learning experiences should directly involve the curriculum which teachers are responsible for implementing (Darling-Hammond & McLaughlin, 1999; Cohen & Hill, 2000; Sykes, 1999). This type of professional development can involve designing new units or courses, implementing a pre-made curriculum, or enhancing a curriculum already in place. Such experiences have been shown to increase teachers’ knowledge of the content they teach (Cohen & Hill, 2000; Loucks-Horsley et al., 2003; Huffman et al., 2003). Furthermore, Miller (1992) claims that, “When teachers are empowered to direct their own curriculum work, they become empowered to direct their own staff development” (p. 102).

**PRACTICE-BASED PROFESSIONAL LEARNING EXPERIENCES**

Researchers also agree that effective professional learning experiences are centered on teacher’s day-to-day practices. These experiences give teachers an opportunity to try out new ideas in the classroom and receive feedback (Birman et al., 2000; Bransford et al., 1999; Davis, 2003; Desimone et al., 2002). Teachers who try out new practices are more likely to continue to use them (Barnett, 1998; Loucks-Horsley et al., 2003). Teachers make their instructional decisions from their base of teaching experiences. New ideas must become part of this base, and must be seen to work (Hawley & Valli, 1999), in order for teachers to accept them.
STUDENT LEARNING-BASED PROFESSIONAL LEARNING EXPERIENCES

Effective professional learning experiences deal with how students learn specific content (Borko, 2004; Borko & Putman, 1995; Darling-Hammond & McLaughlin, 1999; Garet et al., 2001; Sykes, 1999). Understanding student learning is a difficult but essential part of reflective practice. Many teachers have a firm grasp of the content and a sensible strategy for teaching it. But without knowledge of how students learn the material, they will be unable to understand why a particular approach works or fails. Ball and Cohen (1999) argue that teachers need knowledge of multiple theories of student learning in order to better understand and improve their practice. Furthermore, studying student learning has been found to both improve teachers’ content knowledge and their instructional practices (Akerson, 2005; Barnett, 1998; Fennema et al., 1996).

TIME AND COHERENCE IN PROFESSIONAL LEARNING EXPERIENCES

Researchers agree that effective professional learning experiences should be long-term or ongoing. This reflects the self-sustaining, generative change view of how teachers change. It is a slow process which requires constant support (Ball & Cohen, 1999; Franke et al., 1998; Loucks-Horsley, 1999).

As discussed earlier, teachers’ practices are often inconsistent with their stated beliefs. Therefore, giving teachers new instructional goals, even goals which the teachers believe in, will not necessarily change their instructional methods. Similarly, having teachers adopt new teaching methods may not change their vision of what science education is. To be effective, reforms must provide a coherent view of what the goal is, why it is important, and how it can be achieved.

Professional development is just a single part, albeit a crucial one, of any education reform initiative. The changes sought by a professional development program must be supported by the school’s administration and policies in order for a reform to work (Loucks-Horsley, 1999; Pink & Hyde, 1992). The school community as a whole must buy into the reform before it has any hope of succeeding (Davis, 2003; Richardson, 2003).

Some reformers claim that professional development programs should be coordinated with other reforms as part of a simultaneous, large-scale change of an entire school (Elmore & Burney, 1999; Long, 1996; Moore, 2000; Pink & Hyde, 1992). Changing the entire system at once could force teachers to critically evaluate their current practices. But it could also lead them to develop a coping strategy in which they seek out ways to maintain their existing ideas and practices, spoiling the reform.

COLLABORATION AS A PROFESSIONAL LEARNING EXPERIENCE

Collaboration is vitally important to teachers’ learning (Ball & Cohen, 1999; Bransford et al., 1999; Davis, 1998; Goodnough, 2005; Loucks-Horsley et al., 2003; Showers & Joyce, 1996). Kelly (2006) argues that teacher learning can best be understood from a socio-cultural perspective. In this view, knowledge resides in teachers’ practices, not in their heads. Teachers learn through creating shared meanings within the educational community. Expert teachers are those who are able to participate fully within this community. They integrate research results and policy initiatives with their own practical knowledge to strengthen their instruction. They also share ideas with other educators.

In a socio-cultural view, a teacher’s knowledge is formed by his or her participation in culturally determined roles. Teachers with similar experiences can provide uniquely valuable support and advice to one another. Collaboration allows them to focus on their immediate classroom experiences, not on theories and policies created elsewhere. Most importantly, collaboration enables teachers to build their own knowledge. Researchers have found that teachers can increase their own content knowledge through collaborating with one another (Barnett, 1998; Goodnough, 2001; Loucks-Horsley et al., 2003).

CONCLUSION

In order to improve student understanding of science, teachers must make greater use of the inquiry method. A low-quality professional development program may try to bring about this change by merely giving teachers additional hands-on activities to perform. Such a program is unlikely to bring about the kind of changes in teacher practice and student learning that educators agree are necessary.

A professional development program with a higher likelihood of being effective must focus on student learning. It should examine how and why the inquiry method should be used. It should be centered on the teacher’s curriculum and practice, exploring new ways of addressing existing concepts and helping teachers design new activities. It should provide ongoing support, giving teachers a chance to try out and reflect on new lessons and strategies. Finally, it should provide
opportunities for collaboration, allowing teachers to
learn from one another's ideas, successes, and chal-
genches. This type of high-quality professional learning experience empowers teachers to take control of the
reform. They would be enabled and encouraged to
reflect on and continuously improve their own practice.

Identifying the characteristics of an effective profes-
sional development program for science teachers does not, however, make such a program easy to create.
Effective professional learning experiences do not come from a one-size-fits-all program. Instead, each program
must be tailored to the specific needs of the teachers and
school community it serves (Garet et al., 2001; Kimmel,
Deek, Farrell & O'Shea, 1999; Loucks-Horsley et al.,
2003; Pink & Hyde, 1992; Sykes, 1999; Yager, 2005).
The first step to designing an effective professional
development program is to clearly articulate a vision
of science education (Loucks-Horsley et al., 2003). Second, data should be collected both on students' needs (D'Ambrosio, Harkness & Boone, 2004) and teachers' beliefs about teaching and learning science (Van Driel et al., 2001). Third, teachers should help
determine what changes are needed (Darling-
Hammond & McLaughlin, 1999; Guskey & Huberman,
1995) and which of the many available methods of pro-
fessional development should be tried (Shaw, 2003).

References


Educational Resources

Can Participation in a School Science Fair Improve Middle School Students’ Attitudes Toward Science and Interest in Science Careers?
Valerie Finnerty

ABSTRACT
This paper examines the possibility of the use of authentic inquiry in the form of school-based science fair projects to improve middle school students’ attitudes toward science and increase their interest in science and engineering careers. A review of the literature reveals that students’ attitudes toward science are a complex construct that includes personal, school, and social factors. Inquiry-based instruction may create more positive attitudes, although experiences must be both supported and authentic. Students’ interest in science careers is also linked with attitudes toward science, and both are correlated with science self-efficacy. Through successful completion of a science fair project, students’ science self-efficacy may be enhanced. Further research on the effect of involvement in authentic science experiences is proposed.

INTRODUCTION TO THE PROBLEM
Is the United States facing a shortage of scientists and engineers in the near future? A November 17, 2007 news headline in The Seattle Times proclaimed, “Gates sees engineer shortage looming” (Romano, 2007). In the article, Microsoft Chairman Bill Gates described what he sees as a growing crisis in the United States; a mounting deficiency of scientists and engineers. His concern mirrors the report of the Committee on Prospering in the Global Economy of the 21st Century: An Agenda for American Science and Technology (the Committee), Rising above the Gathering Storm (2007), which warns that the United States is facing the loss of competitiveness in the global economy if something is not done to strengthen the infrastructure of science, technology and engineering in America, including the areas of K–12 education, research, higher education, and economic policy. The Committee cautions that, “Without a flourishing scientific and engineering community, young people are not motivated to dream of ‘what can be,’ and they will have no motivation to become the next generation of scientists and engineers who can address persistent national problems, including national and homeland security, healthcare, the production of energy, the preservation of the environment, and the growth of the economy, including the creation of jobs” (p. 112-113). They point out that the goal that needs to be met is the creation of a scientifically literate workforce that will be able to fill future jobs in science, mathematics and engineering education as well as in science and engineering fields.

The U.S. Department of Labor (2007) projects that the need for life scientists will grow by 12.8% between 2006 and 2016; and increase of 34,000 jobs. They estimate that during the same time period, the demand for physical scientists in the United States will increase by 15.7%.

Engineering jobs are also expected to increase on average by 10.5%. The highest growth is predicted to occur in biomedical engineering with an increase of 21.1%, civil engineering with an increase of 18%, environmental engineering with an increase of 25.4%, and industrial engineering with an increase of 19.1%. Although Lowell and Salzman (2007), in their refutation of the claims put forth in Rising above the Gathering Storm (Committee, 2007), claim that the situation is not as dire as the Committee asserts, they do acknowledge that shortages in specific fields may be looming. It should be pointed out that with any increase in the demand for professional scientists and engineers comes a concurrent rise in the need for technicians and other associated science workers, which means that many more jobs will be created to support the professional positions.

On August 9, 2007, President George W. Bush signed into law the America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Act (America COMPETES) (U.S. Office of the Press Secretary, 2007). The Act is a response to the crisis outlined by the Committee (2007). Included among the provisions of the Act are increased funding for scientific research, scholarship funds to universities for students majoring in science and teaching, and proposed methods for improving the teaching and learning of science and mathematics from elementary through graduate programs. In this way, the
Act puts into action the first three recommendations from the Committee (2007) to focus on education, research and higher education.

The Act is a reaction to a growing shortage of American scientists and engineers. The U.S. Department of Labor: Bureau of Labor Statistics (February, 2006) predicts a 28.4 percent growth in professional, scientific and technical service jobs by 2014. According to their projections, of the sixty fastest growing professions requiring an Associate degree or higher, 36 are science- or technology-related. However, in spite of the current and projected demand for workers with scientific knowledge and training, degree-granting institutions awarded fewer science-related bachelor's, master's and doctor's degrees in 2004-2005 than they did in 2001-2002 (U.S. Department of Education: National Center for Educational Statistics, 2001, 2006). This was true for all major categories of degrees in science, health and engineering with the exception of biological science degrees, which increased during this time period.

Research from the field of education makes it seem unlikely that younger students will go on to major in science fields either, if results from the Trends in International Mathematics and Science Study are taken into account. Although about 73% of 8th grade students answered “like” or “like a lot” when asked, “How much do you like science?”, and about 75% of them said that they enjoy learning science, only 47% stated that they would like a job that involved using science (American Institutes for Research, 1999). These data seem to reflect a disconnect between school science and the workforce, since students do not seem to associate liking the science that they study in school to the possibility of enjoying a career in science. A number of studies done since the 1990s (Greenfield, 1997; Hedges & Woodworth, 1996; Moffat, et al., 1992; Shymansky, Weinburgh, 2000) have documented that students of both genders show a deterioration of their attitudes toward science as they enter adolescence, although female students often show more loss of positive attitudes than males (Greenfield, 1997; Jones, Howe & Rua, 2000; Reid & Skryabina, 2003; Weinburgh, 1995, 2000; Woolnough, 1996).

A multitude of studies have found that nontraditional instructional practices, with an emphasis on greater student involvement and project-based activities, are associated with improving attitudes toward science and increasing interest in science careers (Fouts & Myers, 1992; Häusssler & Hoffman, 2002; Reid & Skryabina, 2003; Russell & French, 2002; Shymansky, Hedges, & Woodworth, 1990; Trotter & Jones, 2003; Tuan, Chin, Tsau, & Chang, 2005; Yager, Lim & Yager, 2006; Yerrick, 2000). Other studies (Gibson & Chase, 2002; Schneider, Krajeik, Marx, & Soloway, 2002) have associated both higher science achievement and/or more positive attitudes about science with inquiry-based science programs in particular.

The National Science Education Standards (National Research Council, 1996) first describe inquiry as the pursuit of knowledge the same way that practicing scientists go about their work. It is “the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work” (p. 23). The Standards go on to say that inquiry also describes the activities that students do in classrooms to find out about the natural world, develop their understanding of scientific concepts, and learn about how scientists make discoveries and add to or revise scientific knowledge. Is this duality of the inquiry done by scientists and the inquiry done in classrooms one source of the disconnect between school science and careers in science? Another source of this problem may be that there is a wide variety of classroom practices which are characterized as inquiry-based.

Inquiry-based science education can be categorized by the degree of complexity and student independence. Colburn (2000) described three levels of inquiry; structured, guided, and open, with structured inquiry having more teacher direction and open inquiry having less. On the other hand, Wenning (2005) describes five levels of inquiry-oriented science teaching practices, from discovery learning, with the lowest intellectual sophistication and with more teacher control, to free and hypothetical inquiry activities, which reflect the highest intellectual sophistication and with have student control. Chinn and Malhotra (2002) characterize classroom inquiry by its authenticity; that is, how closely it approaches the inquiry done by scientists.

A science fair project is an example of an open (Colburn, 2000), free (Wenning, 2005) or authentic (Chinn & Malhotra, 2002) inquiry project with a high intellectual sophistication and a high level of student control and independence. A science fair is defined as an opportunity for students to “conduct an investigation using the scientific method to discover the answer to a scientific problem” (VanCleave, 1997, p. 5). It is proposed that engagement in a authentic inquiry activity such as a science fair may help to close the gap between students’ perception of school science and real science, perhaps improving student attitudes toward science and interest in science as a career.
**THE PROBLEM**

The purpose of this literature review will be to explore the factors which contribute to students' attitudes toward science and interest in science careers, and to investigate the effect of inquiry-based instruction on students' attitudes and interest. In light of the projected need for a more scientifically literate workforce and the anticipated deficit in qualified graduates, it is imperative that ways to increase students’ interest in enrolling in science programs at the college level and in pursuing science-related careers be identified. Specifically, this paper will explore the effect of participation in authentic inquiry-based science projects on middle school students' attitudes towards science, including science self-efficacy and their interest in pursuing a future as a scientist.

This topic is important to all areas of education, including policymakers at the school, district, state, and national level. It is relevant to students, their parents, and their teachers, who are all invested in creating a promising, socially-productive future for our children. It is also vital to the business sector; in particular to science and technology-related business that are in existence today and which will be developed in the future. If there is a disconnect between the inquiry-based instruction as practiced in schools and students' rejection of science-related careers, then the future status of the United States as a world leader in science, technology and engineering is in peril. In this paper, the possibility of bridging this gap through the provision of authentic science experiences in the form of a school-based science fair is examined.

**PERTINENT AREAS OF RESEARCH**

Two pertinent areas of research to this problem are attitudes toward science, a construct which includes science self-efficacy and interest in scientific careers, and inquiry-based science instruction with an emphasis on authentic inquiry as defined by Chinn and Malhotra (2002), particularly in the form of a school-based science fair. Relevant subtopics briefly introduced in this paper include self-efficacy theory as described by Bandura (1977, 1984), gender-specific differences in science attitudes and interests, and teachers’ ability to implement and understanding of inquiry instruction in the science classroom.

**REVIEW OF THE RESEARCH LITERATURE**

In 1989, the American Association for the Advancement of Science's Project 2061 released *Science for All Americans*, in which they proposed that in order to create a scientifically literate population prepared to face the demands and meet the needs of the next century, the traditional approach to science curriculum be set aside and replaced with teaching and learning that are consistent with scientific inquiry. With Project 2061's *Benchmarks for Science Literacy* (1993) and the National Research Council’s *Science Education Standards* (1996), a new era of science education reform ensued. State science education standards, derived from the national standards, were written, and science teacher education programs shifted to focus on the development of inquiry-based pedagogy. This new science education reform movement is similar to the hands-on science movement that took place during the 1960s and 1970s (Shymansky, Hedges & Woodworth, 1990), but has important differences in the policy-based support established at the national and state levels.

However, it seems that the changes that have been made in science curricula, teaching, and learning have not necessarily translated into an increase in the number of young people deciding to pursue careers in science and engineering. It is possible that there may be a perceived gap between school science and real science. If this is true, then many of the goals of the inquiry movement may be backfiring.

This paper will look at the construct of students' attitudes toward science, the factors which contribute to these attitudes, and how attitudes toward science are related to interest in a career in science or engineering. The relationship of gender and ethnicity to attitudes and interest will also be explored, although not in great detail due to the narrow focus of this paper.

Next, the classification of inquiry instruction will be described, and the nature and categorization of science fair projects within the realm of inquiry will be discussed. Finally, the effects of involvement in inquiry activities on students’ attitudes and interest will be described, and the differential effects of participation in inquiry projects on males and females will be discussed.

**ATTITUDES TOWARD SCIENCE AND INTEREST IN SCIENCE CAREERS**

An attitude is defined as, “a complex mental state involving beliefs and feelings and values and dispositions to act in certain ways” (Attitude, 2006). Attitudes toward science in particular are defined by Osborne (2003) as, “the feelings, beliefs and values held about an object that may be the enterprise of science, school science, the impact of science on society or scientists themselves” (p. 1053). From his review of the literature, Osborne concludes that attitudes are influenced by four
main factors; gender, classroom and teacher attributes, curriculum, and the perceived difficulty of the subject. Haladyna, Olsen and Shaugnessy (1983) conclude that student perception about the importance of science is the most significant variable in the formation of student attitudes toward science. These findings are somewhat supported by Weisgram and Bigler (2006), who found in one part of their study that belief in the utility of science was a significant predictor of girls’ interest in science careers, but did not find a significant interaction in the second part of their study. Attitudes towards science are important, because they have been shown to influence students’ behaviors, including high school course selection and career choice (Atwater, Wiggins & Gardner, 1995).

A number of studies (Greenfield, 1997; Moffat, et al., 1992; Weinburh, 2000) have documented a deterioration of students’ attitudes toward science and science classes between elementary and high school. Girls in particular, seem to lose interest in science as they enter adolescence (Greenfield, 1997; Jones, Howe, & Rua, 2000; Reid & Skryabina, 2003; Weinburh, 1995, 2000). Many researchers have documented a difference in attitudes between males and females, with boys having a more positive attitude toward science than girls (Moffat, et al., 1992; Greenfield, 1997; Weinburh, 1995, 2000), although Baker (1985), in a much smaller study, found the opposite, with girls having a more positive attitude toward science.

The relationship between attitude toward science and achievement in science is moderate, according to Weinburh’s (1995) meta-analysis of research literature from 1970-1991. Overall, she found that a positive attitude toward science was moderately correlated with higher achievement in science. Osborne (2003) agrees, but proposes that the relationship is unclear; attitudes could influence achievement, or achievement could influence attitudes. He concludes that attitudes are linked somehow with achievement, but that this relationship still needs to be determined.

Boys and girls seem to exhibit different constructs of the relationships between attitudes toward science and achievement in science. Matter and Schau (2002) noted that girls’ attitude toward science tended to remain separate from their achievement in science. Neither seemed affected by the other, so that, even when girls’ improved their achievement in science, their attitudes toward the subject did not change, and vice-versa. Boys, on the other hand, seemed to link their attitudes about science with their achievement in science. When their attitudes improved, so did their achievement, and when their achievement increased, their attitudes toward science became more positive. In contrast, when Weinburh (1995) looked at the attitudes toward science of high-performing girls and compared them to the attitudes of high-performing boys, she found that the girls had a more positive attitude toward science. Lower-achieving girls, on the other hand, often maintained a positive attitude toward science, which reflects Matterand Schau’s finding that girls’ attitudes and achievement constructs tend to be kept separate. In contrast, Baker (1985)’s study of 120 middle school students found that both males and females with A and B grades had a less positive attitude toward science than males and females with C and D grades. These findings suggest that both attitudes and achievement must be addressed in classrooms in order to effectively reach children of both genders and of different abilities.

Attitudes toward science matter, because they are believed to influence behaviors such as course selection and career choice. Atwater, Wiggins and Gardner (1995) studied urban middle school students to find out how their attitudes toward science affected their course selection intentions. They found that students with a high positive attitude toward science reported the intention to take more than the required high school science courses. High positive attitude toward science was also associated with a high science self-concept, which is proposed to be analogous to the concept of science self-efficacy introduced by Bandura (1977, 1984).

Bandura (1994) describes perceived self-efficacy as, “people’s beliefs about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives” (p. 71). An individual’s self-efficacy comes from four sources: mastery of challenging tasks, vicarious experiences, verbal persuasion, and the lowering of anxiety associated with tasks. The most important of these factors for producing self-efficacy is the mastery of challenging tasks. Bandura points out that school is the natural place where children develop intellectual self-efficacy through their successful completion of academic challenges.

Each experience of mastery creates an opportunity for a child to gain positive self image and increase their feelings of efficacy.

Personal self-efficacy affects personal goal-setting behaviors, motivation, and selection processes, including course and career choices (Bandura, 1977, 1994). “Students’ belief in their capabilities to master academic activities affects their aspirations, their level of interest in academic activities, and their academic accomplishments” (1994, p. 79). If this is true, then students’ beliefs about their science abilities will affect their academic accomplishment in science, their involvement in
science-related activities, and the likelihood that they will pursue a career in science.

Science self-efficacy is defined by Tippins (1991) as a collection of beliefs about students’ capabilities in science, including their confidence about their abilities to complete science tasks, solve science problems, complete science class work, and in regards to their science technical skills. Tippins found that her construct of science self-efficacy to be only slightly associated with general academic self-efficacy, thus supporting the idea that individuals’ efficacy beliefs are domain-specific. This suggests that a student with a positive general academic self-efficacy may still have a negative domain-specific self-efficacy.

If self-efficacy affects aspirations, then science self-efficacy may have an effect on students’ attitudes toward science and their interest in pursuing a scientific career. Scott and Mallinckrodt (2005), in a small study of 41 former participants in a science summer enrichment program for girls during high school, found that the 23 young women majoring in science a couple of years after high school graduation, had higher measures of science self-efficacy than the eighteen subjects who were either pursuing other majors or who were not enrolled in college.

Weisgram and Bigler (2006) found that, for girls, positive science self-efficacy was significantly related to their interest in science. Navarro, Flores and Worthington (2007) had similar results from their study of the career goals of male and female Mexican American middle school students; they found that the students’ math/science self-efficacy had a significant effect on both their interest in mathematics and science and their future goals related to mathematics and science.

Science self-efficacy seems to affect students’ intentions to take future science courses. Tippins (1991) noted that, on average, ninth grade males had slightly higher measures of science self-efficacy and intended to take more science courses in the future than ninth grade females. However, Tippins found that together, gender and science self-efficacy only accounted for about 10% of the intentions to take future science courses. Tilleczek and Lewko (2001) found during a four-year study of Canadian high school students that the level of science coursework they participated in had a significant impact on whether or not they decided to pursue a career in science. Students who were enrolled in advanced and honors courses during high school were more likely to indicate interest in science and engineering careers. There seems to be a complex set of attitude, interest, decision-making, and intention constructs that contribute towards an individual’s choosing a career in the sciences. In addition, students’ perceptions of the type and number of courses needed to get into college may play a role in both intentions and actual course enrollment.

In a large longitudinal study involving 3,359 subjects who were in the eighth grade in 1988, it was found that about 50% of the students who identified the intention to pursue a science-related career actually achieved a baccalaureate degree (Tai, Liu, Maltese, & Fan, 2006).

This high correlation between eighth grade career aspirations and actual fulfillment of these intentions suggests that there are important factors that influence career choice which take place prior to young adolescence.

Hill, Pettus and Hedin (1990) identified seven factors that are associated with choosing a career in science; encouragement from teachers and counselors, participation in hobbies and activities related to science, students’ academic self-image, an interest in science-related careers, parental encouragement and support, students’ perceived relevance of science and mathematics, and students’ science and mathematics ability. Another major facet that they identified was personal contact with a scientist. Many of these contributing factors can be addressed in the school setting, and a number can potentially be affected through changes in classroom practices that encourage more students to consider science-related careers.

Woolnough (1996) investigated 654 seventh to eleventh grade students’ interest in science careers. He found that in grade seven, boys and girls started out with similar interest in science careers (about 42% said that they would be interested in a career in science). However, boys’ interest increased from seventh to eleventh grade, while girls’ interest dropped dramatically. The difference between males’ and females’ change in interest is in agreement with findings by other researchers (Greenfield, 1997; Jones, Howe, & Rua, 2000; Reid & Skryabina, 2003; Weinburgh, 1995, 2000), although a number of other researchers documented a concurrent drop in interest for both genders (Greenfield, 1997; Hedges & Woodworth, 1990; Moffat, et al., 1992; Shymansky et al., 1990; Weinburgh, 2000). Woolnough also asked students about factors that he thought might influence their interest in science careers. Student responses indicated that school science was not an influence on students’ desire to pursue scientific careers.

In fact, students revealed that science as done in school fails to motivate and inspire them. One reason cited was the lack of opportunity for open inquiry activities.

Yet another factor that may be influencing students’
The choice of science careers is their beliefs about what career fields are appropriate for each gender. Weisgram and Bigler (2006) found that girls more often express the belief that science is a male field, and Farenga and Joyce (1999) identified a strong stereotypical influence on boys’ and girls’ selection of courses for themselves and for the opposite gender. When given the opportunity to select science courses from a list, boys chose physics and robotics courses for themselves, but selected life science classes for a girl. Girls’ choices echoed those of the boys; they picked life science classes for themselves and chose physics, chemistry, and robotics courses for boys. These studies reveal that students may still harbor gender bias in their perceptions of what they believe to be appropriate science interests and science careers.

Ethnicity may also play a role in selection of science as a career. Weinburgh (2000) found that Black sixth through eighth grade students had less positive views of their science teacher, the utility of science, and less enjoyment of science. They also had lower self-concept in science. In contrast, Hill, Pettus and Hedin’s (1990) analysis of a study of middle school students’ science career choices revealed that Black students actually had a higher interest in science-related careers than did their White peers, which correlates with Mickelson’s (1990) findings that Black students often have more positive attitudes towards school in general. In this and related studies, the researchers found that the main factors predicting science as a career choice were gender and personal contact with a scientist. It is clear that further research needs to be done to explore the interaction of ethnicity, attitude factors and interest in science careers.

Involvement in science inquiry activities is one factor that has been shown to increase both students’ attitudes toward science and their interest in science-related careers. If there is a disconnect between students’ enjoyment of science as done in school, and their ability or willingness to envision themselves as a scientist, then perhaps the integration of authentic inquiry experiences into classrooms will bridge that gap. A discussion of the classification of inquiry activities as done in classrooms, and a review of the literature on the impact of participation in inquiry on students, form the next part of this paper.

**Science Inquiry in Schools**

In The National Science Education Standards, inquiry is described as,

The diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work. Inquiry also refers to the activities of students in which they develop knowledge and understanding of scientific ideas, as well as an understanding of how scientists study the natural world. (National Research Council, 1996, p. 23)

Inquiry, therefore, is the investigation of natural phenomenon in an authentic manner. Scientists and students involved in inquiry must make observations; ask testable questions; conduct background research and write reviews of literature; design investigations that include control of variables; carry out investigations where they gather, analyze and interpret data; draw conclusions, interpretations, and create explanations; and communicate results to peers and others. In the course of their education, students will take part in these inquiry activities separately and in combination, but “they also should develop the capacity to conduct complete inquiries” (p. 23).

The purpose of including inquiry in science classroom is twofold. First, students learn science content by conducting investigations that lead to greater understanding of scientific concepts, and second, students develop the skills needed to conduct scientific investigations (National Research Council, 2000). These are not mutually exclusive goals. Through participation in inquiry activities, students can learn both content and will gain inquiry skills. However, it has been demonstrated that teachers’ understanding of what constitutes inquiry is lacking in depth (Pekmez, Johnson & Gott, 2005; Trumbull, Scarano & Bonney, 2006), and inquiry in the classroom is often interpreted to be the inclusion of hands-on activities (Bencze, Bowen & Alsop, 2006; Tsai, 2006), so it is likely that student achievement is not meeting the goals for inquiry learning.

The inquiry skills that students are expected to acquire are included in the Standards as Content Standard A: Science as Inquiry Standards. The Standards state that all students from kindergarten through the end of high school should develop, “abilities necessary to do scientific inquiry” (p. 105), and “understanding about scientific inquiry” (p. 105). At the middle school level (grades 5 to 8), through scientific inquiry, students are expected to learn to: pose and refine questions that lead to scientific experiments; plan and carry out scientific investigations; select and use the appropriate tools (including computers and mathematical tools) needed to gather, display, analyze, and interpret data; describe their results and then use their data analysis to create interpretations of the results based on evidence; critique scientific investigations, understanding limitations and anomalies; identify and analyze alternative interpretations of
their results; communicate their research and their scientific conclusions; and use mathematics in all parts of scientific inquiry (National Research Council, 1996).

Furthermore, through participation in inquiry activities, students are expected to gain an understanding of what it means to do scientific inquiry. The Standards state that students in grades 5 to 8 are expected to know that there are different types of scientific investigations that arise from different kinds of scientific questions; that science investigations are influenced by historical and cultural context, as well as technological discoveries; that mathematics is essential in all types of scientific inquiry; that technology helps to improve the accuracy of measurements in scientific inquiry, and is used as a tool to analyze results; that scientific conclusions are based on evidence that has been analyzed in a logical manner; that skepticism is important in the advancement of science; and that scientific inquiry often leads to more questions, new ideas for methodology, or the development of new technology for use in data collection (National Research Council, 1996).

The Standards recommend that science classrooms be inquiry-based. According to the National Research Council (2000), there are five essential features of classroom inquiry. First, in a classroom that is inquiry-based, students are engaged by scientific questions. Scientific questions are those which are testable, or which can be explored through scientific investigations. Second, an inquiry-based classroom features an emphasis on the use of empirical evidence. Careful measurements, record-keeping and data analysis are valued. A third feature of an inquiry-based classroom is the logical use of that empirical evidence to draw conclusions and to answer scientific questions. Fourth, students in an inquiry-based classroom critique their explanations in the light of currently accepted scientific knowledge. Finally, in an inquiry-based classroom, students communicate what they have learned, present their conclusions, and participate in peer review and debate (National Research Council, 2000).

The definition of inquiry described in the National Science Education Standards (1996) is echoed in form but not in spirit in the Massachusetts Science and Technology/Engineering Curriculum Framework (Massachusetts Department of Education, 2006), which lists the inquiry expectations for sixth through eighth grade students. By the end of the eighth grade, Massachusetts students are expected to be able to “formulate a testable hypothesis, design and conduct an experiment specifying variables to be changed, controlled, and measured, select appropriate tools and technology (e.g., calculators, computers, thermometers, meter sticks, balances, graduated cylinders, and microscopes), and make quantitative observations, present and explain data and findings using multiple representations, including tables, graphs, mathematical and physical models, and demonstrations, draw conclusions based on data or evidence presented in tables or graphs, and make inferences based on patterns or trends in the data, communicate procedures and results using appropriate science and technology terminology, and offer explanations of procedures, (and) critique and revise them” (p. 11). It is not until the last two years of high school that the Massachusetts Frameworks propose that students undertake completely independent investigations. It is not clear why the Frameworks includes this limitation; students in much younger grades have been shown to be capable of conducting independent student-centered investigations (Abernathy & Vineyard, 2001; Blenis, 2000; Carlisle & Deeter, 1989; Hale, 1967; McDonough, 1995; Yaser & Baker, 2003).

How might young students be taught to conduct full independent investigations? One impediment is the complexity and level of thinking required for scientific inquiry processes. Anton Lawson (1995) identified two different levels of thinking patterns present in school-aged children; empirical-inductive and hypothetical-deductive. These two types correspond roughly to Piaget’s (1969) concrete operational and formal operational stages of cognitive development. Lawson’s levels can be directly applied to students’ ability to engage in particular levels of inquiry. A student who is at the empirical-deductive level of thinking is able to identify simple direct cause-effect relationships, to follow step-by-step directions, to change his or her point of view, to make an unsystematic identification of the variables involved in an experiment, to draw simple references, and to apply a variety of algorithms to solve problems (though not always correctly). They may have difficulty with metacognition. Students who employ hypothetical-deductive thinking can apply combinatorial thinking to both concrete and hypothetical relationships and situations, identify and deal with all needed variables and design experiments with controlled variables, engage in proportional reasoning, use probability to evaluate data and draw conclusions, and can recognize relationships and make comparisons between combinations of factors.

Lawson points out that students must be given opportunities to develop hypothetical-deductive thinking. “Experience helps students to construct structures that can ultimately lead them to think hypothetical-deductively about the world around them” (p. 85). He points out that in traditional classrooms, the teacher uses lab activities to demonstrate and confirm what has
been taught to the students, thus reinforcing students’ simpler empirical-inductive thinking skills rather than encouraging the development of more advanced thinking. He points out that, “If we intend to teach science as science is done—that is, in the way that people spontaneously construct knowledge—then the experiments must come first” (p. 85). In other words, students must have scaffolded opportunities to practice the skills characteristic of a hypothetical-deductive thinker.

One way to enable students to conduct independent science investigations is to design inquiry activities that carefully lead students from teacher-designed demonstration activities to true authentic inquiry activities that are student-centered and require them to stretch their thinking skills. Wenning (2005) reiterates that students must be taught inquiry skills and processes in order to close the gap between traditional activities which are teacher-centered and require a low level of inquiry skills, and more open inquiry activities, which are student-centered and require higher-level inquiry skills. Anders, et al. (2003) and Yerrick (2000) found that students who started out with lower attitudes and abilities in science were successfully engaged in open inquiry activities when the tasks were scaffolded to help them to adjust to the greater cognitive requirements. The teacher’s job is to structure this transition, and to scaffold students’ development of the ability to work on independent scientific investigations.

**Types of Inquiry**

According to the National Research Council (2000), inquiry activities vary by how many of the five features of classroom inquiry described earlier are included, and by whether the features are generated entirely by the students, or are proposed by the teacher. A full inquiry activity is defined as one which has all five of the essential features present. An activity missing one or more of the essential features is said to be partial inquiry. Activities can be further characterized by their degree of learner independence (from more to less) and by the amount of teacher or teacher-provided direction (from less to more). Activities that have the most learner independence and the least teacher direction are open inquiry activities, while those that have the least learner independence and the most teacher direction are guided inquiry activities. Although it is stated that, “more open inquiry will afford the best opportunities for cognitive development and scientific reasoning” (p. 30), the National Research Council points out that few students are ready for completely open inquiry before they have been helped to develop inquiry abilities through guided inquiry (2000).

Bell, Smetana, and Binns (2005) also characterize inquiry activities by the degree of teacher control versus student autonomy. They describe a four-level hierarchy of inquiry that depends on how much information is provided to the student by the teacher. The simplest level of inquiry, Level 1, is also called a confirmation activity, because students are given a question and a procedure with a previously known outcome. A Level 2, or structured inquiry, activity, involves student investigation of a teacher-supplied question, again with a teacher-supplied procedure, but in this case, the outcome is left open. In a Level 3, or guided, inquiry activity, students begin with a teacher-generated question, but design and carry out the procedures. Finally, a Level 4, or open, inquiry activity involves the investigation of student generated questions through student-designed procedures. Colburn (2000), however, does not agree that Bell, Smetana, and Binns’ Level 1 activities constitute inquiry. According to Colburn, inquiry-based instruction is “the creation of a classroom where students are engaged in essentially open-ended, student-centered, hands-on activities” (p. 42). Thus, Colburn’s definition of the highest level of inquiry approaches the inquiry done by practicing scientists.

Chinn and Malhotra (2002) propose that classroom inquiry can be assessed by evaluating how close the cognitive processes and epistemological assumptions of an activity are to those of authentic inquiry, or in other words, the work done by scientists. They divide inquiry tasks into two main categories; authentic and simple. Simple inquiry is further divided into three categories; simple experiments, simple observations and simple illustrations. In simple experiments, students do not select their own research questions or design their own methodology. Variables, analysis, and required cognitive effort are not complex, and much information, such as what and how to measure and which variables to control, are provided by the teacher. Simple observations and simple illustrations are not experimental at all, but instead are designed to provide descriptive and illustrative experiences.

On the other hand, the characteristics of authentic inquiry include the generation of research questions, the selection of variables, the design of procedures, the planning of controlled variables and measures of independent and dependent variables, skepticism, and recognition of observer bias, creative and complex reasoning, and the reading of scientific research reports (Chinn & Malhotra, 2003). According to these researchers, authentic scientific inquiry “refers to the research that scientists actually carry out” (p. 177), and is very dissimilar to most of the science done in schools. In fact, they
stress that students who participate only in simple inquiry tasks may develop an incorrect understanding of how science actually works, with the result that, “they are likely to reject scientific reasoning as irrelevant to any real-world decision making” (p. 214), revealing a lack of understanding of the actual epistemology of science. Could it be that lack of participation in authentic scientific experiences is what causes the previously noted chasm between students enjoyment of school science and their not being able to imagine themselves working as a scientist once they leave school? If this is the case, then schools may be creating students who enjoy the hands-on activities done in their science classes, but who have little understanding and appreciation for the science done outside of school. No wonder that more students do not express an interest in careers in the sciences; it may be outside of their personal realm of understanding.

Science Fairs
If the goal is to increase students’ interest in science and science careers, then perhaps they need to be involved with more authentic science inquiry. One way for students to become involved in authentic science inquiry is through participation in a school science fair. According to Bochinski (2004), a science fair project is an assignment that allows students to use their own ideas to investigate a scientific problem or question that they find interesting. In addition, it allows students to get “to know what it would be like to work in this field as an adult” (p. 5). “Science fairs allow students to see how science works outside the classroom—how scientists investigate and learn about the world in which we live” (Fredericks & Asimov, 1990, p. viii), thus perhaps creating that crucial connection between school science and science in our society.

Science fairs are meant to engage students on a personal level, and to provide them with an opportunity to learn about how scientists solve problems and gain knowledge about how the world works (Fredericks & Asimov, 1990; Van Cleave, 1997; Bochinski, 2005). Janice VanCleave (1997) sees involvement in a science fair project as an opportunity to solve a scientific mystery, with the student playing the role of a scientific detective. Indeed, many students who have been involved in a science fair report that they enjoyed their participation in the fair because it was fun, they enjoyed learning, they liked the competition, and they learned about scientific processes (Abernathy & Vineyard, 2001).

Although science fairs vary in their design, purpose, and effectiveness, most are based on having students follow the scientific method. The scientific method is, “a way to study a scientific problem in order to answer a proposed question or develop a better technique or final product through repeated tests and observations in a controlled environment” (Bochinski, 2005, p. 4). The scientific method is generally described to consist of the following steps:

1. Identification of a problem
2. Formation of a hypothesis
3. Background research and development of the experimental procedure
4. Experimentation and gathering of data
5. Data analysis and formulation of conclusions (Fredericks & Asimov, 1990; Lawson, 1995; Bochinski, 2005)

It should be noted, however, that scientific research does not have to adhere strictly to these steps in this order. Lawson (1995) points out that scientific research is often cyclical and involving much trial and error. During a science fair, for example, students may need to make adjustments to their research question and project design as they proceed through the steps of the scientific method.

A science fair project also adds the elements of presentation, communication, and demonstration of learning. Fredericks and Asimov (1990) describe a science fair project as, “a presentation of an experiment, a demonstration, a research effort, a collection of scientific items, or a display of scientific apparatus. It represents the efforts of a student’s investigation into some area of interest and provides a way for the student to demonstrate the results of those investigations” (p. 1). Clearly, participation in a science fair meets the requirements for engagement in inquiry as defined in the Standards (National Research Council, 1996).

Slisz (1989) used the available research to delineate the desirable characteristics of a school science fair. According to her summary, a science fair that successfully involves students in open inquiry should: have clearly stated goals; have expectations for projects that are adjusted to meet individual needs; allow both individual and cooperative (i.e. group or team) projects; should expect students to work on the projects outside of the classroom; and should require students to complete an experimental (rather than a demonstration or exploration) project.

Several researchers have noted that the most successful school science fairs emphasize the development of all students’ inquiry skills rather than competition (Slisz, 1989 and Balas, 1998). Balas (1998) points out that students’ choice of and success with complex science fair projects is related to their Piagetian (1969)
level of cognitive development. Teachers can help students to select projects that are appropriately challenging. O’Neill and Polman (2004) explain why it is important to teach students to do authentic science, pointing out that, “it is possible to guide student work just enough to maintain an equilibrium between challenge and opportunity” (p. 246).

Effects of Inquiry

It seems that participation in inquiry-based science activities and programs can have positive effects on students. Many studies have found that student involvement in inquiry-based classroom activities leads to improved or at least equal student learning as measured by standardized tests (Chang & Mao, 1998; Schneider, Krajcik, Marx, & Soloway, 2002; Trotter & Jones, 2003), by classroom assessments (Ertepinar & Geban, 1996; Freedman, 2002; Haussler & Hoffmann, 2002; Anders, et al., 2003; and Yager, Lim & Yager, 2006) and by student knowledge of inquiry skills (O’Neill & Polman, 2004). Shymanisky, Hedges and Woodworth (1990), in their metaanalysis of 81 studies of science reform curricula from the 1960s and 1970s, found that, overall, these curricula had a greater impact on student achievement than the traditional curricula of the same time period.

In addition to academic advantages, students who have participated in inquiry-based projects have more positive attitudes toward science (Shymanisky, Hedges & Woodworth, 1990; Fouts & Myers, 1992; Haussler & Hoffmann, 2002; Yager, Lim, & Yager, 2006), interest in science and science-related careers (Gibson & Chase, 2002), motivation towards and involvement in science learning (Trotter & Jones, 2003; Tuan, et al., 2005), and science self-efficacy (Yeck, 2000). Houtz (1995) found that 7th grade students showed an increase in positive attitudes toward science after the first year of instructional strategy change from traditional to inquiry-based, but 8th grade students showed no change.

Little research has been done on the impact of participation in science fairs on student achievement and attitudes. However, the research that has been done has had mixed results. According to Balas (1998), involvement in a school science fair increases students’ science self-efficacy, and helps them to develop an appreciation for the relevance of science in everyday life. Blenis (2000), on the other hand, found that low achieving fifth grade students who were mandated to participate in noncompetitive science fairs showed a significant increase in interest in science, but no significant change in attitudes toward science. Hale (1967) found that ninth graders who elected to participate in a science fair showed no difference in their interest in science careers when compared with their peers who did not participate. In one study, high achieving fifth graders were found to respond with a decrease in interest when involved in noncompetitive science fairs (Blenis, 2000), but in another study (Yaser & Baker, 2003), competition was found to have a negative effect on seventh grade students’ attitudes toward science. When he examined the responses students who opted to participate in their school science fair over multiple years, Hale (1967) found a positive relationship between the number of years that students participated in science fairs and their interest in science-related careers. Other factors, such as parental support for involvement in a science fair and liking of topic (McDonough, 1995) also play a role in students’ attitudes.

There is some evidence that participation in inquiry activities affects males and females differently. A small number of studies have been conducted on the effect of inquiry-based teaching on girls’ attitudes toward and interest in science. As with the studies of the effects of science fairs, the results are varied. Kelly (1999), in a small study of the effect of inquiry-based interventions in K-6 classrooms, found that both boys and girls maintained similar perceptions about whether boys or girls should be scientists or mathematicians. Shymansky’s (1990) metaanalysis identified a gender difference in the effect of inquiry curriculum’s effect on student perceptions of science. Boys’ perception seemed to become more positive, while girls’ perception seemed to remain unchanged. Freedman (2002) noted that involvement in laboratory experiences increased girls’ achievement in science, but did not find a difference in attitude between the laboratory group and the control group. In addition, gender seems to play a role in participation in a science fair, as Lawton and Bordens (1995) noted that 1.3 times more K-2 boys than girls participated, and 1.7 times more boys than girls in grades nine to twelve participated. The numbers were more equal for students in grades three to five and six through eighth, but it should be noted that it is more likely that science fair participation is mandated in these grades, and that the increased numbers of participants may reflect a required project rather than an equal amount of interest. These mixed results suggest the need for further research, perhaps with careful comparison of the effect of different types of inquiry activities.

Conclusion and Potential Research Questions

This paper explores the factors that contribute to students’ attitudes toward science and interest in scientific careers. The development of attitudes and interest
is complicated, and seems to be related to school science experiences, gender-related experiences and expectations, science self-efficacy, math and science ability and achievement, and other factors such as science opportunities and experiences that occur outside of school and teacher, parent, and role model effects.

Inquiry-based instruction is one possible intervention that shows potential in improving students’ attitudes toward science and interest in science-related careers. However, one problem that is apparent is that there is not a consistently-applied definition of inquiry instruction. The inquiry-based instruction that is described in studies includes the mere addition of laboratory activities (Ertepinar & Geban, 1996; Freedman, 2002), changing a single laboratory experience to make it open-ended (Anders, et al., 2003), implementing a short-term inquiry-based instructional program (Chang & Mao, 1998), involving students in a science summer camp or one day science enrichment program (Gibson & Chase, 2002; Weisgram & Bigler, 2006), increasing the hours of unspecified inquiry-based instruction (Houtz, 1995; Trettter & Jones, 2003; Tuan, et al., 2005), development of an interest-guided curriculum (Hauessler & Hoffmann, 2002), and a science, technology and society approach (Yager, Lim & Yager, 2006). A number of studies (Abernathy & Vineyard, 2001; Blenis, 2000; Goodnough & Cashion, 2006; Hale, 1967; McDonough, 1995; O’Neill & Polman, 2004; Schneider, et al., 2002; Yaser & Baker, 2003; Yerrick, 2000) describe what would be classified as open, Level 4 or authentic inquiry as defined by Colburn (2000), the National Research Council (2000), Chin and Malhotra (2002), and Bell, Smetana and Binns (2005). This lack of agreement on what constitutes inquiry instruction makes it difficult to interpret the actual effects of inquiry-based instruction on students’ achievement, attitudes and interest in science.

By definition, most science fair projects fit the definition of authentic inquiry as defined by Chinn & Malhotra (2002) because students are in charge of selecting a question for investigation, designing and conducting the experiment, analyzing results, drawing conclusions, and communicating their findings. Along with a high degree of student control, a science fair project represents a significant intellectual challenge. Through completion of a science fair project, students experience the work that practicing scientists do, create a deep understanding of their topic, and hone and improve the skills necessary for doing inquiry. Of course, students’ efforts to engage in authentic science experiences like a science fair project must be carefully supported, and inquiry skills must be taught in order to scaffold students’ baby steps from simple inquiry activities to student-centered investigations. However, if this is done effectively, a positive effect on students’ science self-efficacy, their attitudes toward science, and their interest in science-related careers may be seen.

According to Bandura, self-efficacy is most affected by an individual’s mastery of challenging tasks. By extension, mastery of challenging science tasks should have the ability to positively improve students’ science self-efficacy. A science fair project is a long-term project that involves higher-level thinking and integration of skills and knowledge from many different disciplines. Therefore, it can be seen as a challenging science task, and it follows that a student who is successful at completing a science fair project should experience an improvement in their science self-efficacy. In addition, the careful structuring of the science fair tasks, instruction in the inquiry process, and guidance of student project selection will also positively affect science self-efficacy by decreasing student anxiety associated with the project.

Completion of a science fair project allows students to experience authentic science, and to perhaps create a vision of themselves as a potential scientist. Through their concentrated work on a particular topic, students should gain an appreciation for the work that scientists do and the usefulness of science in our society. As Isaac Asimov writes, “This is the point of science fairs—to engage the interest of youngsters—to introduce them to the scientific method—to encourage them to understand science and possibly to become a scientist or engineer” (Fredericks & Asimov, 1990, p. vii).

This paper has explored the connections between student involvement in classroom inquiry activities, their attitudes toward science and interest in science careers. When the research is examined, it is clear that there is a gap where the effects of carefully defined authentic inquiry experiences need to be further explored. The following research question is proposed:

What is the effect of participation in an authentic inquiry experience such as a school-based science fair on middle school students’ attitudes toward and interest in science?

In addition, a number of subquestions that explore associated factors including gender, ethnicity, and socio-economic factors are put forth:

1. How do middle school students’ attitudes toward and interest in science before their participation in an authentic open inquiry activity such as a noncompetitive school-based science fair compare to their attitudes toward and interest in science after successful completion of the activity?
2. How are middle school students' specific attitude factors (science self-efficacy, enjoyment of science classes and perception of the utility of science), their intention to enroll in high school science courses and their interest in a science-related career affected by their successful completion of a science fair project?

3. How do the attitudes toward science and interest in science careers of middle school girls at the beginning and the end of participation in a school-based science fair compare to those of boys at the beginning and end of the project?

4. How do the attitudes toward and interest in science of middle school students of different ethnicities or socio-economic groups compare at the beginning and the end of successful participation in a school-based science fair?

5. How do other factors such as how well the student liked their project, their perception of the amount of difficulty they experienced with it, their perception of the amount of assistance they received, their perception of how successful their project was, and whether they worked independently or with a partner or partners affect their attitudes towards science and science careers?

REFERENCES


Involving Scientists in K–12 Education: Influencing Students’ Understanding of the Nature of Science
Michael J. Wadness

ABSTRACT
The following literature review addresses the problem of science literacy, specifically students’ understanding of the nature of science. The literature demonstrates that there is a need for explicit instruction of the nature of science within a context. One such context is through the participation of scientists in K–12 education. Although there is an abundance of literature describing how scientists have been involved in K–12 education, very few of the programs have been formally evaluated. Specifically, there is a significant gap in the literature related to whether or not scientists’ involvement in K–12 education as guest lecturers can affect students’ understanding of the nature of science.

INTRODUCTION TO THE PROBLEM
One of the major goals of K–12 science education is to create science literate citizens. The American Association for the Advancement of Science (AAAS), suggests that a science literate citizen is someone who understands major principles and concepts in science, understands the nature of science, recognizes the connections between science, mathematics, and technology, maintains a positive attitude towards being able to understand science, and is able to use scientific knowledge and habits of mind for solving personal and social purposes (1995). AAAS further delineates the reasons for cultivating a science literate population by suggesting that science provides the knowledge base to develop solutions for local and global problems, science creates a respect for nature, and science fosters the habits of mind needed in life when making informed, critical decisions. Most importantly, the future life-enhancing benefits of science and technology may never occur unless there is a new generation of scientists to carry on the scientific endeavor and there is a science literate general population, which understands the decisions that need to be made to support the scientists’ work (AAAS, 1989).

There are many scientific issues that have turned into controversial political issues. Cloning, stem cell research, pesticide use, global warming, and evolution are all examples of issues requiring an understanding of scientific evidence. Citizens have a responsibility to be able to understand and interpret the claims made by both scientists and politicians (National Research Council, 1996).

As one examines what it means to be science literate and its importance, the argument for educating the public on the science that tax dollars support becomes stronger. Science educators have argued that achieving the goal of science literacy among the populace is greatly aided by the teaching of the Nature of Science (NOS) throughout the K–12 curriculum (Smith & Scharmann, 1999).

Although there is disagreement among philosophers of science concerning NOS (Alters, 1997), some of the commonly agreed upon tenets include: the use of evidence, the stability and tentativeness of scientific knowledge, that science is a human enterprise with strengths and weaknesses, the use of imagination and creativity by scientists, that science is a complex social activity, that science is theory laden, and an understanding of the functions and relationships of models, laws, and theories (AAAS, 1989; Abd-El-Khalick & Akerson, 2004; NRC, 1996). Essentially, NOS addresses how scientific knowledge is developed, what scientific knowledge represents, and the recognition of science being a social, human enterprise.

Much research has focused on how to best help students understand NOS. One example is Irwin’s use of historical examples in high school science classrooms to illustrate to students how scientific knowledge develops (2000). Another example is Abd-El-Khalick and Akerson’s application of conceptual change theory to foster a more comprehensive view of NOS for preservice elementary school teachers (2004). In addition, both of these studies demonstrated that traditional science instruction that does not explicitly include NOS instruction does not result in students developing a comprehensive view of NOS, therefore suggesting that NOS is not implicitly learned through studying science content. Research has also demonstrated that for many teachers explicit instruction about NOS is a low priority due to their own lack of understanding of NOS, lack of confidence in their ability for NOS instruction, and a lack of available teaching materials and resources (Abd-El-Khalick, Bell, & Lederman, 1998). One solution is to utilize curriculum extension projects that do not require a significant amount of class time, but aid teachers in the explicit teaching of NOS.
During the last three years, one attempt to increase students’ understanding of NOS within high school physics courses has been through a program known as TheoryNet. TheoryNet, funded in part by the National Science Foundation, pairs high school physics teachers with local theoretical physicists for classroom visitations and guest lectures. Specifically, TheoryNet seeks to help students get answers to curiosity-based questions, see that scientists are real people, get an idea of what theoretical physics research is about, and to learn more about theoretical physics topics (Taylor, 2006). Not only does TheoryNet hope to improve students’ understanding of the nature of scientific work, it also hopes to improve students’ interest in the field of physics.

THE PROBLEM

The need for a science literate populace underlies the problem addressed by this paper. Specifically, the lack of understanding of the nature of science among high school students will be addressed. It will be suggested that by involving scientists in K–12 education, one outcome may be to influence students’ understanding of NOS; an integral aspect of science literacy.

This problem is of significance because, as earlier described, there are many social and political issues that demand a science literate populace which understands the meaning and impact of scientific claims. For example, if a local community is to decide how evolution is to be taught in schools, then a decision can only be reached if the public understands the difference between a scientific theory, a religious belief, and faith. It is not only important to teach students the outcomes of scientific inquiry, but also to teach students how the scientific knowledge was constructed, the nature of that knowledge, and that science is a human and social enterprise dependent upon imagination and creativity. If teachers are to include explicit instruction of NOS in curricula already overflowing with content, then there is a demand for projects that can be easily implemented without a significant impact to class-time.

PERTINENT AREAS OF RESEARCH

The areas of research pertinent to this problem are:

1. The Nature of Science (NOS) and Science Literacy
2. Research into Teaching and Learning about NOS
   a. Students’ view of NOS
   b. Teachers’ view of NOS
   c. Teachers’ attitudes towards NOS instruction
   d. Explicit instruction in NOS
3. Scientists’ Involvement in K–12 Education
   a. Summer research experiences for students
   b. Mentoring programs
   c. Working with classes for inquiry-guided learning
   d. Scientists involved as guest speakers.

For the purposes of this paper, the following two strands will be reviewed in depth:

Research into Teaching and Learning about the Nature of Science, and
Scientists’ Involvement in K–12 Education.

REVIEW OF RESEARCH LITERATURE

This section is divided into two parts, both of which examine the literature related to the problem of students’ understanding of NOS. The first part of this literature review examines research into teaching and learning about the Nature of Science (NOS). This section will address the research on: (a) students’ understanding of NOS, (b) teachers’ understanding of NOS, (c) teachers’ attitudes towards NOS instruction, and (d) the methods/approaches that have been attempted to improve students’ understanding of NOS.

The second part of this literature review examines scientists’ involvement in K–12 education. This section will review the literature describing the current programs, which involve scientists working in K–12 education.

Each area is discussed thoroughly to illustrate the important need for research related to the research problem addressed in this paper.

RESEARCH INTO TEACHING AND LEARNING ABOUT THE NATURE OF SCIENCE

A full treatment of the development of our contemporary understanding of the nature of science is beyond the scope of this paper. However, research by Alters (1997) illustrated that tremendous disagreement still exists among philosophers of science as to what the major tenets of NOS are. Alters surveyed half of the 418 membership of the Philosophy of Science Association to determine what the consensus position is for the basic tenets of NOS. The researcher’s instrument consisted of 15 frequently expressed tenets that were taken verbatim from the NOS literature. Out of the 15 tenets only six tenets had agreement greater than 75% and one tenet had disagreement greater than 75%. The remaining eight tenets showed no clear agreement or disagreement. Therefore it can be concluded that even among philosophers of science there is no consensus about the
nature of science. Samaraungvan, Westby, and Bodner (2006) have further suggested that scientists express a view of NOS that differs from philosophers.

Although there are disagreements among philosophers and scientists as to the basic tenets of NOS, this paper will draw upon those described in the writings of national professional science societies (AAAS, 1989; NRC, 1996). These commonly agreed upon tenets include: the use of evidence, the stability and tentativeness of scientific knowledge, that science is a human enterprise with strengths and weaknesses, the use of imagination and creativity by scientists, that science is a complex social activity, that science is theory laden, and an understanding of the functions and relationships of models, laws, and theories. Essentially, NOS addresses how scientific knowledge is developed, what scientific knowledge represents, and the recognition of science being a social, human enterprise (Abd-El-Khalick & Akerson, 2004).

Teaching about the nature of science was at one time considered unnecessary because it was assumed that students would implicitly understand the features of the domain through participation in lab-based courses; however, research on preservice science teachers, with backgrounds in studying science in lab-based courses have demonstrated insufficient understandings of NOS (Abd-El-Khalick 2005; Scharmann, Smith, James, & Jensen 2005). These research studies will be reviewed in detail in a later section of this paper. It has therefore been suggested by national professional science societies that NOS be taught explicitly in the classroom (AAAS, 1989; NRC, 1996). In the following section, research examining students’ understanding of NOS will be reviewed to identify why explicit methods of NOS instruction are needed.

Students’ Understanding of NOS

Research suggests that students’ views of NOS differ from the commonly accepted view. In a study by Samaraungvan, Westby, and Bodner (2006) it was determined through interviews, that high school students’ views of science were influenced by their own research experiences, which consisted mainly of non-inquiry based labs. The students did not recognize the role of anomalies as an avenue to advance knowledge, but rather viewed them as mistakes made by the investigator. The researchers also reported that students did not see a need to critically evaluate the development of scientific knowledge. This study only examined the ideas of 19 high school chemistry students in a Midwest high school and thus the generalizability is called into question. However, in support of this study’s description of students’ views of anomalous data in science is research by Sere et al. (2001). The researchers found that high school and university students in France and Spain also did not maintain a comprehensive view of NOS that criticizes the epistemological and ontological base of scientific data. In other words, the students did not subscribe to a view of NOS that questions how scientific knowledge is constructed and its relationship to reality.

More detailed accounts of students’ views of NOS were revealed in research by Tao (2003). Tao’s study involved a large sample size of 150 secondary grade one, high ability boys. Tao demonstrated that many students in Hong Kong held entrenched, inadequate views of NOS in which they portrayed the work of scientists as doing experiments to test explanations. However, few students believed scientists had an expectation of what might happen in advance of conducting the experiment. In other words, students believed scientists carry out experiments that result in chance, unexpected results. In addition, the students saw scientific knowledge as facts that are absolute.

In a study by Lederman (1999), 50 students were randomly selected from five science teachers of varying experience. The students were surveyed and interviewed to determine their understanding of NOS after a year of instruction that did not include explicit NOS instruction. All of the five teachers subscribed to comprehensive views of NOS and two of the five teachers taught inquiry lessons that the researcher claimed were consistent with ideas of NOS, such as demonstrating the tentative nature of scientific knowledge. After a year of instruction, Lederman reported that 46 of the 50 students lacked a comprehensive view of NOS. Lederman reported that the students did not understand the tentative nature of science and saw creativity, imagination, and subjectivity having only a minor role in the construction of scientific knowledge. This study suggests that students do not implicitly gain a comprehensive view of NOS through the instruction of science content or through inquiry labs. In addition, the study suggests that NOS is not implicitly conveyed through exposure to teachers’ who hold comprehensive views of NOS. In other words, NOS is not conveyed through osmosis as a result of being exposed to teachers whose view of NOS is comprehensive.

In stressing the importance of students understanding a comprehensive view of NOS, research by Sadler, Chambers, and Zeidler (2004) demonstrated, through the use of surveys and interviews of 84 high school biology students, that although 80% of the students did recognize that science uses evidence, almost half of the students did not have a complete understanding of how
science uses evidence. In addition, the researchers studied how students react to socio-scientific issues. In the study, the researchers presented students with scientific arguments that support global warming and scientific arguments that debunk global warming. The positions were both presented with data and arguments consistent with the contemporary debate. Although many of the students recognized the social, political, and economic influences on the issue, the majority of the students did not consider the scientific merit of the presented data in expressing their view of global warming. Therefore, this study suggests a need for explicit NOS instruction to aid students in making informed, critical decisions in regards to socio-scientific issues.

The above research indicates that students do not recognize the tentative nature of scientific knowledge, do not recognize the use of anomalous evidence, do not recognize the role of investigations to evaluate scientific claims, do not recognize that science is theory laden, and do not recognize the role of imagination and creativity. In addition, the above data were collected before any explicit instruction in NOS was conducted. Therefore, these studies also suggest that a comprehensive view of NOS is not implicitly conveyed through instruction of science content, laboratory investigations, or through exposure to a teacher with comprehensive view of NOS. Although more research in this area is needed, it can be suggested that there is a need for explicit methods of teaching NOS to students.

Teachers' Views of NOS

In the previous section, the reviewed research demonstrated students' views of NOS to be incomplete when compared with the tenets described by the AAAS and NRC. One reason for this may be that teachers fail to teach NOS or may even hold incomplete views of NOS themselves. In this section research concerning teachers' views will be addressed. Much of the research has focused on preservice science teachers' view of NOS rather than on the views of practicing teachers.

Abd-El-Khalick (2005) used a questionnaire and interviews to ascertain the views of 56 preservice science teachers from the Midwest. The researcher reported that his sample of preservice teachers, before explicit instruction of NOS, held views which included: theories become laws when proven true, scientific knowledge is not tentative, scientific knowledge represents an absolute truth, and a theory was simply a guess made by one scientist to describe an unknown phenomenon. The preservice teachers also did not understand that only indirect evidence can be used to support theories, and did not understand the informed view of the explanatory and predictive functions of theories for guiding research. In addition, only a minority of the preservice teachers recognized the roles played by inference, creativity, and imagination in NOS. The researcher continued by observing that only a minority of the preservice teachers appreciated the theory-laden nature of observations and investigations. The preservice teachers did not understand the role of prior knowledge, assumptions, theoretical commitments, and guiding frameworks in influencing scientists' interpretations of evidence. Although the preservice teachers did see science as universal, they dismissed the role of personal and social factors in the validation of knowledge. The preservice teachers viewed science as purely being about the collection of facts, in search for absolute knowledge, and that science does not involve speculation. It can be implied from this research that a comprehensive view of NOS is not developed implicitly by studying science. These findings also beg the question, if preservice science teachers have inadequate views of NOS themselves, how will they improve their understanding once they become science teachers? If new teachers do not develop a comprehensive view of NOS, then the problem of students' lack of understanding NOS will not be alleviated.

Abd-El-Khalick's findings were confirmed by research by Scharmann, Smith, James, and Jensen (2005). They found the majority of preservice science teachers in their sample, prior to explicit NOS instruction, also did not maintain views consistent with contemporary, accepted views of NOS. The researchers stated that many of the students expressed strong religious views and skepticism about evolutionary claims. The students did not understand the function of a theory and viewed ideas such as Intelligent Design as scientific.

An interesting finding related to teachers' holding comprehensive views of NOS was illustrated in a study by Benze, Bowen, and Alsop (2006). Their work suggested a correlation between teachers' views of NOS and inquiry orientation. The researchers found that the more social constructivist the teachers' view of NOS was, the greater the tendency to include student-directed, open-ended science inquiries. Although inquiry will not be addressed in this review it is interesting to note a potential relationship between teachers' views of NOS and their use of inquiry.

These studies demonstrate that preservice science teachers do not subscribe to a view of NOS consistent with that described by national professional science societies. As the majority of these preservice science teachers have backgrounds in science, it can be implied that merely engaging in the study of science does not
result in an understanding of NOS. In other words, an understanding of NOS does not arise implicitly from participation in science. Therefore it can be assumed that if preservice teachers gained their inadequate view of NOS while studying science, then high school students studying science under similar conditions will not fare any better. It can also be suggested that if teachers do not have a comprehensive view of NOS, then they will need assistance in explicitly instructing their students in NOS. Therefore explicit methods are needed for NOS instruction. Unfortunately research by Backhus and Thompson (2006) showed that most science teacher preparation programs do not address NOS. The researchers reported that less than one-third of the programs evaluated actually included a course focused on explicit instruction of NOS. The remaining programs however, did include NOS as a part of their methods course. The study suggests that if the majority of teacher education programs do not emphasize explicit instruction of NOS then it may never translate to explicit classroom practices. However, even if teachers do understand NOS and explicit methods are available, then what are their attitudes towards including explicit NOS instruction in the classroom?

Teachers’ Attitudes Towards NOS Instruction

To address the question of what teachers’ attitudes are towards NOS instruction, research by Abd-El-Khalick, Bell, and Lederman (1998) reported in their study of 14 preservice science teachers during their final semester of student teaching, that explicit teaching of NOS was not a priority. Interestingly, the researchers stated that the student teachers maintained views of NOS generally consistent with the contemporary view. In addition, all the student teachers stated that NOS was important to teach; yet almost nobody explicitly planned or taught NOS. The reason given for not including NOS explicitly was that the student teachers did not feel that NOS instruction was a priority compared to other classroom issues during the student teaching experience. Other reasons given were that the student teachers did not think students would be interested in NOS instruction, content was more important, a lack of confidence in teaching NOS, not enough NOS instructional material, not enough control over the curriculum to include NOS, and a lack of time. In addition, the student teachers also believed that they had implicitly included NOS by doing science; however, they never assessed their students’ views of NOS. The results of this study imply that even if preservice teachers enter the classroom with adequate views of NOS, there is no guarantee that NOS instruction will explicitly occur. In addition, this study suggests that there is a need to support teachers in the explicit instruction of NOS.

Further research by Abd-El-Khalick (2005) demonstrates that many preservice teachers who do not have extensive training in NOS do not have positive attitudes towards its instruction. Working with two groups of preservice secondary science teachers, Abd-El-Khalick exposed one group of 46 preservice science teachers to ideas about NOS during a methods course, while another group of ten studied NOS extensively in a philosophy of science course, in addition to their methods course. The methods course and the philosophy of science courses were both taught by the researcher. It was found that only one third of the preservice teachers in the methods group noted that NOS needed to be addressed in order to dispel myths and foster an appreciation towards science. In addition, approximately one third of the student teachers expressed hesitation with respect to NOS instruction. These student teachers believed their authority, as a science teacher would be compromised if science were presented as a messy, disorganized process. One fourth of the methods group stressed that NOS should be included, but felt that it would not prove interesting to students, or would be too difficult for students to understand, and were worried about the time needed to address NOS. In contrast, the preservice teachers who had studied NOS extensively in a one-semester philosophy of science course, contemplated changes in their future teaching to include assignments that incorporated NOS. This research suggests that preservice teachers who do not have extensive training in NOS may not be willing to promote its explicit teaching, while those preservice teachers who have more focused training in NOS may be more willing to explicitly include NOS in their teaching. Therefore, it can be suggested that in order to foster a more positive attitude towards the explicit instruction of NOS, teachers must maintain a more thorough understanding of NOS.

In contrast to Abd-El-Khalick’s findings, Lederman (1999) suggests that teachers’ understanding of NOS is not a significant factor in his/her willingness to explicitly include it in the curriculum. Lederman studied five high school biology teachers, of varying experience, who all subscribed to comprehensive views of NOS to determine the role in which it played in their teaching. Through the use of observations, formal and informal interviews, and analysis of lesson plans over a one-year period, it was determined that none of the teachers included explicit NOS instruction. Although, NOS was part of their school districts’ curriculum, the teachers did not feel that NOS instruction was a priority. The
teachers believed that it was more important to spend class-time teaching science content and that NOS was too abstract for their students. The researcher reported that the inexperienced teachers expressed frustration in regards to including NOS instruction but saw classroom management and content to be of greater concern.

Bartholomew, Osborne, and Ratcliffe (2004) also suggest that teachers’ understanding of NOS is not a significant factor in his/her willingness to explicitly include it in the curriculum. The researchers state the more significant variable in determining confidence and approach in NOS instruction relates to the teachers’ conception and understanding of learning goals. The researchers found that the use of teaching methods was a greater factor in determining teachers’ confidence towards NOS instruction than an understanding of NOS. It was observed that those teachers who engaged students in open discussions, used reflective discourse, and integrated science processes and content into relevant lessons, had a greater attitude towards NOS instruction. In addition, it was also observed that those teachers were more successful in teaching a comprehensive view of NOS. This research suggests that good pedagogy is more important than the teachers’ initial understanding of NOS. Therefore the variable that determines teachers’ attitudes towards explicitly teaching NOS may not be the teachers’ understanding of NOS but having explicit methods that the teachers clearly understand and a command of good teaching practices.

The above research suggests that although in some circumstances a teachers’ understanding of NOS affects their attitudes towards teaching NOS, the more significant variable is the teachers’ understanding of lesson outcomes and good pedagogy. One solution may be to make available resources for the explicit teaching of NOS that are firmly grounded in good pedagogy and that are easily implemented by teachers. If teachers are willing to explicitly include NOS instruction, then this may address the problem of students’ lack of understanding NOS. This leads to the question: what are the available methods for explicit NOS instruction? Explicit Instruction in NOS

As discussed earlier, research has demonstrated that many students and preservice teachers with backgrounds in science who have not had explicit instruction in NOS hold a view of NOS that can be considered far from adequate (Abd-El-Khalick, 2005; Samaraungavan, Westby, and Bodner, 2006; Scharmann, Smith, James, & Jensen, 2005; Sere et al., 2001; Tao, 2003). These studies therefore suggest that a comprehensive view of NOS may not occur implicitly through the instruction of science content. To test this hypothesis, Irwin (2000) used historical cases in teaching science to 14 year-old students in London, UK. Irwin taught two science classes consisting of relatively equivalent students. These two classes were identified as a control group and a treatment group. Both groups were given instruction on atoms and the periodic table in the same number of lessons. However the treatment group included NOS instruction using a historical context. After instruction, Irwin reported that there was no statistically significant difference in the mean scores for the two groups’ understanding of science content, but following focus group interviews, he determined that members of the treatment group used historical examples to demonstrate their understanding of many of the tenets of NOS, while the control group maintained an inadequate view of NOS. Although this study did not include a large, random sample of students it still demonstrates several points. These points include: an understanding of NOS was not implicitly conveyed through the instruction of science content in the control group; NOS instruction was explicitly conveyed in the treatment group by providing a historical context; because both groups received the same instruction time, NOS can be integrated within the historical context to teach content without a sacrifice to time; and the inclusion of explicit NOS instruction did not affect the students’ understanding of content compared to the control group. Therefore this study demonstrates an explicit method for addressing the problem of students’ understanding of NOS.

To further investigate explicit methods for NOS instruction, some researchers (Abd-El-Khalick & Akerson, 2004; Scharmann, Smith, James, & Jensen, 2005; Tsai, 2006) have drawn upon conceptual change theory (Posner, Kenneth, Strike, Hewson, & Gertzog, 1982). Although discussion of conceptual change theory will not be addressed in this literature review, the main points refer to students resisting accommodation of new concepts until there is significant dissatisfaction with their preconceptions. In addition, the new concepts must be considered intelligible, plausible, and fruitful. Conceptual change theory also addresses factors that affect students’ resistance to discarding preconceptions within their conceptual ecology, which includes prior experiences and epistemological views. Therefore, researchers suggest that students will only undergo conceptual change to accommodate a more comprehensive view of NOS if their original views are challenged within the context of explicitly designed instruction. Scharmann, Smith, James, and Jensen, (2005) reported in their study of undergraduate students, in their last
semester before student teaching, that when including explicit NOS instruction within a context and providing several opportunities for reflection and disequilibrium, students gained a more comprehensive view of NOS. The students were enrolled in a course that emphasized studying NOS. Throughout the course, students were given multiple opportunities to discuss issues in NOS within a context of historical examples. Although, the students were undergraduates it is possible that the results may be applied to K–12 education.

It is important to recognize the use of good pedagogy in NOS instruction. Tao's study (2003) used historical science stories to teach NOS without the use of disequilibrium and reflection. The researcher reported very little, if any, improvement between students’ pre and posttests. In fact, it was determined that many students used the stories to confirm and reinforce inadequate views of NOS. The students did engage in discourse, however, the co-construction of ideas did not foster conceptual change. This study suggests that students have entrenched ideas concerning NOS that need to be overcome through skilled pedagogy. Therefore, although the previously mentioned study by Irwin (2000) demonstrated successful NOS instruction through a historical context, Tao's study suggests that by just exposing students to stories without explicit instruction, students may not go through the process of conceptual change and the problem of students’ inadequate views of NOS will not be addressed.

Summary

Research indicates that both students and preservice teachers have inadequate views of NOS. In order for the problem of students’ lack of understanding of NOS to be addressed, teachers must not only understand NOS themselves but must also include the explicit instruction in NOS during their teaching of a science course. An example of explicit instruction includes the use of a historical context to demonstrate key aspects of NOS. However, research indicates that students’ naive views of NOS need to be challenged during classroom discourse. Students need to recognize the inadequacies of their conceptions of NOS and seek to replace them with more comprehensive views.

One way to foster conceptual change for NOS may come from introducing high school students to contemporary scientists and exposing them to their work. In a study by Abraham (2002), students worked alongside scientists in a summer internship. Abraham reported that students had a positive change in their views of science and scientists. The researcher claims that students saw scientists as real people and recognized science as a human endeavor as a result of their experience. The next section of this paper will review research concerning scientists’ involvement with K–12 education.

Scientists’ Involvement in K–12 Education

Students’ lack of an understanding of the nature of science is a problem in K–12 education, but one possible way to help students develop a comprehensive view of science and scientific work may come from introducing students to working scientists. In this section, the available literature concerning the involvement of scientists in K–12 education will be reviewed. It should be stated that research (Dolan, Soots, Lemaux, Rhee, & Reiser, 2004) suggests that in order for an outreach program to be successful, teachers must be involved in the development and implementation to ensure the goals of the program are to be met. Some scientists are unaware of the challenges in the teaching profession and have no training in pedagogical techniques; therefore teachers’ input is vital (Thompson, Collins, Metzgar, Joeston, & Sheperd, 2002). The research reviewed suggests that there are four variations in which scientists are involved in K–12 education. These include summer research experiences for students, mentor programs, classroom inquiry projects, and guest speakers.

Summer Research Experiences for Students

One of the more common avenues for scientists to be involved in K–12 education is through the use of summer research experiences. There are many variations of this method, but all involve a summer program in which students travel to a research institution. The students generally will live on campus while they conduct research. One example is the previously mentioned study by Abraham (2002). Abraham’s study was to determine the effects of research apprenticeships on 75 high school students who spent two to three weeks working in a lab alongside a scientist. Abraham used a pre- and post-questionnaire to determine the students’ views towards science. Abraham reported, as a result of the research experience, that students demonstrated an increased interest in pursuing science and that 52% of the students had a positive change in regards to their views of scientists as members of the scientific community. The students saw scientists as real people and recognized science as a human endeavor. Abraham stated that students initially expected the scientists to exhibit stereotypical behaviors and were surprised that, “…not all science fields are relegated to those with pocket protectors and bad haircuts” (p21). In addition, the students were surprised to discover that scientists could also enjoy literature, food, and foreign culture.
Although this research does not address every major tenet of NOS, it does suggest that students’ exposure to scientists can provide a significant amount of disequilibrium to promote conceptual change in students’ views of the human aspect of NOS. However, more research is needed to assess the impact of exposure to working scientists on the understanding of other major tenets of NOS.

Croft’s study (1999) also involved students working with scientists, during a summer program, only his interest was in measuring students’ understanding of the moral and ethical responsibilities of scientists. Thirty high school students attended a four-week summer program in meteorology at Jackson State University in Mississippi. The students interacted with research scientists in the field of meteorology through lectures, labs, and field trips. The goals of the program were summarized to include: the development of science skills, increased awareness of the interdisciplinary nature of meteorology, increased awareness of the role of a meteorologist as a researcher, teacher, and communicator, increased awareness of the opportunities for students to choose a career in meteorology, and increased awareness of the moral and ethical responsibilities of researchers in meteorology. The program was evaluated through the use of survey questions and journal entries. Croft reports that the students had gains in both knowledge content of meteorology as well as attitudes towards science. Croft claims that all of the goals of the program were met with the exception of increased awareness of moral and ethical responsibilities. Although Croft does describe data related to the gains in content knowledge, the author does not present detailed data relevant to the other goals of the program. Therefore, it is unknown what the students’ actual views were of the moral and ethical responsibilities. In addition, the claim of increased interest in meteorology can be called into question due to the fact that the students had to undergo a selective application process. Since the students’ initial interests were never measured, it is possible that the students may have already been interested in pursuing a career in meteorology. It is unfortunate that the evaluation did not thoroughly study the goals of this program related NOS. Croft concludes the evaluation by claiming the activities increased students’ understanding of NOS, however the researcher never presented data to support this statement. Therefore this program demonstrates the need for more rigorous research concerning how exposure to working scientists’ affects students’ understanding of NOS.

A last example of scientists interacting with students through a summer research experience is a two-week summer program at Stockholm University described by Johansson, Nilsson, Engstedt, and Sandqvist (2001). In this program, the students were assigned a mentor who introduced a research topic in astronomy or particle physics and aided the student in making a research plan, carrying out research, and writing a report. The authors claimed that as a result of student participation in the program, students developed new insights into the scientific method as a result of the interaction with the scientists. In regards to NOS, the authors claimed that students learned it was more important to understand the methods of data collection and the uncertainty in the measurements, than what the right answer is. Students were encouraged to take initiatives and to question established facts. Unfortunately, the authors of this paper never referred to any formal evaluation to support these claims, thus questioning the validity of the study.

The three programs described above refer to summer research experiences for students to interact with scientists. Although the research does suggest a shift in students’ perceptions of the human aspect of science, there is very little data to support any of the other claims regarding NOS. As a result of these studies, a need presents itself for formal research regarding students’ understanding of NOS. In addition, the above programs also involved an application process, which introduces the threat of self-selection and it was never determined if these students initially subscribed to a comprehensive view of NOS before beginning the programs. If the improvement of NOS did occur, was it due to the research experience or due to the interaction with the scientists as a mentor? The next section reviews the research concerning scientists involved as mentors.

Mentor Programs

Mentor programs are another example of how scientists are involved in K–12 education. Waltner (1992) describes a mentor program in which students, with a grade of “B” or better, are paired with local scientists. The mentors and the students meet throughout the year, both formally and informally. In formal meetings, the pairs participate in activities that include problem solving, programming and trouble shooting computer networks, preparing and staining specimens for microscopes, and investigating variables. Waltner claims that the relationships had continued well after the program had ended and had also led to summer jobs and college internships. Although the goals of the program were not evaluated, the program claimed to improve students’ science and technical literacy, demonstrate real applica-
tions of science, math, and communication skills, highlight the critical need for skillful problem solvers, and help students make informed college and career choices.

Although a mentor program can be a useful program for many students, it presents a problem in rural areas where there are few working scientists available to act as mentors. O’Neill, Wagner, and Gomez (2006) present the solution to this problem with online mentors. The authors describe a mentor program in which graduate students are paired with middle and high school students to interact through email and other computer conferencing methods to help students work on long-term science projects. The goal of program is to develop “an audience of scientists who can offer students advice and criticism on an ongoing basis” (p. 39). Although, there was no formal evaluation of the program, the authors claimed that students came closer to realizing their potential, that the stereotypes of working scientists were dispelled, and that the graduate students provided an intellectual resource for both teachers and students. This article is of interest; because if the claims were to be supported, then it is possible that the exposure to working scientists may help students understand the human aspect of the nature of science.

Since it is unclear if students’ exposure to working scientists through mentor programs is the determining factor in affecting students’ understanding of NOS, there is a need to evaluate if it is indeed the research experience that affects students’ understanding of NOS. Thus, the next section describes the literature involving scientists working with K–12 science classes to help promote inquiry learning.

Working with Classes for Inquiry-Guided Learning

A third example of scientists’ involvement with K–12 education is through aiding teachers in providing inquiry-guided learning. The first type of inquiry program focuses on students collaborating with scientists in research topics, which fall outside the traditional curriculum. Abbott and Swanson (2006) describe a program in Alaska in which students collect data on Alaskan lake-ice and interact with scientists in the analysis of the data. In this program, the scientist visits the classroom and instructs students on how to collect data and explains the value of the study. The goal of the program is to connect classroom science with authentic applications. The authors claim that as a result of the program, students will recognize the relevance of science, develop a sense of pride, and become more excited about science. Unfortunately, the authors of this paper do not describe any evaluation of the program. Therefore it is unknown if the exposure to working scientists through inquiry had any effect on the students’ attitudes towards science or on their understanding of NOS.

In a more structured form of inquiry, which involves students collecting and analyzing data with scientists, is a program described by Stork, Okuda, and Cowan (1999). The program involves middle school, high school, and university students collaborating with NASA scientists and engineers to design a study, collect images, and analyze data collected by a digital camera mounted on the space shuttle. The images of the Earth were to be incorporated into an inquiry lesson for learning science content. The authors evaluated the program to determine the students’ prior knowledge, the knowledge gained, the reactions, opinions and views of the learners, how the program was implemented, and the associated costs of the program. Data were collected through the use of journals, questionnaires, observations, and formal assessment for content objectives. The authors claim that the students did have gains in content knowledge and mathematics, however the control group had similar gains. The authors claim that as a result of the program, the students did have growth in their motivation and interest in Earth science and space exploration. Unfortunately, the authors never addressed the students’ views of NOS. Therefore, it is unknown how an inquiry project affects students’ understanding of NOS.

In a paper that does address students’ understanding of the human aspect of NOS, Sabo, Sarquis, and Ennis (1997), describe a program in which students are to solve an engineering problem involving industrial plastics. The goal of the program was to improve chemistry education through curriculum development and outreach experiences that would allow students to interact with scientists and engineers in industry. The program involves a scientist coming to the school to present two one-hour seminars in which the students were encouraged to ask questions. The seminars were intended to present the problem and provide the necessary background knowledge. The students were given a tour of an industrial laboratory, conducted lab work at school, and brought samples to local universities for analysis before writing their research reports. Although no formal methods of evaluation were described, the authors claimed that students’ preconceptions of scientists changed as a result of the program, especially due to the tour of the laboratory where the students were able to interact with scientists. Students were surprised to see scientists as real people who work for a living. The article described that students’ expectations were to see the scientists as “a bunch of old, white-haired scientists standing around looking at the periodic chart all day.
long” (p. 450). The importance of this article is that if the claims of the authors can be supported, then exposure to scientists may affect students’ understanding of the human aspect of NOS.

In contrast to a researcher having classes collect data in the field, the second type of inquiry program involves content, which falls within the traditional curriculum. Siegel, Mlynarczyk-Evans, Brenner, and Nielsen (2005) describe a program in which scientists can interact with students in more traditional experiences. The authors describe a program in which graduate students in biology, are partnered with a high school biology teacher to design four interactive inquiry lessons for advanced biology students. The graduate students were in the classroom engaging the students while they were collecting and interpreting data. Although no description of evaluation methods or analysis were presented, the authors claimed that the students benefited by gaining an increased ability for integrating authentic inquiry, that students gained a better understanding of how to conduct a scientific investigation, and that students gained a better understanding of NOS. Specifically, the authors claimed that students saw scientists as real people, recognized that scientists employ a variety of approaches in investigations, and viewed science as a creative process. Therefore it is possible that the students’ exposure to scientists may have affected students understanding of NOS.

A final example of involving scientists with K–12 education through inquiry experiences is the RE-SEED program (Zahopotos, 2003). The RE-SEED program brings retired scientists and engineers into classrooms to work with students and teachers to conduct inquiry-based activities. The goal of the program is to increase science literacy by improving science education in public schools. The scientists and engineers volunteered their time and underwent 40 hours of training before entering the classroom. Zahopotos reported that in an evaluation, not performed by the author, 82% of the 849 middle school students had an increased interest in science. Zahopotos also reports that 86% of the students wanted the volunteer to return because he/she made learning fun and made science easier to understand. However the 14% that did not want the volunteer to return, stated that the volunteer was too hard to understand or was boring. Zahopotos does not report if students understanding of NOS was ever assessed.

It is interesting to note that the RE-SEED program, like many of the other inquiry programs described in this review, had the goal of science literacy yet did not explicitly evaluate NOS. It can also be questioned, if the scientists need to perform inquiry lessons with students to increase their understanding of NOS? The next section examines programs in which scientists come to the classroom and act as guest speakers.

**Scientists Involved as Guest Speakers**

Unfortunately, not all students can attend summer research programs and many schools may not have access to programs that involve individual mentorship or have access to programs that have scientists conducting inquiry lessons. Another option that does not demand a large time commitment by teachers or by scientists is a program known as TheoryNet (Taylor, 2006). In this program, theoretical physicists are paired with high school teachers to visit schools and guest lecture. The goals of the program are to provide students with answers to curiosity-based questions, supplement and enhance the current subjects being studied, develop an understanding of what theoretical physics is about, and to dispel the myths of what a scientist is and how they research science. At the present time, this program has not been evaluated.

Due to the need for explicit methods for the instruction of NOS, the TheoryNet program may provide an avenue for the instruction of NOS if its claims can be supported. It is possible that students’ understanding of NOS, such as the human, social, creative, and imaginative tenets, could be improved if they were explicitly instructed through the context of a TheoryNet presenter discussing his/her own experiences in science research. This is an area that lacks data and therefore demonstrates a significant need for research.

**Summary**

The articles reviewed above describe programs in which scientists are involved in K–12 education through summer programs, mentoring programs, inquiry activities, and as guest lecturers. Unfortunately, very few of the above articles included formal evaluations, which could determine the validity of their claims. The above articles imply that the involvement of scientists in K–12 education may affect students’ understanding of NOS, however there is a demand to verify these claims with formal research.

**CONCLUSION AND POTENTIAL RESEARCH QUESTIONS**

Individuals in our society are called upon to make critical decisions when science overlaps with social policy. To participate effectively in such decision-making requires an understanding of how scientific knowledge is developed and accepted. In other words, individuals need to understand the nature of science and this
knowledge is most appropriately developed during the course of K–12 science instruction.

The research reviewed has shown that students do not subscribe to an understanding of NOS that is consistent with the tenets suggested by national professional science societies. Unfortunately, the research indicates that the problem may not be solved easily due to inadequate views of NOS held by preservice and inservice teachers. Only one-third of science teacher preparation programs include courses specifically devoted to instructing preservice science teachers on NOS and on methods for NOS instruction. In addition, research indicates that even if preservice and inservice teachers understand NOS, their attitude towards NOS instruction may be less than favorable due to a lack of confidence, time, and available resources for teaching it. The research reviewed indicates that successful NOS instruction occurs within a context through explicit methods firmly grounded in good pedagogy.

One aspect of understanding NOS is recognizing that science is performed by curious human beings who engage in careful exploration of ideas and communicate their findings to the professional community. The research reviewed in this paper has indicated many means and outcomes of the involvement of scientists with K–12 students, but few studies have examined whether students can develop an understanding of some or all aspects of NOS through exposure to scientists and their work.

Consequently, research into determining if scientists’ involvement in K–12 education affects students’ understanding of NOS is warranted. Specifically, the area of scientists as guest speakers in the classroom needs to be researched, as this is a popular method of linking the world of science to the lives of students. The TheoryNet program involves theoretical physicists visiting high school physics classrooms and this program will form the focus of this study.

The following research questions are suggested:

1. How do theoretical physicists from TheoryNet view NOS?
2. Which aspects of NOS, if any, do TheoryNet presenters make explicit while presenting to high school students?
3. Do theoretical physicists from the TheoryNet program speaking to high school physics classes affect students’ understanding of NOS?
4. Do theoretical physicists from the TheoryNet program speaking to high school physics classes affect teachers’ understanding of NOS?
5. Do theoretical physicists from the TheoryNet program influence students’ views of scientists and science as a career?

Researching these questions will provide insight into the role of guest scientists in helping to develop students’ understanding of the nature of science.

REFERENCES


Use of Portable Technology in a Teacher Preparation Program for the Purpose of Enhancing Teachers’ Understanding of Scientific Inquiry

Sachiko Tosa

ABSTRACT

In spite of the strong emphasis on teaching science as inquiry in the science education community, very little inquiry-based teaching is taking place in actual K–12 science classrooms. One of the reasons for this deficiency may be ascribed to the lack of experience that teachers have with authentic scientific inquiry in the course of their own education. This paper addresses a need for providing beginning teachers with authentic inquiry experiences as well as subsequent instruction that links their experiences to teaching situations. It is argued that teacher education programs should include these two components so that teachers may better develop the knowledge and skills necessary for teaching science as inquiry. Two areas of research literature are reviewed: scientific inquiry and the use of technology to enhance teachers’ understanding of scientific inquiry. Through the literature review, this paper develops a rationale for examining the effectiveness of a teacher preparation program that makes use of a programmable data-collecting device as a tool to enhance teachers’ understanding of scientific inquiry.

INTRODUCTION TO THE PROBLEM

The National Science Education Standards (NSES) (National Research Council [NRC], 1996) emphasize clearly the importance of scientific inquiry in K–12 science education. The overview to the standards states that “inquiry is central to science learning” (p. 2). The NSES envision inquiry-based science classrooms in which students actively develop their understanding of science by combining scientific knowledge with reasoning and thinking skills. Based on the idea that students’ learning is greatly influenced by how they are taught, the NSES also describe in detail strategies for inquiry-based teaching. The NSES call for dramatic changes in science education in school systems, and advocate inquiry-based learning and teaching for students and teachers in the United States.

Professional science societies and the science education research community have been responding to this call. The National Research Council (2000) describes examples of inquiry-based science practices that are to be found in K–12 classrooms. A number of inquiry-based curricula and projects have been developed for elementary and secondary science (Chinn & Malhotra, 2000). Research studies examining various aspects of scientific inquiry in science education have been conducted to help teachers and educators to envision what scientific inquiry is and how it can be incorporated into contemporary educational contexts (Flick & Lederman, 2004; Minstrell & van Zee, 2000).

In spite of the emphasis on scientific inquiry in the science education research community, it has been pointed out that very few inquiry-based lessons are taking place in actual science classrooms (Lotter, 2004; Reiff, 2002; Simmons et al., 1999). Teachers are often central figures who transmit definitions and facts in a didactic manner. Students are typically busy looking for ‘the right answers’ by checking their textbooks (Roehrig & Luft, 2004). Laboratory work is often described in a cookbook style, where every step of the experiment is given and students are expected to follow precisely what they read or are told (McComas, 2005; Singer, Hilton, & Schwierengeruber, 2006). It appears that teachers are not translating the guidelines for inquiry-based teaching in the NSES into their own classroom practice. If teachers are not teaching science concepts through inquiry, then students have little opportunity to experience learning science through inquiry in their science classrooms.

Scientific inquiry is regarded as the heart of science education because inquiry helps students to develop deeper understanding of scientific concepts (NRC, 1996). Students construct their own understanding of scientific ideas when they make sense of what they are learning (Bybee, 2000; Krajcik, Blumenfeld, Marx, & Soloway, 2000). Through the active process of inquiry, students are expected to develop “the knowledge and understanding of scientific concepts and processes required for personal decision making, participation in civic and cultural affairs, and economic productivity” (NRC, 1996, p. 22). However, if very little inquiry-based teaching is taking place, it seems self-evident that the science classroom will not be a place for students to develop understanding of scientific concepts through their own meaning-making processes.

One of the reasons for the lack of inquiry-based teaching in science classrooms may be ascribed to the limited experience that teachers have had during their own science education. Even while studying for undergraduate degrees in science, science teachers in general have little chance to engage in open-ended explorations.
College introductory courses in science departments are typically loaded with routine attendance to lectures and problem-solving exercises. Laboratory activities that are required in undergraduate courses are mostly focused on the verification of stated notions with sets of given materials (McDemott & DeWater, 2000). While teachers who hold master's degrees in science have probably engaged in autonomous inquiries, teachers who have master's degrees in science education normally do not experience scientific investigations in their education courses (McDemott & DeWater, 2000). As a consequence, many teachers have not been given opportunities for experiencing authentic scientific inquiry.

It should not be thought, however, that inquiry experiences instantly make a science teacher competent in teaching inquiry-based lessons. If they do, any scientist could successfully conduct inquiry-based teaching. When science teachers are given opportunities to engage in authentic inquiry, what they need next is to transform their inquiry experiences into forms that are useful for student learning. This transformed knowledge is called pedagogical content knowledge (PCK) (Schulman, 1986), and is a crucial component in teaching (Gess-Newsome, 1999). Teachers need to reflect on how their understanding of a particular topic has evolved and try to identify what are the critical issues that need to be addressed with students for meaningful learning to occur (McDemott & DeWater, 2000).

A number of programs have been developed to provide in-service and pre-service teachers with authentic scientific inquiry experiences (Crawford, Zembal-Saul, Munford, & Friedrichsen, 2004; NRC, 2000). These programs mainly aim to develop teachers’ abilities and understanding of scientific inquiry. They often incorporate features such as field work in science settings at excavation sites or science museums, or collaboration with university faculty in research laboratories (NRC, 2000). More recently, the use of various technologies to enhance teachers’ understanding of the nature of scientific inquiry is being explored.

The potential for developing scientific knowledge through the use of technology was first explored by Seymour Papert (1980). Through his observations of students' activities, Papert realized that students change their way of thinking when they engage in computer programming to explore their own questions. He argued that students deepen their understanding of scientific concepts when they actually design and construct their own projects (Papert, 1993). Papert's work has evolved and joined into a broader research field in which various technologies, such as hand-held computers with sensitive probes, modeling, or World Wide Web, are used for educational purposes (Novak & Krajick, 2004). One of the technological products developed in Papert's group is a device called Handy Cricket (Martin, Mikhak, & Silverman, 2000). The Cricket is a fully programmable portable device that can be used in robotics by attaching motors, lights, and sensors. Resnick, Berg, and Eisenberg (2000) initiated the use of Cricket in order to impact students' understanding of the nature of scientific inquiry. Crickets can also be used as a tool to collect data for testing various scientific ideas by attaching different sensors.

Martin and Greenwood (2007) embedded the Cricket technology in a teacher preparation program. The researchers conducted a preliminary study on the effect of beginning teachers' awareness of the nature of scientific inquiry through the exploration of their own scientific questions with the Handy Cricket. Their results indicated a significant impact on teachers’ awareness of the nature of scientific inquiry. When the teachers actually experienced generating questions, collecting data, and interpreting the data, they became more cognizant of the fact that science is not a simple process and involves various aspects that are not normally described in textbooks. For example, the teachers learned that they themselves had to look carefully at their data points in order to make sense of what happened in their experiments.

Unfortunately, Martin and Greenwood (2007) did not document the change in teachers’ perceptions in detail and no data are available to investigate how and in what aspects the teachers’ views toward scientific inquiry have been impacted. Moreover, their teacher preparation program did not include a component that helps teachers transform their inquiry experience into forms that are useful for teaching science through inquiry. Therefore, the focus of the present study will be on the change of teachers’ understanding of the nature of scientific inquiry when they go through a teacher preparation program that provides them with authentic inquiry experience with the Cricket as well as instructions that help them develop their PCK for planning inquiry-based lessons.

THE PROBLEM

The problem this paper will address is the continuing lack of emphasis on scientific inquiry in many science classrooms. One factor contributing to this may be that K-12 science teachers have not themselves experienced scientific inquiry. It is suggested that teachers need to engage in an authentic inquiry experience and
transmform their experience into forms that are useful for inquiry-based teaching if they are to better understand the nature of scientific inquiry advocated in the National Science Education Standards (NRC, 1996). Specifically, the paper focuses on how effective the use of a programmable data-collecting device called the Handy Cricket will be for the purpose of enhancing preservice and novice teachers’ understanding of scientific inquiry in a teacher preparation program. Particular attention is placed on the need to integrate an authentic inquiry experience facilitated by the use of new technology, with instructions that aim to help teachers transform their inquiry experiences into appropriate PCK for inquiry-based teaching. This problem is significant for K–12 science education because it may suggest a new direction for developing teachers’ understanding of inquiry during their teacher preparation program.

Pertinent Areas of Research

The areas of research pertinent to this problem are:

• Scientific inquiry
• The relationship between teachers’ subject matter knowledge, pedagogical content knowledge and their understanding of the nature of scientific inquiry
• Constructionism and the use of technology to enhance understanding of scientific inquiry

For the purpose of this paper the following two strands will be reviewed in depth:

• Scientific inquiry
• Use of technology to enhance understanding of scientific inquiry

Review of Research Literature

Two areas of research literature are used to develop a rationale for a proposed study. The first part of the literature review is devoted to the examination of research studies in the area of scientific inquiry. Three perspectives will be discussed: (a) meanings of scientific inquiry, (b) importance of scientific inquiry, and (c) practices of scientific inquiry. This section will conclude by examining research studies that suggest that teachers who have not had both experiences with authentic inquiry in their own course of education and appropriate PCK for teaching inquiry-based lessons may have difficulty in teaching scientific concepts through inquiry.

Providing teachers with opportunities for inquiry experiences together with opportunities for developing appropriate PCK is a logical solution for the above problem. PCK is commonly recognized as a synthesis of subject matter knowledge and general pedagogical knowledge (Gess-Newsome, 1999). Inquiry-based science teaching, in particular, requires that teachers integrate multiple knowledge domains actively for planning and conducting effective lessons (NRC, 1996). It is therefore necessary for researchers to examine how different knowledge domains are integrated in order to develop appropriate PCK for inquiry-based teaching. The discussion of the relationships between the knowledge domains necessary for teachers’ understanding of scientific inquiry is an important issue for the present problem. However, because of the limited format of this paper, it is not discussed here.

The second area of the literature review examines research that discusses the use of technology in developing teachers’ and students’ understanding of the nature of inquiry. In particular, it focuses on: (a) the effectiveness of the use of technology in developing understanding of nature of scientific inquiry, and (b) practices that have been utilized in education programs for science teachers. The term ‘constructionism’ is introduced as a philosophical framework in this research area.

Each area of research is discussed thoroughly to illustrate the importance of the proposed study and the need for research related to the research problem addressed in this paper.

Scientific Inquiry

Definitions

The National Science Education Standards (NRC, 1996) use the term ‘inquiry’ in two ways: (i) to describe students’ activities by which they come to understand science, and (ii) to depict the teaching strategies associated with such activities (Bybee, 2000). The two definitions of inquiry are discussed separately.

Inquiry as student activity

The NSCS define the term inquiry as the following:

Scientific inquiry refers to the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work. Inquiry also refers to the activities of students in which they develop knowledge and understanding of scientific ideas, as well as an understanding of how scientists study the natural world (NRC, 1996, p. 23).
This definition makes it clear that scientific inquiry is the scientists’ ways of knowing the natural world. Scientific inquiry in science education also means activities of students. Students employ scientific inquiry in order to develop their understanding of scientific concepts in science classrooms. To distinguish the scientists’ ways of knowing from the activities of students, the former is often referred as ‘authentic scientific inquiry’ (Chinn & Malhotra, 2000). This use of the term is adopted in this paper.

The content standards of the NSES (NRC, 1996) describe scientific inquiry in terms of two components: abilities necessary to do scientific inquiry and understanding about scientific inquiry. Students should develop these abilities and understanding as the result of their activities during K–12 science education. At the end of the 12th grade, students should be able to conduct complete scientific inquiries by identifying questions, designing and conducting their investigations, using technology and mathematics, formulating their explanations, recognizing alternative explanations, and communicating their findings. It should be noted that abilities of scientific inquiry include not only process skills of doing tasks materially, but also cognitive skills of critical thinking and reasoning (Bybee, 2000).

As for developing students’ understanding about the nature of scientific inquiry, the NSES (NRC, 1996) describe that by the end of the 12th grade, students should know that conceptual principles guide scientific inquiries, that scientists conduct investigations for a variety of reasons, that technology and mathematics are essential in scientific inquiry, that scientific explanations must be supported by logic and evidence, and that the results of scientific inquiry emerge from different types of investigations. In short, students should develop knowledge about authentic scientific inquiry as a result of their K–12 science education.

**Inquiry as teaching strategies**

One of the assumptions on which the teaching standards of the NSES (NRC, 1996) based is that “what students learn is greatly influenced by how they are taught” (p. 28). Therefore, the NSES provide a comprehensive perspective for science teachers who actually teach students science lessons. The teaching standards of the NSES describe guidelines for teachers to follow in order to implement effective inquiry-based lessons. For example, Teaching Standard A describes how teachers should plan inquiry-based lessons, while Teaching Standard B describes how to guide and facilitate students through their inquiry processes.

It is important to note that the NSES (NRC, 1996) does not recommend a single approach to teaching science. Instead, the NSES advocate the use of diverse teaching strategies:

Although the Standards emphasize inquiry, this should not be interpreted as recommending a single approach to science teaching. Teachers should use different strategies to develop the knowledge, understandings, and abilities described in the content standards. Conducting hands-on science activities does not guarantee inquiry, nor is reading about science incompatible with inquiry. (p. 23)

The NSES state clearly that inquiry-based teaching does not mean a single teaching method or strategy. Teachers should combine and use various instructional strategies in order to make students engage in learning science through inquiry. While hands-on activities may increase the probability of more engaged learning (NRC, 2000), the NSES insist that they are neither sufficient nor necessary for inquiry-based learning of scientific concepts.

In this context, it is important to note that the NRC (2000) provides a working definition of scientific inquiry for the purpose of giving teachers a clearer idea about what scientific inquiry means in a classroom. The working definition consists of five essential features of classroom inquiry, but classroom inquiry does not necessarily encompass each one, resulting in different levels of inquiry as indicated in the table of ‘essential features of classroom inquiry and their variations’ (NRC, 2000) (see Table 1, next page). Depending on the amount of structure, guidance, and coaching that the science teacher provides for students, the inquiry-based classroom practice can range from the most student-directed open inquiry to the lesser student-directed guided inquiry and the most teacher-directed structured inquiry (Colburn, 2004).

Inquiry-based teaching is not a specific teaching approach prescribed in the standards. A broad range of instructional strategies should be included for inquiry-based teaching. It is the teachers who must select appropriate strategies for their classrooms in order to develop student knowledge and understanding of scientific ideas effectively through inquiry.

**Importance of Scientific Inquiry in K–12 Science Education**

Schwab (1962) argued that school science has departed from the science that scientists do, and that the situation has to be changed to cognize “science as a product of fluid inquiry” (p.5). Since Schwab, there has been a renewed emphasis on the importance of teaching
Table 1

Essential Features of Classroom Inquiry and Their Variations

<table>
<thead>
<tr>
<th>Essential Feature</th>
<th>Variations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Learner engages in scientifically oriented questions</td>
<td>Learner poses a question</td>
</tr>
<tr>
<td>2. Learner gives priority to evidence in responding to questions</td>
<td>Learner determines what constitutes evidence and collects it</td>
</tr>
<tr>
<td>3. Learner formulates explanations from evidence</td>
<td>Learner formulates explanation after summarizing evidence</td>
</tr>
<tr>
<td>4. Learner connects explanations to scientific knowledge</td>
<td>Learner independently examines other resources and forms the links to explanations</td>
</tr>
<tr>
<td>5. Learner communicates and justifies explanations</td>
<td>Learner forms reasonable and logical argument to communicate explanations</td>
</tr>
</tbody>
</table>

More ————————————————— Amount of Learner Self-Direction ————————————————— Less
Less ————————————————— Amount of Direction from Teacher or Material ————————————————— More


Science through inquiry (American Association for the Advancement of Science [AAAS], 1990; 1993; Bybee, 2000). The idea of teaching science as inquiry was firmly established in the National Science Education Standards in 1996.

The NSES (NRC, 1996) argue that students can actively develop their understanding of scientific concepts when they are engaged in inquiry. The addendum by the NRC (2000) provides images of inquiry in K–12 classrooms in the form of vignettes. The students are portrayed as deepening their understanding of scientific concepts and learning to think logically and critically about the world around them in the inquiry-based learning environment. Opinions that support the effectiveness of inquiry-based teaching are plenty. For example, Colburn (2004) states that “by prompting students to come up with their own answers, inquiry-based instructions lead to a deeper understanding of scientific concepts” (p. 63). Krajcik et al. (2000) state that “inquiry promotes development, transformation, and representation of ideas and helps learners understand how knowledge is generated in different disciplines” (p. 284). However, research into whether inquiry-oriented teaching really results in a higher level of student understanding of science is equivocal. For example, Shymansky, Kyle, and Alport (1983) reported effectiveness of 27 inquiry-based curricula in elementary, middle, and high schools through their systematic analysis of research studies. However, in a research study led by Shymansky, Yore, and Anderson (1999), no significant effect was found on academic achievement of 976 grade 3 and 4 students as a result of a major reform project in elementary science in the Iowa City Community School District. Von Secker and Lissitz (1999) conducted a statistical analysis of the data obtained from the 1990 National Education Longitudinal Study. For the sample of 7,642 grade 10 students, they found no evidence that instructional practices that emphasize critical thinking are associated with significant differences in student academic achievement.

The NSES (NRC, 1996) also argue that engaging students in inquiry helps them to develop an understanding of the nature of science. On this point also, research evidence is insufficient. Sandoval and Morrison (2003) found that 87 high-school students did not show a significant change in their understanding of the nature of science over the course of a four-week intervention of an inquiry-based biology unit. Moss, Abrams, and Kull (1998) conducted a qualitative research study on seven high school students. They found that a series of inquiry-based projects in an integrated science course over an entire school year did not
help them to develop their understanding of the nature of science.

It may not be surprising to find that research evidence is inconclusive, considering that there are a number of factors influencing the process of evaluating inquiry-based teaching (Harlen, 2004). For example, the level of student achievement may be influenced by the length of time during which an inquiry-based program has been implemented. A research method assessing long-term effects or immediate changes may also influence the findings. In fact, the National Research Council called for a meeting on the status of evaluation of inquiry-based science in 2004 (NRC, 2004). The purpose of the meeting was to build a framework for evaluating the effectiveness of inquiry-based science programs.

At the present point, this researcher has not yet uncovered enough evidence for the effectiveness of inquiry-based teaching. More examination of the literature will be required.

**Using Scientific Inquiry in the Classroom**

In spite of the emphasis on the importance of teaching science as inquiry in the standards (NRC, 1996), research literature suggests that very little inquiry-based teaching is taking place in classrooms. In a large-scale study involving 116 beginning secondary teachers, Simmons et al. (1999) found only 10% of them implemented student-centered inquiry-based instruction in their first year of teaching. Roehrig and Luft (2004) found that only four teachers out of 14 beginning secondary science teachers in their study were implementing guided inquiry lessons. In contrast, they identified eight of them as traditional teachers who asked students to find the right answers through textbook readings and laboratory exercises. Lotter (2004) reported that only three out of 13 secondary pre-service teachers in her study observed an inquiry-based lesson during their classroom observation hours in their teacher preparation program.

In school-science laboratories, inquiry-based teaching is seldom taking place as the committee organized by the National Research Council describes in their report (Singer, Hilton, &Schwiengruber, 2006). The report states that in most of high school science laboratories today, “instead of focusing on clear learning goals, teachers and laboratory manuals often emphasize the procedures to be followed, leaving students uncertain about what they are supposed to learn” (p. 133). McComas (2005) further contends that school laboratory exercises are typically cookbook type and aimed to verify classroom content that is already discussed.

One of the reasons for the lack of inquiry-based instruction in today’s science classrooms may be ascribed to the limited experiences that the teachers have with inquiry during their course of their own education. For example, one of the elementary pre-service teachers in a study conducted by Reiff (2002) wrote in the reflection that “I did not have much experience with inquiry throughout my schooling. I know most of my science classes in junior high and high school consisted of looking up a definition and answers to questions” (p. 7). Many of the elementary pre-service teachers in Reiff’s study expressed their fears about inquiry-based teaching because of their limited exposure to scientific inquiry. One of them wrote that “The question of inquiry is not an easy one. My lack of experience coupled with my limited exposure has left me questioning the entire method” (p. 6).

On the other hand, it seems that when teachers had experience working as a scientist, they had less difficulty with inquiry-based teaching. Crawford (1999) reports that a secondary pre-service teacher in her case study did not find difficulty in planning inquiry-based biology activities because of the 10-year experience the teacher had as a research technician. This teacher had a strong desire to engage students in scientific investigations because that was the way she experienced science. She also had appropriate knowledge and understanding about choosing activities for investigating a specific topic of the subject.

As a matter of fact, research studies show that inexperienced teachers feel more confident about inquiry-based teaching after they had an inquiry experience by themselves. For example, a high-school physics teacher who had initial reservations about teaching through inquiry changed his view toward inquiry after he had gone through a three-year master’s program, in which he was involved in a long-term inquiry project (NRC, 2000). In another episode provided by the NRC, an elementary teacher who participated in a university-based professional development program every summer wrote, “How does one begin to develop some expertise in these strategies we call inquiry? For me, I can only suppose that it began by reflecting upon my personal experiences” (p. 97).

The necessity of involving teachers in the experience of active scientific investigations is clearly advocated in the National Science Education Standards (NRC, 1996). The Professional Development Standard A for science teachers in the NSES advocates that science teachers of all levels must develop their scientific content knowledge through their inquiry experiences. By
acquiring scientific knowledge through inquiry, teachers can have the same opportunities as their students will have in the process of developing understanding of science in the classroom. In addition to subject matter knowledge, what teachers themselves must develop in the inquiry process are the abilities and understanding of how to conduct a scientific investigation.

Standard B for professional development in the NSES (NRC, 1996) advocates that teachers must learn to integrate scientific knowledge with knowledge about learning, pedagogy, and students. This special knowledge is called pedagogical content knowledge (PCK) (Shulman, 1986), and its possession distinguishes science teachers from scientists. Teachers must transform subject matter knowledge into forms that are useful for students to build their own subject matter knowledge. Crawford (1999) found that the biology pre-service teacher in her study who had research experience had difficulty in managing the classroom because of her lack of teaching experience. Some of the activities that the pre-service teacher skillfully planned for the topic were not at an appropriate level of the students. Crawford’s finding suggests that PCK for inquiry-based lessons is not a direct product of teachers’ inquiry experiences. It is rather a synthesis of knowledge about content and teaching that novice teachers must develop through teacher preparation programs and subsequent teaching experiences.

From the descriptions of the NESE (NRC, 1996), it becomes clear that besides general pedagogy, inquiry-based science teaching requires teachers to develop the following kinds of knowledge and skills. First of all, teachers should have solid scientific content knowledge. Second, specifically in their content knowledge, they should have the abilities necessary to conduct scientific inquiry as well as understanding of how authentic scientific inquiry is conducted. Third, they should possess PCK for inquiry-based lessons. In order to become competent in teaching science as inquiry, teachers must gain these kinds of knowledge through their education. The first two kinds of knowledge should be developed through teachers’ own inquiry experiences (NRC). The last kind of knowledge, PCK for inquiry-based teaching, should be developed through teacher preparation programs based on their inquiry experiences (Crawford, Zembal-Saul, Munford, & Friedrichsen, 2005). Incidentally, Roehrig & Luft (2004) found that an understanding of the nature of scientific inquiry, content knowledge, and PCK were three of the five factors that they identified impacting inquiry-based teaching of the beginning teachers.

**Summary**

Inquiry-based teaching is important because it has the potential for helping students to develop deeper scientific understanding. The National Science Education Standards (NRC, 1996) advocate its importance strongly. However, in actual classrooms very little inquiry-based teaching is taking place. It seems as if teachers have difficulty in translating what the standards say into teaching practices. Research studies suggest that one of the reasons for this difficulty may be the lack of inquiry experience that teachers have in the course of their own education. As a result, many beginning teachers do not have a general idea of what inquiry looks like, let alone skills and knowledge of teaching inquiry-based science lessons.

Literature suggests that providing opportunities for inquiry experiences, followed by instruction that helps teachers transform their experiences into appropriate PCK may improve their understanding about scientific inquiry. Literature also suggests that many professional development programs have been developed and implemented for the purpose of giving science teachers inquiry experience. One of the approaches developed recently for teacher educational programs is the use of technology in inquiry projects. Research studies in this area will be discussed next.

**USING TECHNOLOGY TO ENHANCE UNDERSTANDING OF SCIENTIFIC INQUIRY**

**Effectiveness of the Use of Technology in the Inquiry-Based Science Classrooms**

The use of technology for the purpose of promoting learners’ understanding of the nature of scientific inquiry did not start in teacher preparation programs; rather, it started in order to help children develop a deeper understanding of scientific concepts. Seymour Papert (1980) argues that children can change their way of thinking when they engage in computer programming. He developed a simple programming language called Logo. The simplicity of the language allows children to write their own programs. Papert found that children could come up with new ideas when they engaged in writing their own Logo programs, for example, to draw line graphics with an on-screen curser called “turtle”. He further developed his idea of learning by construction into a theory called “constructionism” (Papert, 1993). He advocates that children deepen their understanding of scientific concepts when they build a real-world concrete structure in addition to a mental
image in their head. Papert encourages children to engage in actively designing their own projects. His book describes a number of children who benefit from the constructionist approach, for example, by designing a doll house with blinking lights or a Lego motor that travels when it vibrates.

Since Papert, the idea of using technology to promote learning has evolved into a broader field (Novak & Krajcik, 2004). Notably, it has included the idea of using microcomputer-based activities to foster scientific inquiry. Resnick, Berg, and Eisenberg (2000) initiated the use of a small programmable device called the Handy Cricket for educational purpose. The Cricket is a fully programmable portable device which can control motors, lights, and data-collection from sensors (Martin, Mikhak, & Silverman, 2000). A user can write a Cricket Logo program and download it from a desktop computer to the Cricket through the infra-red transceiver. The data can be collected away from the computer and uploaded to the computer through the transceiver for an analysis. In the case study conducted by Resnick et al. (2000) three elementary students designed and constructed their own projects using the Cricket. The researchers found that the students engaged deeply in their projects. The students experimented with different designs to test their hypotheses. The freedom of writing their own programs allowed them to conduct a personalized analysis of data, and the children generated various explanations of the phenomena. The findings showed evidence of critical thinking and development of process skills for scientific inquiry.

The use of portable data-collecting devices allows students to investigate science outside the classroom, where they can find problems connected to the real world. Tinker (2001) argues that the use of the portable technology allows students to deepen their understanding of the real-world phenomena by reflecting on the data they actually collect at the site of the investigation.

In addition, the use of computer technology can assist students' inquiry processes in a way that would not be possible without it (Novak & Krajcik, 2004). The use of probes attached to a computer allows students to collect data that cannot be easily measured with ordinary tools. For example, the Cricket can measure light or temperature every half a second for 20 minutes. The Cricket can be also programmed to collect data once every minute for 40 hours. Both of these measurements are difficult to do with a thermometer, a stopwatch, a pen and paper. The computer-based data analysis further allows students to see the data in a graph. Graphs assist children to grasp the meaning of the data and generate in-depth understanding of physical phenomena (Linn, Layman, & Nachmias, 1987).

Krajcik and Starr (2001) reported that the use of portable technology in a middle school science class enhanced student conceptual understanding about water quality. In their study, 71 seventh grade students who participated in the project measured pH, temperature, and other variables in the nearby stream with probes attached to the hand-held computer. The researchers found that the students showed a significant gain in their conceptual achievement test scores and that their concept maps revealed their integrated understanding about scientific concepts of water quality.

Metcalf and Tinker (2004) found the effectiveness of their technology-based middle school science curriculum. A total of 149 students and 7 teachers participated in the 10-week implementation of the Transfer of Energy unit, which involved student use of hand-held computers called Palms and probes. The researchers found that the mean of the achievement test scores increased in all the teachers' classes.

Compared to the hand-held computers that were used in the above research studies, the Cricket is a primitive device. The user has to write a Logo program on a computer, and download it to the Cricket. When the user finishes collecting data, he/she has to upload the data, save the data, and open the data file with a spreadsheet so that he/she can see the data on a graph. It is true that the use of Crickets requires users more steps in the data processing than the use of, for example, a Palm hand-held computer with a Vernier sensor. However, the primitiveness of Crickets can offer users opportunities for developing a sense of conducting an active investigation. Crickets also give the users opportunities for understanding what a data collection in an authentic inquiry looks like by requiring to create their own data-collecting protocols (Martin & Greenwood, 2007). It should also be noted that Crickets are much more inexpensive than hand-held computers. This last feature of Crickets is important when one considers a possibility of providing a Cricket for each student and allowing students to take it home for data collection.

**Using Technologies in Teacher Preparation Programs**

From the discussion in the first part of the literature review, it becomes clear that science teachers should develop their understanding of science content and of the nature of scientific inquiry through their own authentic inquiry experiences. Research studies suggest that a professional development program that does not provide teachers with an authentic inquiry experience would not enhance their understanding of the nature of
scientific inquiry. Sandoval, Deneroff, and Franke (2002) found qualitatively that an activity-centered professional development program did not influence most of the 20 secondary teachers’ understanding about scientific inquiry. Luft (2001) found that a so-called “inquiry-based professional development program” that focused on discussions about inquiry without providing actual inquiry experiences for 14 secondary teachers did not improve their understanding about the nature of scientific inquiry.

There have been efforts to provide authentic inquiry experiences for pre-service and in-service teachers in K–12 science education. One example is to place teachers in scientists’ research labs and engage them in scientific inquiry processes (Schwartz, Lederman, & Crawford, 2004). Schwartz et al. found that most of the 13 secondary pre-service teachers who went through the laboratory experience showed substantial development in their understanding of the nature of authentic scientific inquiry. However, arranging for teachers to work in various laboratories in universities or in research institutes may not be always feasible. In addition, laboratory work may be different from one lab to another and it may not give all the teachers equally desirable experiences of inquiry (Crawford, Zembal-Saul, Munford, & Friedrichsen, 2005). It is necessary for higher educational institutes to develop a feasible teacher preparation program which provides them with authentic inquiry experiences without relying on other resources.

Crawford et al. (2005) investigated the effectiveness of a software-based professional development program about evolution. The program is called The Galapagos Finches and was used with 21 prospective secondary teachers in an advanced science methods class. Crawford et al. found that there was evidence that the program enhanced prospective teachers’ understanding of evolutionary concepts as well as their understanding of the nature of scientific inquiry. They claimed that the use of real-word database in their software allowed the teachers to experience authentic scientific inquiry.

Martin and Greenwood (2007) conducted a pilot study about the impact of a teacher preparation program in which pre-service and beginning in-service teachers were involved in a scientific investigation using the Cricket. About 20 teachers in a graduate-level science methods course participated in the study. They were asked to generate personally relevant scientific questions to conduct an investigation with the Cricket. The teachers designed and conducted their own data collection protocol. They also analyzed the data in order to generate a possible explanation of the phenomena.

For example, one of the teachers measured temperature in the attic and garage to see the effectiveness of the insulation that was newly installed in the teacher’s house. The researchers found that the teachers generated a wide variety of questions based on their personal interests, and could perform scientific work analogous to scientists. The researchers also found that the project gave the teachers a good opportunity for developing their scientific reasoning skills. For example, when a teacher had three line graphs that showed temperature change in a cup with ice cubes, the teacher needed to interpret the data in a way that was plausible and consistent with the actual situation. The process required the integration of scientific knowledge and critical thinking skills.

Unfortunately, Martin and Greenwood (2007) did not collect detailed data from the teachers. They did not have a systematic way of assessing teachers’ thinking and feelings about how they were impacted by the inquiry experience.

Furthermore, Martin and Greenwood (2007) did not include explicit considerations about developing teachers’ instructional strategies of teaching science through inquiry. From the discussion in the first part of this literature review, it becomes evident that in addition to inquiry experience, teachers need to develop their PCK in order to become competent in teaching science as inquiry. An inquiry experience alone does not normally lead teachers to develop teaching strategies for inquiry-based lessons (Crawford, 1999). In fact, a few prospective teachers in the study conducted by Crawford et al. (2005) expressed their concerns that they did not see a connection between the content in the science methods course and the inquiry experience that they had to go through with The Galapagos Finches software. There must be instructions that connect teachers’ experiences of scientific inquiry with teaching situations. The program that Martin and Greenwood developed lacks this pedagogical component. Crawford et al. advocate that authentic inquiry experiences should precede the teaching of pedagogical component in the methods course in a teacher preparation program. Ideally, a methods course in science education should provide prospective teachers with authentic inquiry experiences using the technology first, and then move on to the discussions about teaching strategies that would be effective in inquiry-based classrooms.

**Summary**

The use of technologies in a scientific investigation has the potential for enhancing students’ and teachers’
understanding of scientific ideas and of the nature of scientific inquiry. When beginning teachers conducted a scientific investigation based on their personal interest, the use of the programmable data-collecting device called the Handy Cricket had noticeable impact on their understanding of the nature of scientific inquiry. However, the preliminary nature of the study done by Martin and Greenwood (2007) did not allow much information to be revealed for the evaluation of the effectiveness of the program. Furthermore, in order for an authentic inquiry experience to impact the classroom teacher’s practice, research indicates that the teacher may need assistance in the development of appropriate PCK.

CONCLUSION AND POTENTIAL RESEARCH QUESTIONS

In spite of the emphasis on inquiry-based teaching advocated in the National Science Education Standards (NRC, 1996), very few science teachers in K–12 education teach science as inquiry. Classroom methods are mostly teacher-centered and students are asked to find the right answer by checking their textbooks. Laboratory exercises are commonly cookbook type where students blindly follow written instructions to verify content they already take for granted.

The research reviewed has indicated that one of the reasons for this lack of inquiry-based teaching may be ascribed to the limited experience that teachers have with scientific inquiry during their own education. It is suggested then, that teacher preparation programs must address this deficiency by providing prospective and practicing science teachers with authentic inquiry experiences and the subsequent instructions for transforming the experiences into appropriate PCK.

The research reviewed has also indicted that the use of technologies to support inquiry-based instructions is effective in enhancing students’ and teachers’ understanding of scientific inquiry. The literature has indicated that students and teachers develop understanding of scientific inquiry when they engage in designing and conducting their own experiments with the help of technology. Martin and Greenwood (2007) embedded the use of a programmable device called the Cricket in a teacher preparation program. The present review has indicated that the program provided pre-service and beginning in-service teachers with authentic inquiry experiences in a feasible way.

Theory and research into other areas pertinent to the problem still need to be reviewed. However, the present review indicates that research into the effectiveness of the teacher preparation program that makes use of the Cricket technology is equivocal. More data collection followed by detailed analysis needs to be conducted in order to examine the effectiveness of the program. Moreover, the program itself needs to be developed further to incorporate explicit instructions focused on the transformation of the teachers’ inquiry experiences into appropriate PCK for planning an inquiry-based classroom.

Consequently, research into the effectiveness of the Handy Cricket as a tool to develop teachers understanding of authentic inquiry, as well as research which explores how scientific inquiry experiences are translated into PCK is warranted. The following research questions are suggested:

1. How do teachers change their understanding of the nature of scientific inquiry in the course of a teacher preparation program in which the Handy Cricket is utilized?
2. How do teachers’ authentic inquiry experiences impact their beliefs and understanding of the feasibility of inquiry-based teaching?
3. What are the essential features of a teacher preparation program that help teachers transform their inquiry experience into appropriate PCK for planning an inquiry-based classroom?

Answering these questions might help educators in higher-education to develop a teacher preparation program that is effective and feasible with the use of Cricket technology. If teachers have a better understanding of the nature of scientific inquiry, it is expected that students will have a better chance of learning science through inquiry in science classrooms.

REFERENCES


2009 Annual Colloquium on Research in Mathematics and Science Education

Call for Papers

Educators, researchers, and graduate students are invited to submit papers that will be presented at the Twelfth Annual Colloquium on Research in Mathematics and Science Education and published in the Colloquium Journal, vol. XIV. The papers must discuss issues and trends in Mathematics and Science Education.

WHEN SUBMITTING A PAPER, PLEASE USE THE FOLLOWING GUIDELINES.

1. Submit an electronic version of the paper and one hard copy, an abstract, approximately 150 words, and a biographical sketch, about 30 words. All pictures and diagrams must be submitted in a separate document.

2. Use double spacing with one-inch margins.

3. For references, diagrams, etc. follow the style described in the Publication Manual of the American Psychological Association (APA), Fifth Edition.

4. Paper length must not exceed 30 pages, including pictures, tables, figures, and list of references.

5. Paper must be received by November 15, 2008.

6. Authors will be notified about the status of their paper by January 15, 2009.

7. A Colloquium will be scheduled for April 2009.

SUBMIT PAPERS AND CORRESPONDENCE TO:

Dr. Regina M. Panasuk
Professor of Mathematics Education
The Graduate School of Education
University of Massachusetts Lowell
61 Wilder Street, O’ Leary 5th Floor
Lowell, MA 01854

Phone: (978) 934-4616
Fax: (978) 934-3005
Regina_Panasuk@uml.edu