Highlights from The Trends in International Mathematics and Science Study (TIMSS) 2007:

Mathematics and Science Achievement of U.S. Fourth and Eighth-Grade Students in an International Context

December 2008

Executive Summary

The 2007 Trends in International Mathematics and Science Study (TIMSS) is the fourth administration since 1995 of this international comparison. Developed and implemented at the international level by the International Association for the Evaluation of Educational Achievement (IEA)—an international organization of national research institutions and governmental research agencies—TIMSS is used to measure over time the mathematics and science knowledge and skills of fourth- and eighth-graders. TIMSS is designed to align broadly with mathematics and science curricula in the participating countries. This report focuses on the performance of U.S. students relative to that of their peers in other countries in 2007, and on changes in mathematics and science achievement since 1995. Thirty-six countries or educational jurisdictions participated at grade four in 2007, while 48 participated at grade eight. This report also describes additional details about the achievement of U.S. student subpopulations. All differences described in this report are statistically significant at the .05 level. No statistical adjustments to account for multiple comparisons were used.

Key findings from the report include the following:

- In 2007, the average mathematics scores of both U.S. fourth-graders (529) and eighth-graders (508) were higher than the TIMSS scale average (500 at both grades). The average U.S. fourth-grade mathematics score was higher than those of students in 23 of the 35 other countries, lower than those in 8 countries (all located in Asia or Europe), and not measurably different from those in the remaining 4 countries. At eighth grade, the average U.S. mathematics score was higher than those of students in 37 of the 47 other countries, lower than those in 5 countries (all of them located in Asia), and not measurably different from those in the other 5 countries.

- Compared to 1995, the average mathematics scores for both U.S. fourth- and eighth-grade students were higher in 2007. At fourth grade, the U.S. average score in 2007 was 529, 11 points higher than the 1995 average of 518. At eighth grade, the U.S. average mathematics score in 2007 was 508, 16 points higher than the 1995 average of 492.

- In 2007, 10 percent of U.S. fourth-graders and 6 percent of U.S. eighth-graders scored at or above the advanced international benchmark in mathematics. At grade four, seven countries had higher percentages of students performing at or above the advanced international mathematics benchmark than the United States: Singapore, Hong Kong SAR, Chinese Taipei, Japan, Kazakhstan, England, and the Russian Federation. Fourth-graders in these seven countries were also found to outperform U.S. fourth-graders, on average, on the overall mathematics scale. At grade eight, a slightly different set of seven countries had higher percentages of students performing at or above the advanced mathematics benchmark than the United States: Chinese Taipei, Korea, Singapore, Hong Kong SAR, Japan, Hungary, and the Russian Federation. These seven countries include the five countries that had higher average overall mathematics scores than the United States, as well as Hungary and the Russian Federation.

- In 2007, the average science scores of both U.S. fourth-graders (539) and eighth-graders (520) were higher than the TIMSS scale average (500 at both grades). The average U.S. fourth-grade science score was higher than those of students in 25 of the 35 other countries, lower than those in 4 countries (all of them in Asia), and not measurably different from those in the remaining 6 countries. At eighth grade, the average U.S. science score was higher than the average scores of students in 35 of the 47 other countries, lower than those in 9 countries (all located in Asia or Europe), and not measurably different from those in the other 3 countries.

- The average science scores for both U.S. fourth- and eighth-grade students in 2007 were not measurably different from those in 1995. The U.S. fourth-grade average science score in 2007 was 539 and in 1995 was 542. The U.S. eighth-grade average science score in 2007 was 520 and in 1995 was 513.

- In 2007, 15 percent of U.S. fourth-graders and 10 percent of U.S. eighth-graders scored at or above the advanced international benchmark in science. At grade four, two countries had higher percentages of students performing at or above the advanced international science benchmark than the United States: Singapore and Chinese Taipei. Fourth-graders in these two countries were also found to outperform U.S. fourth-graders, on average, on the overall science scale. At grade eight, six countries had higher percentages of students performing at or above the advanced science benchmark than the United States: Singapore, Chinese Taipei, Japan, England, Korea, and Hungary. These six countries also had higher average overall eighth-grade science scores than the United States.

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 December 1.

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

Inquiry-Based Teaching for American and Japanese Middle-School Science Teachers: Two Views, Two Practices  
*Sachiko Tosa* .............................................................................................................. 1

Enhancing the Teaching of the Nature of Science  
*Michael Wadness* ........................................................................................................ 11

A Study of the Relationship Between Group Perception of School Climate and Incidences of Bullying at the Junior High/Middle School Level  
*Kimberly Douglass* .................................................................................................... 24

## Qualifying Papers

Increasing the Effectiveness of Mathematics Instruction Through Lesson Planning  
*Noreen Flanagan Johnson* .......................................................................................... 39

Interactive-Engagement and Reflective Writing  
*Michael Doherty* ........................................................................................................ 53
CONTRIBUTORS

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 since 1995, and is involved in a technology-based educational project in the Computer Science Department at UML. Sachiko is a recipient of Graduate Research Scholar award in 2008 and currently working on the final stage of her dissertation.

Kimberly Douglass is currently completing her dissertation for the Leadership in Schooling Program. She graduated with a B.S. in Special Education in 1994 and an M.S. in School Psychology in 2000. Kimberly is currently employed as a school psychologist for the Northborough, Massachusetts School Department.

Michael Wadness is a doctoral student in mathematics and science education and is making his final preparations to begin his dissertation research. Michael has been teaching physics for eleven years at Medford High School.

Noreen Flanagan Johnson teaches Mathematics at the Greater Lawrence Technical School. She is currently working on her dissertation proposal in the Math and Science Education doctoral program at UML.

Michael Doherty has been teaching Physics, Chemistry, and Electronics at Andover High School. For the past two years he has been involved in UML Teams Academy teaching Baseball Bat Design Engineering. He is currently working on his dissertation proposal in the Math and Science Education doctoral program at UML.

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

Vera Ossen, Ed.D. Regina Panasuk, Ph.D.
Vera_Ossen@uml.edu Regina_Panasuk@uml.edu
The Director of Educator Licensure Programs Mathematics Education

David Lustick, Ed. D. Michelle Scribner-MacLean, Ed.D
David_Lustick@uml.edu Michelle_ScribnerMacLean@uml.edu
Science Education Science Education

Application packets may be obtained on line or by contacting the Graduate School of Education at (978) 934-4601

2008–2009 Academic Year
Mathematics and Science Education Program

<table>
<thead>
<tr>
<th>DISSERTATION DEFENSE STAGE</th>
<th>DISSERTATION PROPOSAL STAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neil Hatem</td>
<td>Wesley Johnson</td>
</tr>
<tr>
<td>Sachiko Tosa</td>
<td>Matthew Beynavenard</td>
</tr>
<tr>
<td>Michael Wadness</td>
<td>Kate McLaughlin</td>
</tr>
<tr>
<td></td>
<td>Alex Ballantyne</td>
</tr>
<tr>
<td></td>
<td>Valerie Finnerty</td>
</tr>
<tr>
<td></td>
<td>Jaimie Stone</td>
</tr>
<tr>
<td></td>
<td>Noreen Flanagan Johnson</td>
</tr>
<tr>
<td></td>
<td>Michael Doherty</td>
</tr>
</tbody>
</table>
Inquiry-Based Teaching for American and Japanese Middle-School Science Teachers: Two Views, Two Practices

This work was in part supported by CAREER grant provided by the National Science Foundation under the grant number 0546513.

Sachiko Tosa
University of Massachusetts Lowell

ABSTRACT

This study examines similarities and differences in how Japanese and American middle-school science teachers think and feel about inquiry-based teaching. Their attitudes toward the use of inquiry in science teaching were measured through a survey instrument (N=191). The results showed that American teachers value the importance of inquiry-oriented teachers’ actions more highly than Japanese teachers do. Japanese teachers reported feeling less comfortable helping students with their questions and activities. These findings were further examined through school visits in Japan (N=15) and the United States (N=9). Preliminary analysis of data indicated that Japanese teachers were not generally helping students construct their own understanding of scientific concepts in spite of well-organized lesson structures and activity set-ups. On the other hand, the American lessons often lacked meaningful science content in spite of the high level of inquiry-oriented teaching strategies that many teachers exhibited. Implications regarding the cultural differences in science education will be discussed.

Since the publication of the National Science Education Standards (NSES) (National Research Council [NRC], 1996), the importance of scientific inquiry has been clearly emphasized in K-12 science education in the United States. Scientific inquiry is regarded as the heart of science education because inquiry helps students develop deeper understanding of scientific concepts (NRC). 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, p. 22). The effectiveness of inquiry-based teaching for developing students’ deeper understanding of scientific concepts has been suggested by various researchers (Anderson, 2002; Colburn, 2004; Ertepinar & Geban, 1996; Krajcik et al.; Shymansky, Kyle, & Alport, 1983; Von Secker, 2002).

In spite of the emphasis on scientific inquiry in the science education research community, the literature suggests that very few inquiry-based lessons are taking place in actual science classrooms (Lotter, 2004; Reiff, 2002; Simmons et al., 1999). The Third International Mathematics and Science Study (TIMSS) 1999 Video Study of science (National Center for Education Statistics [NCES], 2006) showed that 66% of 8th-grade science teachers’ lessons developed science content mainly by student acquisition of facts, definitions, and algorithms. The study showed that only 17% of the U.S. lessons developed science content through inquiry. It appears that U.S. teachers are not translating the guidelines for inquiry-based teaching described in the NSES.

In contrast to the situation in the United States, the literature suggests that Japanese science lessons seem to be more inquiry-oriented (NCES, 2006). In the TIMSS 1999 Video Study (NCES), 72% of Japanese science lessons for 8th graders developed science content mainly by making connections between students’ activities and the explanations. The study further showed that 57% of the Japanese lessons developed content by making connections through inquiry. However, there is some evidence that Japanese science teachers are just following prescribed ways of teaching science as inquiry, and that their teaching is not inquiry-based teaching in the sense described by the NSES (Ogura, 2004). In the independent research study based on the TIMSS 1999 videotaped lessons, Ogura found that U.S. science teachers were critical about Japanese lessons; U.S. teachers who watched the Japanese video-taped lessons found that Japanese teachers were typically so concerned with giving students “the right answer” as the result of their activity that they rushed toward the conclusion at the end of the lessons.

Given a set of contradicting findings, it is important to conduct a study on science teachers’ thoughts and feelings about inquiry-based teaching in Japan and the United States. The TIMSS Video Study (NCES, 2006) showed that inquiry was more widely used in middle school science in Japan than in the United States based on observational data collected by the governmental agencies. The next step may be to more closely study what science teachers think and feel about inquiry-based teaching in these countries. By studying individual teachers’ thoughts and feelings toward inquiry, we
may be able to understand the situation of inquiry-based teaching from teachers’ points of view. By observing lessons in a more private setting, we may be able to collect data from their actual daily practices. Moreover, by matching up teachers’ thoughts and feelings with their practice, we may be able to draw a better picture of the current status of inquiry-based teaching in middle school science in each country. Furthermore, by understanding how the meaning of inquiry-based teaching is similar and different to American teachers and Japanese teachers, we may be able to identify what elements are missing for promoting inquiry-based teaching in both countries.

Consequently, the following research questions regarding U.S. and Japanese middle-school science teachers guided the present study:

1. What is their understanding of inquiry-based teaching?
2. What are the similarities and differences between their understandings of inquiry-based teaching?
3. How do their teaching practices reflect their understanding of inquiry-based teaching?
4. What are their attitudes toward the use of inquiry-based teaching?
5. What are the similarities and differences between their attitudes toward the use of inquiry-based teaching?
6. How do their teaching practices reflect their attitudes toward inquiry-based teaching?

The significance of this study is to inform science educators what might be lacking in science teaching in Japan and the United States. As Stigler, Gallimore, and Hiebert (2000) described in their paper of a comparative study, teaching is a cultural activity. It is difficult to see some of the ubiquitous teaching practices and beliefs in one’s own culture. It is often the case that educators are not aware of lacking elements because they may be buried in the tradition of teaching in the particular country. By comparing U.S. and Japanese teachers in terms of their understanding of, and attitudes toward the use of inquiry, this study reveals what is taken for granted as inquiry-based teaching in each place. This study further provides U.S. and Japanese science educators in higher education with information about what should be more strongly emphasized in teacher education programs in order to increase the use of inquiry in middle school science.

METHOD

Research Methods

This study measured U.S. and Japanese science teachers’ understanding of, and attitudes toward inquiry-based teaching. Teachers’ attitudes were measured by a survey instrument. The survey data were analyzed statistically. Data on teachers’ understanding of inquiry-based teaching were collected qualitatively through lesson plans, classroom observations, and interviews. The qualitative data provide information about the context of each of the lessons, including the content, the teacher, and the students.

Participants

Fifty-seven American and 134 Japanese teachers completed the survey. They were the members of listserv networks to which the researcher had access. Participants for classroom visits were first recruited through mailings to randomly selected middle schools. Participants were also recruited through a request that was posted at the end of the survey. Fifteen Japanese and ten American teachers agreed to participate in classroom observations and interviews. Four of the American teachers dropped from the participation in the process, and three more American teachers were recruited using a snowball sampling technique.

Instrumentation

Instruments to measure teachers’ attitudes toward inquiry-based teaching were not available in literature (Smolack, Zembal-Saul, & Yoder, 2006). Thus, prior to this study, the researcher created and piloted a survey instrument by clearly defining attitudes toward inquiry-based teaching into four categories. The scale measured teachers’ (a) general preference of inquiry-based teaching, (b) personal evaluation toward performing typical behaviors in inquiry-based teaching, (c) beliefs about the possible obstacles to inquiry-based teaching, and (d) self-efficacy about the use of inquiry in science teaching. The instrument contains 45 questions: 37 closed-ended questions with ordered response option, one open-ended question about inquiry-base teaching, and seven demographic questions. The survey instrument items and their categories can be found in the Appendix.

The researcher also created an interview protocol. The interview protocol included questions regarding the observed lessons as well as questions regarding teachers’ thoughts about inquiry-based teaching. The interview started with a question about the lesson that was just completed. In this way, the interviewer was able to have the participants describe their thoughts about the use of inquiry on a concrete basis. Furthermore, using the interview protocol the researcher was able to collect data that could be sorted.
out in six categories for participants' understanding of
inquiry-based teaching: (a) definitions of inquiry-based
teaching, (b) what constitutes student learning, (c)
teachers' depth of understanding of subject-specific
knowledge, (d) teachers' understanding of effectiveness
of different instructional strategies in inquiry-based
teaching, (e) teachers' depth of understanding of the
ways to transform scientific concepts into teachable
forms in inquiry-based teaching (pedagogical content
knowledge), and (f) teachers' understanding of the
nature of scientific inquiry.

The rubric for analyzing classroom observations
was also created by the researcher. The observed lessons
were coded by using a rubric that describes how each of
the essential features of inquiry-based teaching (NRC,
2000) were practiced. The rubric included three points
(X) based on who completed the activity (teachers 1,
students and teachers 2, students alone 3), and five
points (Y) for the depth of conceptual link displayed.
The two kinds of points were given for each of the
essential features of inquiry-based teaching.

Newly-created instruments were reviewed by a
group of experts for content validity. Following the
experts' comments, the instruments were revised several
times before they were submitted to the Institutional
Review Board for permission for conducting research on
human subjects.

Procedures

The request for middle-school science teachers to
participate in the survey was distributed through U.S.
and Japanese listserv networks separately in late May,
2008. The responses were collected through the use of
online survey instrument provided by a commercial
company. Both American and Japanese data collection
took place for two-and-a-half weeks.

The researcher visited each of the interview partici-
pants' schools once to conduct a classroom observation,
an interview, and a collection of artifacts including the
lesson plan and worksheets. The researcher observed
15 Japanese science teachers' lessons at 13 public mid-
dle schools in the Tokyo area in June and July, 2008.
Nine American teachers' classes at eight public middle
school schools in the state of Massachusetts were
observed in September and October in 2008.

RESULTS

Cronbach's alpha for the whole set of survey data
(N=191) was .796, indicating high internal consistency
reliability. Alpha for the Japanese data alone (N=134)
was .815, indicating even higher internal consistency
reliability. The Cronbach's alpha for the U.S. data
(N=57) alone was slightly lower, yet still reliable at
.771.

Factor analysis

A factor analysis was performed for the whole set of
data (N=191). The analysis identified 10 factors with
eigenvalues greater than one. The matrix rotation con-
verged and the items that share similar characteristics
were identified. The percentage variation explained is
65. The scree plot is shown in Figure 1. It shows a dis-

![Scree Plot](image)

Figure 1. Scree plot for the factor analysis
distinctive elbow.

By considering common features of the items grouped in the factors, each of the factors can be named as they are shown in Table 1.

The numbers in Table 1 indicate the average numerical values (and standard deviation) for the items grouped in the factors. The numerical values 5 to 1 were given to each of the response options for the 36 items in the instrument, where strongly agree = 5, agree = 4, neither agree nor disagree = 3, disagree = 2, and strongly disagree = 1. The asterisk (*) in the last column in Table 1 indicates that the difference between the U.S. and Japan is statistically significant for the factor (p<.01).

Four of the items that address the obstacles to inquiry-based teaching were factored out in Factors 3 and 8. The items that ask teachers’ self-efficacy about inquiry-based teaching were loaded on particular factors (Factors 2, 4, and 5). It is interesting to observe that underlying categories in defining teachers’ attitudes toward the use of inquiry in teaching was revealed in the analysis. This result implies that the scale well represents the underlying latent variables in the research project.

The results of the factor analysis indicate that both American and Japanese teachers highly value the importance of the elements of inquiry-based teaching (Factor 1). They both agreed, for example, that it is important individual students formulate their explanations from evidence in their science classes (Q13). The factor analysis also shows that American teachers agree with the importance of inquiry-oriented teachers’ actions more strongly than Japanese teachers. It also shows that American teachers were more satisfied with the extent to which they use inquiry in their science classes than Japanese teachers (Factor 9). One of the reasons for the Japanese teachers’ reservations toward the use of inquiry might be the time constraint in the Japanese curricula; Japanese teachers have only 105 science lessons in a school year (Ministry of Education, Culture, Sports, Science, and Technology [MEXT], 2007), whereas American teachers typically have 135 hours (NCES, 2008). The Japanese teachers’ concern about the time constraint as an obstacle to inquiry-based teaching is clearly shown in the factor analysis results (Factor 3).

The factor analysis also reveals that Japanese teachers felt less comfortable helping students with their questions and activities (Factors 4 and 5). This somewhat surprising result will be discussed in the next section.

The observed lessons were coded using the rubric

<table>
<thead>
<tr>
<th>Factor</th>
<th>Name of the factor (survey items grouped in the factor)</th>
<th>Mean Score (SD)</th>
<th>Significance of difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td>Standard-based concepts of inquiry-based teaching (Q10, Q12, Q13, Q14, Q18, Q21)</td>
<td>4.46(0.37)</td>
<td>4.20(0.53)</td>
</tr>
<tr>
<td>Factor 2</td>
<td>Students’ self-directedness and teachers’ confidence in teaching inquiry-based lessons (Q4, Q9, Q25, Q36, Q37)</td>
<td>4.17(0.51)</td>
<td>3.93(0.59)</td>
</tr>
<tr>
<td>Factor 3</td>
<td>Time and material constraints as obstacles to inquiry-based teaching (Q33, Q34, Q35)</td>
<td>2.76(0.85)</td>
<td>3.48(0.79)</td>
</tr>
<tr>
<td>Factor 4</td>
<td>Helping students and the importance of inquiry-based teaching (Q2, Q24, Q29)</td>
<td>4.46(0.72)</td>
<td>3.99(0.57)</td>
</tr>
<tr>
<td>Factor 5</td>
<td>Dealing with students’ questions and higher level of thinking (Q22, Q26, Q27, Q28)</td>
<td>4.41(0.41)</td>
<td>3.95(0.52)</td>
</tr>
<tr>
<td>Factor 6</td>
<td>Correctness of the scientific concepts (Q16, Q23)</td>
<td>3.35(0.69)</td>
<td>3.41(0.62)</td>
</tr>
<tr>
<td>Factor 7</td>
<td>Secondary important concepts of inquiry-based teaching (Q6, Q19, Q20)</td>
<td>4.35(0.46)</td>
<td>4.24(0.47)</td>
</tr>
<tr>
<td>Factor 8</td>
<td>Presence of mandate tests as an obstacle to inquiry-based teaching (Q30)</td>
<td>3.49(1.16)</td>
<td>3.19(1.17)</td>
</tr>
<tr>
<td>Factor 9</td>
<td>Teachers’ satisfaction (Q3)</td>
<td>3.42(1.08)</td>
<td>2.93(0.92)</td>
</tr>
<tr>
<td>Factor 10</td>
<td>Importance of whole class understanding (Q17)</td>
<td>4.18(0.87)</td>
<td>4.33(0.68)</td>
</tr>
</tbody>
</table>

*p<.01

Table 1. Name of Factors and Mean Scores for the Items in Each Factor by Country.
that describes how each of the essential features of inquiry-based teaching was practiced. Two numbers were given to each essential feature found in the lesson. When the particular features were not found in the observed lesson, the researcher asked in the interview if the teacher intended on including the feature in the following lesson. If no intention was indicated, the lesson received (0, 0) as a set of scores for the feature. If the teacher indicated his/her intention to carry over the feature into the next lesson, it was left as blank for the feature. Table 2 shows the preliminary results of the mean scores for each of the essential features observed in American and Japanese lessons.

Table 2 indicates that on average it was the teachers who gave students the questions to be investigated during the lesson in both American and Japanese lessons (X scores of Feature 1). It was also the teachers, on the average, who connected explanations to scientific knowledge (X scores of Feature 4). In contrast, when it came to collecting data, students did the activities in both countries (X scores of Feature 2). However, for formulating explanations from data, more American teachers let students do the task (X scores of Feature 3). Table 3 also shows that on the average little conceptual links were displayed in both American and Japanese lessons (Y scores of overall mean). However, it is interesting to notice that the American lessons received higher Y scores for the features that were exhibited in the beginning of the lessons (Features 1 and 2) than those exhibited later in the lessons (Features 3 and 4) . On the other hand, the Japanese lessons received the same average scores throughout the features from 1 to 4.

**DISCUSSION**

This study examined differences and similarities between U.S. and Japanese science teachers in terms of their thoughts and feelings toward the use of inquiry-based teaching. The findings from the survey analysis showed that American teachers agreed more strongly with inquiry-based teaching than Japanese teachers. The findings also showed that there was a statistically significant difference between U.S. and Japanese teachers in terms of the willingness of helping students with their questions and problems during their science classes.

These findings make more sense when descriptive observation data were taken into consideration. The researcher found that except for one lesson, which was devoted to skills training of connecting an electric voltage meter and a current meter in a circuit, the observed Japanese lessons started with scientifically meaningful questions. The questions were examined by students in experiments or observations. The Japanese middle-school students were often instructed to perform sophisticated experiments using equipments such as gas burners and microscopes. After the experiments, Japanese students were instructed to write down the conclusions that they drew from the activities. At the end of the lessons, teachers made connections between what they found from the activities and scientific con-

<table>
<thead>
<tr>
<th>Essential Feature of Inquiry</th>
<th>Mean scores (SD) for who complete the activity X and how conceptual links were made Y</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>US (N=9)</td>
</tr>
<tr>
<td>1. Students engage in scientifically oriented questions</td>
<td>1.06 (0.17)</td>
</tr>
<tr>
<td>2. Student gives priority to evidence in responding to questions (Student collects and analyzes data)</td>
<td>2.06 (0.18)</td>
</tr>
<tr>
<td>3. Student formulates explanations from evidence</td>
<td>2.25 (0.76)</td>
</tr>
<tr>
<td>4. Student connects explanations to scientific knowledge</td>
<td>1.14 (0.90)</td>
</tr>
<tr>
<td>5. Student communicates and justifies explanations</td>
<td>0.67 (0.52)</td>
</tr>
<tr>
<td>Overall Mean</td>
<td>1.49 (0.30)</td>
</tr>
</tbody>
</table>

Note. X scores indicate who did the activity: 1 = teacher only, 2 = teacher and students together, 3 = students only. Y scores indicate the depth of conceptual links displayed by teacher and students during the lesson, ranging from 1 = no conceptual links to 5 = multiple links as conceptual networks. Sig. indicates the significance of difference between countries.
cepts in a whole-class discussion. In spite of the well-organized lesson structure, the researcher did not observe incidences in which Japanese teachers helped students construct their own meaning of the scientific concepts. This trend was indicated by the flat Y scores of 2.00 for the features 1 through 4 in Table 2. It shows that Japanese teachers were making connections between activities and concepts at each step of the content development. However, the connections were made not through logical reasoning steps, but by given facts. The Japanese teachers rarely took time to relate the questions to students’ prior experiences or concepts that they have learned already. Many of the teachers even omitted stating what the purpose of the experiment was before students started the task. Many students often expressed in private conversations that they did not know what they were expected to do. During the student activities, the Japanese teachers were often busy checking if the experiments were performed in the right way. During the interviews, however, the Japanese teachers expressed the importance of helping students make connections between the outcomes of their activities and the scientific concepts. It seems that the Japanese teachers are unfamiliar with actual ways how to help students construct their knowledge.

Compared to the Japanese lessons, the observed American lessons had more variations in the Y scores. The larger standard deviations in Y for U.S. teachers in Table 2 support this point. Some of the teachers in the observed American lessons helped students well in their construction of the meaning of the activities through a whole-class discussion. These teachers related the topic with what students learned already. Some of the teachers also prompted students to come up with the experimental procedures by a series of questionings. In spite of the high level of pedagogical skills that some of the teachers exhibited, many American lessons did not contain scientifically meaningful content. The researcher found little scientific concepts involved in three American lessons, and they received low Y scores; if there were no scientific concepts to be connected with the results of the student activities, there would be no conceptual links made. It also happened in some of the U.S. lessons that the experiments did not show what the teacher wanted to show. When the experiments did not work, students were having a hard time generating explanations for the phenomena and the lessons received low Y scores for Features 3 and 4.

The differences between U.S. and Japanese science lessons that were observed in this study may be attributed to the different experiences that the teachers had during the course of their education in terms of the amount of science content they have learned. All of the Japanese teachers hold undergraduate degrees in either science or science education. Undergraduate students in science education in Japan are required to take science courses in a similar way as students who are in science and engineering departments in order to be certified as middle-school science teachers (MEXT, 2008). They are often expected to develop teaching materials as undergraduate research projects. For example, a Japanese teacher who participated in the interview reported that as his undergraduate research project he developed a way to measure time when a falling object passed two distant points using sensors. Even at the graduate level, Japanese students in science education are required to take science classes. A typical research project at the Master's level in Japanese science education is to study the effectiveness of the use of materials that were developed to teach specific science content. Three of the Japanese teachers who participated in the classroom observations hold master's degrees in science related fields. It was clear that all the Japanese teachers who participated in the observations had strong background in science content and materials in the course of their education. On the other hand, American teachers who participated in this research had less experiences studying science at higher education institutions. Four of the American teachers did not have science background as undergraduate or graduate degrees; two of them majored in elementary education, while the other two studied humanity subjects and were certified to teach science after graduation. The other five teachers studied science as undergraduates and studied education at the graduate level. However, none of them took science courses at their graduate schools.

Teachers’ experiences of learning science content are not the only place where differences between U.S. and Japanese middle-school science teachers are found. When the demographic data is examined in terms of the amount of teachers’ experiences in learning pedagogy, another difference is revealed. Seven of the Japanese teachers who hold degrees only in science did not study education as their major subject of study. In contrast, only one of the American teachers did not have experience of learning education as the major field. As far as the observation participants were concerned, the American middle-school science teachers had more experiences than the Japanese teachers in learning pedagogy in their own course of education.

Table 3 shows the result of the analysis when demographic data of survey participants and observation participants was examined in terms of their educational background.
Table 3.

Percentage of Teachers Who Have Academic Degrees in Science and/or Education in the USA and Japan

<table>
<thead>
<tr>
<th>Educational Background</th>
<th>Survey Participants % (N)</th>
<th>Observation Participants % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>US (N=45)</td>
<td>Japan (N=46)</td>
</tr>
<tr>
<td>Science and Education</td>
<td>44% (20)</td>
<td>37% (17)</td>
</tr>
<tr>
<td>Science only</td>
<td>11% (5)</td>
<td>63% (29)</td>
</tr>
<tr>
<td>Education only</td>
<td>40% (18)</td>
<td>0% (0)</td>
</tr>
<tr>
<td>Neither Science nor Education</td>
<td>4% (2)</td>
<td>0% (0)</td>
</tr>
</tbody>
</table>

Note. The numbers in the parentheses show the actual numbers of respondents in the categories.

The numbers of respondents for the survey participants shown in Table 3 were much smaller than the numbers of whole survey participants due to the fact that many teachers did not provide the names of the fields they studied. Table 3 indicates that the trend that was found for the observation participants also holds for the survey participants: all the Japanese participants have science in their educational background, whereas more than 40% of American teachers did not have science as the major field of study in their own course of education. In terms of background of learning pedagogy, 63% of Japanese teachers did not have much experience, whereas only 15% of American teachers were in this category.

The findings of this study imply that the lack of content in U.S. middle-school science lessons might be due to the lack of science education in teachers’ background. The findings suggest that one of the ways to improve U.S. science teachers’ practice would be to include more science courses for teacher preparation programs. Without solid content knowledge, teachers would not be able to use pedagogical knowledge for inquiry-based teaching that they have acquired through their education trainings. On the other hand, one way to improve Japanese teachers’ practice for promoting inquiry-based teaching would be to put more emphasis on pedagogy than on materials and procedures in their teacher preparation programs. It should be noted, however, that these conclusions drawn from the observational data were based on small sample sizes and thus should be taken cautiously.

Further analysis of observation data and interview data is underway. It is expected that interview data will reveal what teachers assumed as inquiry-based teaching within their culture. It would be interesting to find out how teachers’ assumptions about inquiry-based teaching were connected with their practice. Through these analyses, it may be possible to further suggest effective ways to improve teacher education programs in both countries.

References


---

**APPENDIX**

**Table A1.**

**Survey Instrument and Categories for Measuring Middle-School Science Teachers’ Attitudes Toward Inquiry-Based Teaching.**

<table>
<thead>
<tr>
<th>Item</th>
<th>Statement</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2</td>
<td>I think it is important that the use of inquiry is emphasized in middle-school science education.</td>
<td>A. General preference</td>
</tr>
<tr>
<td>Q3</td>
<td>I am satisfied with the extent to which I use inquiry in my science classes.</td>
<td>A</td>
</tr>
<tr>
<td>Q4</td>
<td>Students learn scientific concepts with deeper understanding through inquiry-based lessons than through lecture type lessons.</td>
<td>A</td>
</tr>
<tr>
<td>Q5</td>
<td>I prefer to teach lecture type lessons rather than to teach lessons in which students engage in hands-on activities.</td>
<td>A</td>
</tr>
<tr>
<td>Q6</td>
<td>In my science classes, it is important to me that students acquire the skills for conducting scientific investigations.</td>
<td>B. Personal evaluation toward inquiry-oriented behaviors</td>
</tr>
<tr>
<td>Q7</td>
<td>In my science classes, it is important to me that students have opportunities to work on open-ended questions whose answers are not known.</td>
<td>B</td>
</tr>
<tr>
<td>Q8</td>
<td>In my science classes, it is important to me that students have opportunities to discover scientific concepts by engaging in hands-on activities.</td>
<td>B</td>
</tr>
<tr>
<td>Q9</td>
<td>It is important to me that students have opportunities to investigate their own scientific questions in my science classes</td>
<td>B</td>
</tr>
<tr>
<td>Q10</td>
<td>It is important to me that students have opportunities to generate their own hypotheses in my science classes.</td>
<td>B</td>
</tr>
<tr>
<td>Q11</td>
<td>It is important to me that students have opportunities to collect data to test their hypotheses in my science classes.</td>
<td>B</td>
</tr>
<tr>
<td>Q12</td>
<td>Variations in students’ experimental results are an important element of discussion in my science classes.</td>
<td>B</td>
</tr>
<tr>
<td>Q13</td>
<td>It is important to me that individual students formulate their explanations from evidence in my science classes.</td>
<td>B</td>
</tr>
</tbody>
</table>
Q14 It is important to me that individual students generate their explanations through discussions in my science classes. B
Q15 In my science classes, it is important to me that students’ explanations are consistent with data they have collected through experiments or observations. B
Q16 In my science classes, it is important to me that students test their explanations against descriptions found in the textbooks or other reliable sources. B
Q17 It is important to me to guide the whole class to achieve understanding of scientific concepts that I teach in a lesson. B
Q18 It is important to me to help students make connections between outcomes from their hands-on activities and scientific concepts in my science classes. B
Q19 It is important to me to explicitly state the scientific concepts I was trying to teach in the lesson at the end of the lesson. B
Q20 It is important to me to help students make connections between scientific concepts and real-life situations during the course of a science unit. B
Q21 It is important to me that students learn to communicate their findings in my science classes. B
Q22 It is important to me that students learn to justify their explanations in my science classes. B
Q23 It is important to me to follow suggestions described in the teacher guidebooks in my science lessons. B
Q24 In my science classes, it is important to me to offer help to individual students that matches their level of understanding of scientific concepts. B
Q25 It is important to me that students are encouraged to explore their own ways of investigating scientific questions in my science classes. B
Q26 It is important to me that students are able to recognize the strengths and weaknesses of the methods they took to explore scientific questions in my science classes. B
Q27 I feel comfortable in handling students’ questions that I cannot answer immediately. D. Teacher’s self-efficacy
Q28 I use students’ questions which I cannot answer as an opportunity for further student explorations. D
Q29 I feel comfortable helping students when they have trouble with their experiments or observations. D
Q30 Curricular demands placed by high stakes exams administered by the state make it difficult for me to use scientific inquiry in my science classes. C. Obstacles
Q31 Students get out of control easily in an inquiry-based classroom. C.
Q32 Assessment of student learning outcomes in inquiry-based teaching is difficult. C.
Q33 Teaching inquiry-based lessons takes too much preparation time. C.
Q34 Teaching inquiry-based lessons takes too much instructional time. C.
Q35 Managing materials for hands-on activities is an issue for me in inquiry-based teaching. C.
Q36 I feel comfortable teaching inquiry-based science lessons to middle-school students. D.
Q37 Inquiry-based teaching is feasible in my middle-school science classes. D.
Q38 What are your thoughts and feelings about inquiry-based teaching? Please write freely.

Annual Colloquium Journal vol. XIV, Spring 2009
Please provide the following information by marking ? in the appropriate box.

<table>
<thead>
<tr>
<th>Q39</th>
<th>Are you male or female?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Female</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q40</th>
<th>How old are you?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20s</td>
</tr>
<tr>
<td></td>
<td>30s</td>
</tr>
<tr>
<td></td>
<td>40s</td>
</tr>
<tr>
<td></td>
<td>50s</td>
</tr>
<tr>
<td></td>
<td>60s</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q41</th>
<th>How many years have you been teaching?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Less than 1</td>
</tr>
<tr>
<td></td>
<td>1-2 years</td>
</tr>
<tr>
<td></td>
<td>3-5 years</td>
</tr>
<tr>
<td></td>
<td>6-10 years</td>
</tr>
<tr>
<td></td>
<td>More than 10 years</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q42</th>
<th>What grade(s) are you currently teaching?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6th</td>
</tr>
<tr>
<td></td>
<td>7th</td>
</tr>
<tr>
<td></td>
<td>8th</td>
</tr>
<tr>
<td></td>
<td>Other: _________________________________</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q43</th>
<th>What subject(s) are you teaching?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Science</td>
</tr>
<tr>
<td></td>
<td>Mathematics</td>
</tr>
<tr>
<td></td>
<td>Both science and mathematics</td>
</tr>
<tr>
<td></td>
<td>Other: _______________________________</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q44</th>
<th>What are the academic degrees you earned during the course of your higher education and what were the fields of study?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st Degree: ___________________________ Field of study: ________________________________________________</td>
</tr>
<tr>
<td></td>
<td>2nd Degree: ___________________________ Field of study: ________________________________________________</td>
</tr>
<tr>
<td></td>
<td>If you have more than 2 degrees, please indicate degree and field: ______________________________________</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q45</th>
<th>What area(s) and level(s) are you certified to teach?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Area: __________________ Level(s): ____________________</td>
</tr>
<tr>
<td></td>
<td>If you are certified to teach more than one area, please specify: _________________________________</td>
</tr>
</tbody>
</table>
Enhancing the Teaching of the Nature of Science

Michael J. Wadness
Medford High School

ABSTRACT
This paper addresses the problem of science literacy, focusing specifically on students’ lack of understanding about the nature of science. In February of 2009, research will be conducted to determine if QuarkNet’s Particle Physics Masterclass provides a fruitful program for students to learn about the nature of science. The Masterclass is a national program where students come to a local area research institute and interact with particle physicists through lectures, informal discussions, and work together to analyze real particle physics data. This paper highlights the research question and methodology for the February, 2009 study.

INTRODUCTION
One of the major goals of K–12 science education is to create scientifically literate citizens. The American Association for the Advancement of Science (AAAS) suggests that a scientifically 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 scientifically literate population by suggesting that science provides the knowledge base for citizens to develop solutions for local and global problems, that science creates a respect for nature, and that science fosters the habits of mind needed in life when making informed, critical decisions. Most importantly, the possible 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 along with a scientifically literate population who understands the decisions that need to be made to support scientists’ work (AAAS, 1989).

There are many scientific issues that have turned into controversial social and political issues. Cloning, stem cell research, pesticide use, global warming, and evolution are all examples of issues that demand an understanding of scientific evidence. All 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 the meaning and importance of being scientifically literate, the case for educating the public on the science that their tax dollars support is strengthened. Science educators have argued that achieving the goal of science literacy among the populous is greatly aided by the teaching of the Nature of Science (referred to as NOS) throughout the K–12 curriculum (Smith & Scharmann, 1999).

During the last three years, one attempt to increase understanding of NOS for high school physics students has been through a program known as the Particle Physics Masterclass. The Masterclass allows students the opportunity to interact with physicists at a research institution where students attend lectures, tour the laboratory, work alongside physicists to analyze real particle physics data, and communicate their findings to each other in a conference-type atmosphere. Specifically, the Particle Physics Masterclass seeks to increase interest in science and to demonstrate the processes of real-life scientific research (Johansson, Kobel, Hillebrandt, Engeln, & Euler, 2007).

PURPOSE OF THE STUDY
The purpose of this study is to evaluate the 2009 U.S. Particle Physics Masterclass to determine if it is indeed an effective program for student learning of NOS. Although research has been conducted regarding student learning of NOS (Irwin, 2000; Scharmann, Smith, James, & Jensen, 2005), and anecdotal claims have been made about science outreach programs aiding in the learning of NOS (Croft, 1999; Abraham, 2002; MJ Young & Associates 2008, July), rigorous research has not been conducted to determine if student exposure to working scientists can aid in the learning of NOS. Specifically, research has not been conducted to determine if the 2009 U.S. Particle Physics Masterclass is an effective program for student learning of NOS.

RESEARCH QUESTION
The following research question will be examined in an effort to determine if the 2009 U.S. Particle Physics Masterclass is an effective program for student learning of NOS.

As a result of attending the 2009 U.S. Particle Physics Masterclass at Fermilab, do secondary science students change their views about the nature of science?
SIGNIFICANCE OF THE STUDY

Although much can be learned from science textbooks and classroom laboratory activities, these two approaches focus almost exclusively on the acquisition of content knowledge and skills. Learning science knowledge and skills may not be enough for today's students to understand and appreciate the scientific enterprise. What is generally studied in high school science classrooms are the finished products of science. Although science content does provide the foundation for science literacy, it may not be enough for students to understand the basis for which scientists make claims. Unfortunately, due to the social controversies surrounding the teaching of Evolution and the Big Bang Theory, there is a need to teach students what scientific knowledge represents and how scientific knowledge is developed. In order for the problem of science literacy to be addressed, there is a need for the development of available programs for the instruction of NOS. Research has shown that although traditional classroom activities teach students about scientific knowledge, they do little to convey the socially constructed nature of science (Lederman, 1999). The information eventually printed in textbooks almost never includes the alternative hypotheses, the anomalous data, the dissenting views, the many confirming studies, and the social processes that occur in discussing the meaning of the observations. What is primarily taught in science classrooms are the final products of the scientific process and not NOS (Abd-El-Khalick, Bell, & Lederman, 1998; Abd-El-Khalick, 2005). Although there are a number of programs involving scientists in K–12 education, very few of them have been formally evaluated to determine if they provide an adequate learning of NOS. Therefore, this study is significant because if student learning of NOS occurs during the 2009 U.S. Particle Physics Masterclass then it will suggest that other science outreach programs may also be a useful context for student learning of NOS.

OPERATIONAL DEFINITIONS

The following operational definitions will be used in this study:

Nature of Science

Although there are many variations on the tenets of the nature of science, the tenets used in this study will be drawn upon NSES (NRC, 1996), Project 2061 Science for all Americans (AAAS, 1989), and from Abd-El-Khalick, Bell, & Lederman (1998). The nature of science describes the epistemology of scientific knowledge and of the scientific enterprise itself. The following will be defined as the tenets of the nature of science.

The Nature of Science includes the nature of scientific knowledge and the nature of the scientific enterprise.

1. Scientific knowledge is tentative yet stable.
   a. Scientific knowledge is subject to on-going testing and revision.
   b. Scientific knowledge occasionally undergoes major revolutionary changes.
   c. Scientific knowledge often encounters minor changes.
   d. New theories that conflict with current accepted theories may encounter harsh criticism.

2. Scientific knowledge is based on observations that are indirect, subjective, and theory laden.
   a. Scientific knowledge attempts to explain and predict observations.
   b. Scientific knowledge is subject to observations.
   c. Until consensus, conflicting theories can result from identical data.
   d. The goal of observations is to be objective yet it is impossible to be completely objective.
   e. The interpretation and measurement of observations is based on existing theories.
   f. Personal, cultural, social, and metaphysical factors can influence the interpretation of data.

3. Science is a social, human enterprise with strengths and weaknesses.
   a. Scientists work alone and in teams but never in isolation.
   b. Scientists publish and evaluate each other's work (peer review) and suggest alternative explanations for observations.
   c. New ideas often encounter harsh criticism.
   d. Scientific knowledge is subject to social and cultural bias.
   e. Research is guided by the availability of funding.
   f. Scientific knowledge is constructed by people who are not significantly different from the general population.


5. There is no universal scientific method.

6. Scientific knowledge incorporates an understanding of the function and relationship between hypotheses, models, laws, and theories.
Learning of NOS

The learning of NOS will be defined as an increase in score from the pre- and posttest instruments used during this study.

REVIEW OF THE LITERATURE

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 incorporates 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 as being a social, human enterprise.

Explicit Instruction in NOS

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; Samarapungavan, Westby, & Bodner, 2006; Scharmann et al., 2005; Sere et al., 2001; Tao, 2003). These studies 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 alone in the control group; NOS instruction was in fact 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 classroom learning 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, et al., 2005; Tsai, 2006) have drawn upon conceptual change theory (Posner, 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 conceptions. In addition, the new concepts must be considered intelligible, plausible, and fruitful. Conceptual change theory also addresses factors that affect students’ resistance to discarding conceptions 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 et al. (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 also 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 merely exposing students to stories without proper explicit instruction, students may not go through the process of conceptual change and the problem of students’ inadequate views of NOS will not be properly addressed.

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 this experience.

Research suggests that other programs in which students interact with scientists have also led to a better understanding of NOS. The programs in which these claims have been made include the exposure of students to scientists through research programs, mentor programs, classroom inquiry projects, guest lectures, and laboratory tours (Sabo, Sarquis, & Ennis, 1997; Croft, 1999; Berkson & Harrison, 2002; Siegel, Mlynarczyk-Evans, Brenner, & Nielsen, 2005). Although not all of the tenets of NOS described were addressed by the above science outreach programs, many have discussed how the interactions between students and scientists affected how students viewed science as a human enterprise and have changed their view of the nature of observations.

Unfortunately many of these outreach programs have questionable evaluation methods, therefore presenting the need for more rigorous research. However, the common ingredient in the above programs is that they provide a context for students to experience firsthand the tenets of NOS. Because the Particle Physics Masterclass provides multiple contexts for explicit learning of NOS, it may provide a suitable vehicle for students to change their view of NOS.

The Particle Physics Masterclass is a program in which students spend one day at a research facility interacting with particle physicists to learn about particle physics. This is accomplished primarily through lectures and an inquiry-based activity where students analyze real data alongside physicists and present their results in a scientific conference environment. In addition, students will tour the research facility to learn how current data is collected and analyzed. The Particle Physics Masterclass is a program run by QuarkNet, in association with Fermilab. QuarkNet is a program partially funded by the National Science Foundation and the U.S. Department of Energy. The goal of QuarkNet is to provide particle physics educational outreach to secondary science students.

The Particle Physics Masterclass

The Particle Physics Masterclass began in 1997 in the U.K. and has expanded to 18 European countries administered at 58 universities and research centers, attracting over 3,000 students (Johansson et al. 2007). In 2005, the Masterclass was administered by the European Particle Physics Outreach Group (EPOG) in an effort to provide opportunities for students to participate in an authentic research process to improve their understanding of scientific research. Specifically, the aims of the Masterclass were to “stimulate interest in science, demonstrate the scientific research process, make data of modern particle physics experiments available for students, and to explore the fundamental forces and building blocks of nature” (p. 637).

To accomplish these goals, students spent the day at a university or research center attending an introductory lecture to learn about particle physics and how particle physicists analyze data from high-energy experiments. Following the lectures, students analyzed data from CERN’s Large Electron Positron Collider, which collected data in the 1990’s outside of Geneva, Switzerland. The students analyzed data alongside the physicists in an attempt to understand Z-decays, using a web-based program similar to what physicists actually use. At the conclusion of the day, an international videoconference occurred between four to six sites at a time to collaborate the students’ data. Physicists at CERN moderated the videoconference in an attempt to provide an environment similar to current international collaborations carried out by research physicists.

To evaluate the effectiveness of the Masterclass, a questionnaire was provided to the students at the conclusion of the 2005 Masterclass. Unfortunately, the authors do not discuss their instrument in detail or make claims of validity or reliability. The sample size consisted of 1,291 high school students from 18 European countries with most students being between 16 and 19 years old. The students’ understanding of particle physics before the program was smoothly distributed with one third of students reporting that they knew little to nothing about particle physics, one third claiming to know something about particle physics, and one third of the students claiming to know very much about particle physics. Of the sample, 69% of the students were male and 31% were female. The majority of the students were either nominated by their teacher to
attend or attended on their own initiative. This implies that the sample may not be representative of the general population.

The authors claim that 82% of the students reported liking the Masterclass or very much liking it and did not find a correlation between previous knowledge of particle physics and enjoyment of the program. In addition, the authors found no significant difference in responses between males and females. Therefore, the authors claim that an understanding of particle physics was not necessary to enjoy the program and was not gender biased. The majority of the students reported that the Masterclass increased their interest in physics and their desire to learn more about particle physics. In regards to NOS, 70% of the students claimed that the Masterclass helped them to learn much or very much about the organization of scientific research. In a correlation of the students’ replies, a weak correlation of .15 was found between students’ appreciation of Masterclass and learning about the organization of research. Interestingly, a moderate correlation of .510 was found between the appreciation of Masterclass and finding the lectures interesting. However, 68% of the students reported the lectures to be easy or very easy to understand. This suggests that although the focus of the design of the Masterclass is in the data analysis portion, the design of the lectures may hold a significant role in students’ reporting their appreciation of the Masterclass.

Although the above study only reports on the students’ perceptions in regards to how Masterclass helped them to learn about the organization of science research, this study is significant because it is related to the nature of science. It would be interesting to investigate what the students’ understanding of the organization of science was before and after the Masterclass, in addition to other tenets of NOS.

In a related study, a U.S. Masterclass was conducted in 2008. The U.S. Masterclass was conducted by QuarkNet, a program partially funded by the National Science Foundation and the U.S. Department of Energy. QuarkNet is a particle physics educational outreach program for high school physics teachers. The format of the U.S. Masterclass was similar to the European Masterclass with the exception of high school physics teachers conducting an hour or more of preparation with their students to provide a basic familiarity of particle physics terminology. In addition, U.S. students were given a tour of a research lab on the Masterclass day (MJ Young & Associates, 2008). The U.S. Masterclass had similar goals to the European Masterclass with the addition of a broader goal of science literacy. An evaluation of the program was conducted by MJ Young and Associates (2008). The purpose of the evaluation was to determine if the goals of the Masterclass were met, to compare U.S. student responses to European student responses from the above mentioned 2005 study, and to determine if the roles of the teachers created any difference in student responses. The instruments used in this study were a questionnaire identical to the one used by the above mentioned 2005 study, a pre- and posttest to assess students’ gains in content knowledge about particle physics and science literacy in the areas of data organization, analysis, and interpretations, and a transmittal form for teachers to report on the students contextual data. The pre- and posttest was reviewed for face validity but was never piloted. Analysis of the data demonstrated that some test items might have been too easy, too difficult, or too confusing.

The sample of students in the U.S. was similar to the European students. In fact, the ratio of males to females was identical to the European numbers, 69% and 31% respectively. The previous knowledge of particle physics was also reported in a similar distribution of approximately one-third of the students claiming to have little or no knowledge of particle physics, one-third having some knowledge, and one-third knew a lot. Slightly more than half of the students reported having taken an advanced science course and 75% reported taking an advanced math class. Therefore, the sample is not representative of the general population.

The questionnaire had 91 students reporting from 6 of the 13 sites. The results of the questionnaire stated that students reported a strong agreement for wanting to learn more about physics, had an increased interest in physics, and learned how science research is organized and carried out as a result of the Masterclass. These results strongly support the European study’s results. However, the responses of the students are questionable due to the fact that, similar to the European study, the questionnaire was administered to the students at the conclusion of the Masterclass and that the questionnaire did not significantly probe the students’ conceptions.

The pre- and posttest used in this study included items related to science literacy including process items and items about how science is constructed. Unfortunately, after analyzing the data, it was determined by the evaluators that some of the questions were too easy, too difficult, or confusing. However, other items involving content demonstrated students having statistically significant gains regardless of previous knowledge. Therefore, the evaluators concluded that all students did benefit in some manner from the Masterclass. It was also noted that a correlation similar
to the European study arose between students’ enjoyment of Masterclass and their understanding of the lectures. This study also suggests that the lectures may provide a significant role in the students reporting their appreciation of Masterclass. Unfortunately, the evaluators reported that there was not enough data collected to determine how the teachers’ instruction before the Masterclass affected the outcomes of Masterclass.

The significance of the above studies suggests the need for further investigation of the Masterclass activity. The studies suggest that students report a greater understanding of how science research is organized, a main item related to NOS. However, there remains a need to determine what students’ specific understanding of NOS is both before and after Masterclass before concluding whether or not Masterclass provides a context for increased understanding of NOS.

**RESEARCH HYPOTHESIS**

The research hypothesis for this study is that secondary science students attending the 2009 U.S. Particle Physics Masterclass at Fermilab will change their views of NOS towards the accepted view as described in this study.

**RATIONALE**

To examine the rationale for this hypothesis, it is important to briefly summarize the research that demonstrates successful learning of NOS. The research reviewed suggested that students are more likely to change their view of NOS towards the comprehensive view, as described in this paper, when explicitly taught within a specific context. The Particle Physics Masterclass at Fermilab provides many experiences, which may provide a context for learning NOS. These include: a tour of Fermilab, introductory lectures within a historical context in particle physics, a data analysis activity, interactions with working research physicists, and a conference for students to present their results. Many of the individual ingredients of the Masterclass may act as a treatment for NOS. One fruitful context for NOS instruction has been through the use of historical examples (Irwin, 2000; Scharmann et al., 2005). Although the Masterclass does not focus on explicit teaching of NOS through historical examples, the introductory lessons do explicitly include the tenets of NOS and present particle physics content and methodology in a historical context. Therefore, the introductory lessons may promote a better understanding of NOS. Research also suggests that programs in which students interact with working scientists have also led to a better understanding of NOS. The programs reviewed in which these claims have been made to affect the students’ views of NOS include the exposure of students to scientists through research programs, mentoring programs, classroom inquiry projects, guest lectures, and laboratory tours (Sabó et al., 1997; Abraham, 2002; Berkson & Harrison, 2002; Siegel et al., 2005). Due to the Masterclass being a combination of fruitful experiences for NOS instruction, it is possible that the Masterclass will also be fruitful for NOS instruction.

Although not all of the tenets of NOS described in this paper were addressed by the above science outreach programs, many have discussed how the interactions between students and scientists affect how students viewed science as a human enterprise. In addition, the above studies have also suggested that students involved in inquiry projects with scientists have also changed their view of the nature of observations. Although a number of the programs involving scientists in secondary education have questioned research methods, the common ingredient to the above programs is that they provide a context for students to experience the tenets of NOS firsthand. Because the Particle Physics Masterclass provides multiple contexts for the explicit learning of NOS, it may provide a suitable vehicle for students to change their view of NOS.

**PARTICIPANTS**

The students participating in this study will be attending the 2009 U.S. Particle Physics Masterclass at Fermilab, outside of Chicago. This Masterclass will consist of approximately 100 students. The type of sampling for this study will be a nonrandomized convenience sample. The Fermilab site will be subdivided into five subgroups consisting of approximately 20 students each for the data analysis and presentation of results. The sampling bias present is that the sample is not representative of the population due to the fact that some of the students attending Masterclass have been nominated by their teacher and some have volunteered. This might indicate that the students may be more advanced than the average high school student.

**METHODOLOGY AND ANALYSIS**

To determine if the 2009 U.S. Particle Physics Masterclass is an effective program for student learning of NOS, a mixed methodology will be employed. The primary methodology will be based on a quasi-experi-
mental time series design (Campbell & Stanley, 1963; Charles & Mertler, 2002). This design includes a single, nonrandomized group of participants observed multiple times over a period of time and exposed to a treatment in an interval between two of the measurements. The design suggests that if a statistically significant change occurs after the introduction of the treatment, then the treatment may be judged to be the cause of that change. This design is diagramed as:

\[
O \ O \ O \ X_1 \ O \ O \ O
\]

In this study however, due to the limitations of the availability of the participants, there will not be as many observations as implied by the above design. Instead, the participants will be observed once before the treatment and twice after the treatment. Therefore, the design of this study is diagramed as:

\[
O \ X_1 \ O \ O
\]

This variation on the quasi-experimental time series design has been used in studies by Chua (2008), Erten and Tekin (2008), and Madden (2006).

The instrument will be a Likert-type survey measuring student understanding of the tenets of NOS. If students show an increase in their score on the instrument then it can be concluded that learning of NOS occurred. In addition, focus group interviews will occur to further probe students’ understanding of NOS and what specific activities had the greatest impact on their understanding of NOS. Lastly, portions of the Masterclass day will be videotaped to analyze interactions between scientists and students in order to observe evidence of student learning of NOS. The analysis will consist of Case II dependent means, one-tailed t-tests between the means on the instrument designed for this study, and grounded theory methods for the focus group interviews.

**Instrumentation**

The instrument under development for this study is a Likert-type survey with open-ended questions designed to evaluate students’ understanding of NOS (see Appendix A). The majority of the items on the instrument were adapted from two different pre-existing instruments by Liang et al. (2008) and Chen (2006). These instruments have previously been tested for validity and reliability. Due to the specific definition of NOS used for this investigation, questions were either discarded, reworded or added. In order to test for validity and reliability, the instrument used for this study will be piloted with local area high school students.

**Focus Group Interviews**

The focus group interviews will be conducted and analyzed using grounded theory methods as described by Charmaz (2006). Approximately 10 students will be selected to participate in the focus group interview. The purpose of the interviews will be to provide a greater understanding of students’ views of NOS towards the end of the Masterclass and how the Masterclass was associated with that view of NOS. A portion of the questions are designed to attempt to elicit if and how the Masterclass has changed the participant’s view of science. The remainder of questions will be used to determine if participants can use specific examples from the Masterclass in expressing their current view of NOS. The majority of the questions are based on the open-ended questions from the survey. However, these interview questions ask participants to use specific examples from the Masterclass as examples to elaborate their view of NOS (Appendix B).

**LIMITATIONS OF THE STUDY**

One significant limitation of this study includes the bias whereby the researcher who is evaluating the Masterclass is also involved with the planning and implementation of the Masterclass. Another limitation of the study is that the students participating in this study are not representative of the general population. As previously discussed, the students participating from the Chicago area have volunteered or have been nominated by their teacher. The nonrandomized convenience sample creates a significant limitation to this study. These students are more likely to be higher achieving students who have a greater interest in science. Thus, the largest threat to the external validity of this study is population validity. Therefore, it will be more difficult to generalize the results of this study to the general population. Although many of the threats to internal validity are addressed by the design of this study, the threats of history, fatigue, and regression remain.

**CONCLUSION**

The purpose of this study is to determine if as a result of attending the 2009 U.S. Particle Physics Masterclass at Fermilab, do secondary science students change their views about the nature of science. The affirmative research hypothesis is based on the Masterclass being a combination of programs that have claimed to be successful for student learning of NOS. In addition, the Masterclass functions as an explicit expe-
rience for student learning of NOS. The methodology employed in this study is a mixed methodology, which includes a variation of a quasi-experimental time series design and focus group interviews. The analysis will consist of Case II dependent means, one-tailed t-tests between the means on the instrument designed for this study, and grounded theory methods for the focus group interviews. Although many of the threats to internal validity are controlled, history, fatigue, and regression are not. In addition, because the sample is not representative of the general population, the threat of population validity exists towards external validity.

Adequate science literacy requires the development of programs for the instruction of NOS. Although there are a number of programs involving scientists in K–12 education, very few of them have been formally evaluated to determine if they provide adequate learning of NOS. Therefore, this study is significant because if student learning of NOS occurs during the 2009 U.S. Particle Physics Masterclass then it will suggest that similar science outreach programs may also be a useful context for student learning of NOS.

References


APPENDIX A
Masterclass Survey
Views on Science

Directions: For the following questions please circle to the best of your ability your position on the following views on science. Use the code below in your responses. This survey may take you 20-30 minutes to complete. We thank you in advance for your time.

<table>
<thead>
<tr>
<th></th>
<th>SD = Strongly Disagree</th>
<th>D = Disagree</th>
<th>U = Unsure</th>
<th>A = Agree</th>
<th>SA = Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Scientific knowledge is subject to on-going revision.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>2. Scientific explanations may be completely replaced by new explanations in light of new evidence.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>3. Scientific explanations may be changed because scientists reinterpret existing observations.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>4. Scientific explanations based on accurate experimentation will not be changed.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>5. Scientists who present evidence that conflicts with established scientific explanations will encounter harsh criticism.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>6. Some scientific explanations may encounter minor changes due to new evidence.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>7. With examples, explain why you think scientific explanations do not change OR how (in what ways) scientific knowledge may change.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Scientists' observations of the same event may be different because the scientists' prior knowledge may affect their observations.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>9. Scientists' observations of the same event will be the same because scientists are objective.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>10. Scientists' observations of the same event will be the same because observations are facts.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>11. Conflicting scientific explanations can result from identical observations.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>12. Scientists' interpretations of data are based on pre-existing scientific explanations.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>13. The existence of atoms is known because atoms have been directly observed.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>14. The existence of atoms is known because of indirect observations that can only be explained by atoms.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>15. With examples, explain why you think scientists' observations and interpretations of the same event could be the same OR different.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Annual Colloquium Journal vol. XIV, Spring 2009
For questions 16-20 use the following definition. *Societal and cultural influences include*: religious, political, economical, and other values and expectations influenced by ethnic and cultural backgrounds.

<table>
<thead>
<tr>
<th>Question</th>
<th>SD</th>
<th>D</th>
<th>U</th>
<th>A</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td>16. Scientific research is not influenced by society and/or culture because scientists are trained to conduct pure, unbiased investigations.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Cultural values determine <strong>what</strong> science is conducted and accepted.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Cultural values determine <strong>how</strong> science is conducted and accepted.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. All cultures conduct scientific research the same way because science is universal and independent of society and culture.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. With examples, explain how society and/or culture affect OR do not affect scientific research.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. Most scientists work by themselves</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Most scientists work in teams</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. It is important for a scientist(s) to be secretive about their work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. It is important for scientists to be critical of each other's conclusions.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. It is important for scientists to share their results.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. With examples, explain why a scientist(s) would or wouldn't share his/her results.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. Scientists use their imagination when making a hypothesis.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>28. Scientists use their imagination when designing an investigation.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>29. Scientists use their imagination when they analyze data.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>30. Scientists use their imagination when they interpret the meaning of data.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>31. Scientists do not use their imagination because this can conflict with their logical reasoning.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>32. Scientists do not use their imagination because this can interfere with objectivity.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>33. Scientists use their creativity when making a hypothesis.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>34. Scientists use their creativity when designing an investigation.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>35. Scientists use their creativity when they analyze and interpret data.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>36. Scientists use their creativity when they interpret the meaning of data.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>37. Scientists do not use their creativity because this can conflict with their logical reasoning.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
</tbody>
</table>
38. Scientists do not use their creativity because this can interfere with objectivity.  

39. With examples, explain how and when scientists use imagination and creativity OR do not use imagination and creativity in the scientific process.

<table>
<thead>
<tr>
<th>40. Scientists follow the same step-by-step scientific method.</th>
<th>SD</th>
<th>D</th>
<th>U</th>
<th>A</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td>41. When scientists use the scientific method correctly, their results are true and accurate.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>42. There is no so-called the scientific method. Scientists use different types of methods to conduct scientific investigations.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
<tr>
<td>43. Experiments are the only means used in the development of scientific knowledge.</td>
<td>SD</td>
<td>D</td>
<td>U</td>
<td>A</td>
<td>SA</td>
</tr>
</tbody>
</table>

Use the following phrase to answer questions 44-47: 
*In comparison to laws, theories have less evidence to support them.*

| 44. Yes, theories are not as definite as laws. | SD | D | U | A | SA |
| 45. Yes, if a theory stands up to many tests it will eventually become a law, therefore, a law has more supporting evidence. | SD | D | U | A | SA |
| 46. Not quite, some theories have more supporting evidence than some laws. | SD | D | U | A | SA |
| 47. No, theories and laws are different types of ideas. They cannot be compared. | SD | D | U | A | SA |

| 48. Approximately how many scientists do you think fit the following characteristics? |
|----------------------------------|--------|---------|--------|--------|
| a. Male                          | None   | Some    | Half   | Most   | All    |
| b. Female                        | None   | Some    | Half   | Most   | All    |
| c. Young                         | None   | Some    | Half   | Most   | All    |
| d. Middle aged                   | None   | Some    | Half   | Most   | All    |
| e. Old                           | None   | Some    | Half   | Most   | All    |
| f. Black                         | None   | Some    | Half   | Most   | All    |
| g. White                         | None   | Some    | Half   | Most   | All    |
| h. Asian                         | None   | Some    | Half   | Most   | All    |
| i. Hispanic                      | None   | Some    | Half   | Most   | All    |
| j. Wear glasses                  | None   | Some    | Half   | Most   | All    |
| k. Facial hair                   | None   | Some    | Half   | Most   | All    |
| l. Clean shaven                  | None   | Some    | Half   | Most   | All    |
| m. Messy hair                    | None   | Some    | Half   | Most   | All    |
| n. Bald                          | None   | Some    | Half   | Most   | All    |
| o. Wears a lab coat              | None   | Some    | Half   | Most   | All    |
| p. Wears a suit                  | None   | Some    | Half   | Most   | All    |

49. What other characteristics or traits do you associate with scientists that may not be included in this list?
APPENDIX B

Focus Group Questions

1. Many of you came to the Masterclass with a preconception about what science is and how it works. Since coming to Fermilab and experiencing the Masterclass, what surprised you about how science really works? Why?

2. Today you heard a few talks on particle physics, toured Fermilab, analyzed data, presented and discussed results, and interacted informally with particle physicists. Out of those items, what made the biggest impact how you view science really works? Why?

3. With examples from today’s Masterclass can you explain why you think scientific explanations do not change OR how (in what ways) scientific knowledge may change?

4. With examples from today’s Masterclass can you explain why you think scientists’ observations and interpretations of the same event could be the same OR different?

5. With examples from today’s Masterclass can you explain how society and/or culture affect OR do not affect scientific research.

6. With examples from today’s Masterclass can you explain why a scientist(s) would or wouldn’t share his/her results?

7. With examples from today’s Masterclass can you explain how and when scientists use imagination and creativity OR do not use imagination and creativity in the scientific process.

8. What types of methods do scientists use besides experimentation to develop new knowledge?

9. What is the difference between a theory and a law?

10. As a result of today’s Masterclass how has your view of a scientist changed?
A Study of the Relationship Between Group Perception of School Climate and Incidences of Bullying at the Junior High/Middle School Level

Kimberly R. Douglass
Northborough Public Schools, MA

ABSTRACT

Bullying at school has become a significant societal issue. Many students in American schools experience victimization during their school careers. Often, the victims are the students who appear different or have a handicap. Bullies make victims feel unsafe in school and as a result, the education of the victims may be compromised, as these students are less available for learning.

This study focused on the potential link between perception of school climate and incidences of bullying at the junior high/middle school level. The primary research question was what, if any, was the relationship between group perception of the type of school climate identified by the surveyed participants and their perceptions of incidences of bullying in a school. One hundred two teachers/specialists along with 186 7th-grade students and their parents from five schools located in the Commonwealth of Massachusetts were surveyed.

The findings suggest that school climate cannot be used to predict whether a school has an issue with bullying. Therefore, the assumption cannot be made that a school that is viewed by teachers/specialists, students, and parents as having a positive school climate has fewer issues with bullying than does a school with a more negative climate. However, the data do suggest that the perceptions among all teachers/specialists, 7th-grade students, and their parents offer valuable information that educators, administration, and teachers can use to further address bullying issues in their schools.

A serious issue facing students in schools worldwide today is bullying.

For two years, Johnny, a quiet 13-year-old, was a human plaything for some of his classmates. The teenagers badgered Johnny for money, forced him to swallow weeds and drink milk mixed with detergent, beat him up in the restroom and tied a string around his neck, leading him around as a ‘pet.’ When Johnny’s torturers were interrogated about the bullying, they said they pursued their victim because it was fun. (Olweus, 2001, p. 24)

The scenario described above is just one of many involving a student who has suffered at the hands of his/her peers. Bullying is the occurrence in which “a student is repeatedly exposed to negative actions on the part of one or more other students” (Olweus, 2001, p. 24). Several of the first studies on bullying were conducted outside of the United States in countries such as Norway, Great Britain, and Sweden (Stein, 2001). For example, in the early 1970s, Dan Olweus, a professor from Norway, first began studying the issue of bullying in Scandinavian countries in response to the suicides of several adolescent boys, which were believed to have been a reaction to victimization (Olweus, 2001). In recent years, however, researchers in the United States have recognized that bullying has become a major issue in American schools (e.g., school shootings, suicides) and have attempted to identify why bullying occurs. For example, four years ago, the Indicators of Crime and Safety 2005 Report indicated that 28% of students, over a six-month period, reported having been bullied while at school (Dinkes, Cataldi, Lin-Kelley, & Snyder, 2007). Educators, therefore, continue to seek and understand contributing causes to the incidence of bullying.

RATIONALE AND IDENTIFICATION OF THE RESEARCH QUESTION(S)

The purpose of this study was to determine if a relationship existed between teachers/specialists, students, and parents’ perceptions of both climate and levels of bullying in schools by surveying teachers/specialists, students, and parents in different schools. For example, did a school deemed welcoming and supportive by polls of the parent community report incidences of bullying? Did a school in which teachers/specialists reported being supported by the administration and connected to the parent community have issues with bullying? Did students who felt supported by the adults and their peers experience less bullying?

The hypothesis of this study was that in junior highs/middle schools that were identified by teachers/specialists, parents, and students as having a more positive school climate, the three groups would also report fewer incidences of bullying than those in junior highs/middle schools with a negative school climate. By establishing a relationship between group perceptions of school climate and incidences of bullying, the belief was that this information could then be used to assist educators in reducing the percentage of bullied students. By identifying differing interpretations of positive and negative environments, educators may be able
to assess what changes can be made within the school setting, and therefore to the school climate, to reduce incidences of bullying.

**RELATIONSHIP TO THE LITERATURE**

**Bullying**

Certain characteristics of a school can perpetuate bullying. One study found that unmonitored halls lead to more incidences of bullying (DeVoe & Kaffenger, 2005). Another study found that when playgrounds and cafeterias were not well supervised, bullying was more likely to occur (Leff, Power, Costigan, & Manz, 2003). When adult supervision is lacking, students are not held accountable for their behavior. Additionally, if there is no observation of the students, the teachers, or supervisors in the school will likely lack direct data about the occurrence of bullying. The Bradshaw study (2007) found that many teachers, while concerned about the occurrence of bullying in their schools, actually underestimate the issue. Staffs report that they are less likely to intervene if an incidence is reported to them rather than when they directly observe the act (Bradshaw, Sawyer, & O’Brennan, 2007). Therefore, students who report bullying may find no assistance. Students, therefore, learn to avoid certain areas of schools in an attempt to avoid victimization (e.g., stairs, parking lots) (DeVoe & Kaffenger, 2005). Supervision in schools appears to be critical to preventing bullying.

Unfortunately, the victims are also usually the students least likely to respond to the bully, usually enduring the bully’s actions (Eslea et al., 2003). Some victims may believe that they deserve to be the target of a bully (Smith, Talamelli, Cowie, Naylor, & Chauhan, 2004). School then becomes a battleground of sorts for these children, as they are either unable to defend themselves against the bullies or resign themselves to the victimization. These students are also the ones who most often lack a strong support system which could assist them in stopping the bullying (Smith et al., 2004). The research suggests that if peers stand by each other then a reduction in the number of bullying incidences will most likely be observed (Eslea et al., 2003). The inability to stop the bullying may foster the perception in these victimized students of an overall school climate that is negative.

Furthermore, as a result of bullying, as noted above, the victims become less engaged in learning as their attention is diverted from academics to violence at school (DeVoe & Kaffenger, 2005). The students must be on guard; they have to develop their own safety plan in order to be less vulnerable, which may foster a negative school climate for these students (Brown, Birch, & Kancheria, 2005). The victims want to escape from the abuse and eventually, school may lose its purpose for the victims.

**School Climate**

Perceptions of individual members and groups within the school are also affected by school climate, which is defined as “the set of internal characteristics that distinguish one school from another and influence the behavior of each school’s member” (Hoy & Miskel, 2005, p. 185). Several studies have been conducted that examine the role of school climate on student performance (Baker, 1999; Dobrasky, 2004; Wilson, 2004). Peer relationships help students establish a sense of belonging (Wilson, 2004). Student-teacher relationships, parent-teacher relationships, and leader influence all impact the overall morale of, or feelings associated with, the school building (Peterson & Deal, 1998). None of these relationships exist in isolation for all have the potential to impact the overall climate of the school, either positively or negatively.

Research has found that students go to adults as a last resort (Rock, Hammond, & Rasmussen, 2002). A sense of community may be lacking for the students.

**METHOD**

**Participants**

The surveyed population consisted of four groups: seventh grade students, their parents, school administrators, and junior high/middle school teachers and specialists. All participants were from the Commonwealth of Massachusetts at the time the surveys were administered.

One hundred eighty-six students from five middle/junior high schools completed the surveys. At the time of the study, they were enrolled as seventh graders. Two factors that were used in selecting this population were that 1) research showed that bullying behaviors decrease as students enter the later junior high/middle school years and 2) that the majority of the students were enrolled in their current school for at least one school year and therefore had more experience with staff and fellow students (Bradshaw et al. 2007; Seals & Young, 2003).

The parent population was defined as the mother, father, or legal guardian of each seventh grade student. The teachers and specialists included sixth, seventh, and eighth grade teachers as well as special subject area teachers and school specialists. Special area teachers
included those who taught art, music, library, computers, physical education, and technology education. The school specialists category included: school guidance counselors, school psychologists, school adjustment counselors, social workers, special education teachers, behavior specialists, occupational therapists, speech and language therapists, and reading specialists. Building administrator referred to either principal or assistant principal.

Over 30 principals throughout the Commonwealth were contacted by e-mail, seeking permission to conduct student, teacher/specialist, and parent surveys in their schools. Five schools granted permission to administer surveys to all three groups.

School Profiles

School One. School One is located in central Massachusetts. Six hundred sixteen students in grades six through eight were enrolled in this suburban school at the time of the study (Massachusetts Department of Education, 2008). The principal is female. The school has both an anti-bullying curriculum and policy. A school safety officer is available on an as-needed basis.

School Two. School Two is a suburban school located in central Massachusetts. Five hundred fourteen students in grades six through eight attended the school at the time of the study (Massachusetts Department of Education, 2008). The principal is female. There is a curriculum that was designed and is implemented by the school staff to address issues of bullying. There is also a policy against bullying. There is no school safety officer.

School Three. School Three is also located in central Massachusetts. Grades six and seven are housed in this suburban school with an enrollment of 413 students at the time of the study (Massachusetts Department of Education, 2008). The principal is female. The school has an anti-bullying policy and an anti-bullying curriculum. The school has a safety officer; however, this person is rarely on site.

School Four. School Four is located in the North Shore area of Massachusetts. The school is considered an urban school. Enrollment at the time of the study consisted of 429 students in grades six through eight (Massachusetts Department of Education, 2008). The principal is a female. While the school has an anti-bullying policy, there is not a curriculum in place to address issues of bullying. A school safety officer is available on an as-needed basis.

School Five. School Five is a suburban school located near Lowell, Massachusetts. Three hundred ninety-three students in grades six through eight attended the school at the time of the study (Massachusetts Department of Education, 2008). The principal is a male. While there is an anti-bullying policy in place, there is not a curriculum program in place that addresses issues of bullying. The school does not have a school safety officer.

Race/ethnicity. Table 1 breaks down the race/ethnicity for each school for the 2007-2008 school year. The data include all students enrolled in the school at the time, not only the seventh graders that participated in the survey. Tables are from the Massachusetts Department of Education Website (2008).

Instrument

Surveys were developed to examine teacher/specialist, parent and students’ perceptions of school climate and incidents of bullying within their school. The questions relating to school climate were developed as a result of the review of the literature of leading researchers in the field including Hoy and Miskel (2005), Peterson and Deal (1998), Fullan (2001), Kelley, Thornton, and Daugherty, (2005), Kallested and Olweus (2003), and Orpinas, Horne, and Staniszewski (2003). The bullying items for the student and teacher/specialist questionnaires were published by Bradshaw et al. (2007). The parent items were an adaptation of the teacher/specialist items.

Teacher/Specialist survey. The questionnaire for the

<table>
<thead>
<tr>
<th>Race (%) of</th>
<th>School 1</th>
<th>School 2</th>
<th>School 3</th>
<th>School 4</th>
<th>School 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>African American</td>
<td>1.6</td>
<td>0.8</td>
<td>1.9</td>
<td>3.5</td>
<td>2.0</td>
</tr>
<tr>
<td>Asian</td>
<td>5.8</td>
<td>8.0</td>
<td>2.2</td>
<td>2.3</td>
<td>12.7</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1.4</td>
<td>1.8</td>
<td>4.4</td>
<td>32.2</td>
<td>3.1</td>
</tr>
<tr>
<td>Native American</td>
<td>0.0</td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>White</td>
<td>90.9</td>
<td>88.0</td>
<td>91.0</td>
<td>59.4</td>
<td>80.7</td>
</tr>
<tr>
<td>Native Hawaiian, Pacific Islander</td>
<td>0.0</td>
<td>0.6</td>
<td>0.2</td>
<td>0.7</td>
<td>0.0</td>
</tr>
<tr>
<td>Multi-Race, Non-Hispanic</td>
<td>0.4</td>
<td>0.5</td>
<td>0.2</td>
<td>1.9</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Table 1. School Populations by Race/Ethnicity

Annual Colloquium Journal vol. XIV, Spring 2009
teachers/specialists also included questions with Likert scale response options on school climate and bullying. Four questions were used to determine how long teachers and specialists had worked in a particular school, their position (i.e., classroom teacher, special area teacher, specialist), their gender, and how many students they currently taught. Eighteen questions were related to school climate. The questions were used to collect data concerning the following:
1. How and how often teachers communicated with parents of students in their classes,
2. How supported the teachers/specialists felt by their colleagues and the building administration,
3. The teacher/specialists’ feelings regarding the climate or tone (e.g., positive, negative) of the school,
4. Assigned duties and responsibilities.

The next 19 questions were directly linked to bullying. The teacher/specialist items were taken from the teacher questionnaire presented in the Bradshaw study (2007). The questions were used to collect data on the following:
1. Prevalence of bullying in their school,
2. Types of bullying,
3. Attitudes towards students who bully,
4. How seriously bullying reports were taken by school personnel,
5. How safe they felt at their school,
6. Their skill in addressing/handling bullying issues.

**Student survey.** The student survey included both open- and close-ended questions and true/false statements. The first nine questions/statements related to school climate. The questions/statements sought to determine the following:
1. How connected to the staff the students were,
2. How available adults were when students needed to seek assistance,
3. How safe the students felt within the school.

The next 14 questions examined issues regarding bullying. The items were taken from the student questionnaire presented in the Bradshaw study (2007). The questions were used to collect data about the following:
1. How students were bullied,
2. The students’ perceptions about bullying,
3. Adult responses/reactions to bullying.

**Parent survey.** The questions for the parent survey examined both school climate and bullying. The questions were all close-ended. Eighteen questions were related to school climate. The questions were used to collect data concerning the following:
1. How welcome parents felt at the school,
2. If parents felt their child's teacher was willing to communicate,
3. How often parents communicated with the school principal,
4. The number of opportunities for parents to become involved in their children's education (e.g., volunteering, attending school functions, reading the school newsletter, parent organization, helping with homework).

There were eighteen questions that related to bullying. They were based on the teacher questionnaire items taken from the Bradshaw study (2007). The questions were used to collect data on the following:
1. Prevalence of bullying in their child's school,
2. Types of bullying,
3. Attitudes towards students who bully,
4. How seriously bullying reports were taken by school personnel,
5. How safe their child felt at school.

Five other questions were included in order to gain a better understanding of 1) whether mothers, fathers, or legal guardians were the more frequent reporters of their child’s victimization, 2) the sex of the parent’s child that was victimized, 3) work schedules of individual parents, and 4) whether the parent was bullied as a child.

**Administrator survey.** In addition, a survey was developed for the school administrator (principal or vice principal) to complete. Twenty-one questions were used to collect information regarding:
1. Exact job title,
2. Years in the position,
3. Any recent changes in administration,
4. If the administration held regular staff meetings,
5. If their school held grade level meetings and if so, did they, as administrators, attend,
6. If there was a school newsletter,
7. If there were all-school meetings,
8. If there was an anti-bullying curriculum and if so, which one was currently being implemented,
9. Supervision procedures,
10. Whether the school had a safety office and, if so, for how many hours a week,
11. For how long incidences of bullying had been tracked,
12. Number of incidences of bullying reported to the administration.
Procedures

Surveys were distributed to the school principals or the person designated as the contact in the school. Student and parents surveys were sent home with the students; envelopes were also provided with the surveys so that students could return both theirs and their parents’ surveys to their homeroom teacher. Teacher/specialists’ surveys were distributed by the principal or contact and an envelope was placed in the main office into which the completed surveys were returned, unless other arrangements were made (i.e., being distributed and completed at a staff meeting). Students in the study who returned the survey and consent form were offered as an incentive of being eligible to receive a gift certificate for a local store such as Borders or Target (one student from each participating school received a gift certificate). (Note: Some school principals expressed that they preferred that an incentive not be offered.)

RESULTS

The computer-based, data program SPSS was used to analyze the survey data.

Confirmatory Factor Analyses

The findings from each confirmatory factor analysis are discussed below. However an exploratory analysis was conducted prior to the confirmatory analyses. The items that factored into each factor during the initial analysis were used to run the confirmatory analysis. Bullying and school climate were the two factors that appeared across all three analyses.

Teacher/Specialist items. The teacher data resulted in six factors with Eigenvalues over 1.0. However, the elbow in the scree plot indicates that the first five components account for most of the variance in the data. When looking at the components, it becomes clear that only four of those identified factors consist of two or more unique items. When examining how the items clustered together, parent contact, school climate, direct experience with student bullying, and adult response to reported bullying seemed most appropriate for the four factor titles.

Student items. The student data resulted in four factors over 1.0. All four of the identified factors consist of two or more unique items. When examining how the items clustered together, direct bullying, school climate, observed bullying, and adult response to bullying seemed most appropriate for the four factor titles. The direct bullying and school climate factors correspond most closely with the factors regarding direct experience with student bullying and school climate in the analysis of the teacher data.

Figure 1. Confirmatory factor analysis teacher/specialist survey scree plot

Figure 2. Confirmatory factor analysis student survey scree plot

Figure 3. Confirmatory factor analysis parent survey scree plot
Parent items. The confirmatory analysis resulted in six factors over 1.0. However, the elbow in the Scree plot indicates that only the first four of those components account for most of the variance in the data. When looking at the components, it became clear that only those ones consist of two or more unique items. When examining how the items clustered together, school bullying, school climate, parent contact, and reasons for bullying seemed most appropriate for the four factor titles. The school climate factor corresponds most closely to the school climate factors from the teacher and student data. The bullying factor corresponds most closely to the direct experience with student bullying factor from the teacher data and the direct bullying factor from the student data.

Table 2. Confirmatory Factor Analysis Teacher/Specialist Survey Factors

<table>
<thead>
<tr>
<th>Reason</th>
<th>Direct experience with student bullying</th>
<th>Adult response to reported bullying</th>
<th>Parent contact</th>
<th>School climate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasons for contacting parents/ Social</td>
<td></td>
<td></td>
<td>.765</td>
<td></td>
</tr>
<tr>
<td>Reasons for parents contacting teachers/Behavior</td>
<td></td>
<td></td>
<td>.625</td>
<td></td>
</tr>
<tr>
<td>Reasons for parents contacting teachers/Social</td>
<td></td>
<td></td>
<td>.869</td>
<td></td>
</tr>
<tr>
<td>Witnessed bullying</td>
<td></td>
<td></td>
<td>.926</td>
<td>.905</td>
</tr>
<tr>
<td>Sense of safety</td>
<td></td>
<td></td>
<td></td>
<td>.919</td>
</tr>
<tr>
<td>Location of incidences/ Classroom</td>
<td></td>
<td></td>
<td>.695</td>
<td></td>
</tr>
<tr>
<td>Location of incidences/ Hallway</td>
<td></td>
<td></td>
<td>.888</td>
<td></td>
</tr>
<tr>
<td>Location of incidences/ Cafeteria</td>
<td></td>
<td></td>
<td>.689</td>
<td></td>
</tr>
<tr>
<td>Location of incidences/ Other</td>
<td></td>
<td></td>
<td>.765</td>
<td></td>
</tr>
<tr>
<td>Type of bullying/Name calling</td>
<td></td>
<td></td>
<td>.854</td>
<td></td>
</tr>
<tr>
<td>Type of bullying/ Pushing</td>
<td></td>
<td></td>
<td>.720</td>
<td></td>
</tr>
<tr>
<td>Type of bullying/Other</td>
<td></td>
<td></td>
<td>.699</td>
<td></td>
</tr>
<tr>
<td>Type of bullying/ Verbal Threats</td>
<td></td>
<td></td>
<td>.693</td>
<td></td>
</tr>
<tr>
<td>Type of bullying/ Exclusion</td>
<td></td>
<td></td>
<td>.741</td>
<td></td>
</tr>
<tr>
<td>Reason/Students’ race or skin color/Student’s race or skin color</td>
<td></td>
<td></td>
<td>.414</td>
<td></td>
</tr>
<tr>
<td>Reason/ Individual differences</td>
<td></td>
<td></td>
<td>.407</td>
<td></td>
</tr>
<tr>
<td>Reason/ Other</td>
<td></td>
<td></td>
<td>.486</td>
<td></td>
</tr>
<tr>
<td>Response to observed incidence/Intervened with bully</td>
<td></td>
<td></td>
<td>.881</td>
<td></td>
</tr>
<tr>
<td>Response to observed incidence/ Ignored it</td>
<td></td>
<td></td>
<td>.897</td>
<td></td>
</tr>
<tr>
<td>Response to observed incidence/Other</td>
<td></td>
<td></td>
<td>.784</td>
<td></td>
</tr>
<tr>
<td>Response to observed incidence/ Referred to school psychologist/ guidance</td>
<td></td>
<td></td>
<td>.704</td>
<td></td>
</tr>
<tr>
<td>Response to observed incidence/ Referred to an administrator</td>
<td></td>
<td></td>
<td>.707</td>
<td></td>
</tr>
<tr>
<td>Response to report/Talked to administrator</td>
<td></td>
<td></td>
<td>.835</td>
<td></td>
</tr>
<tr>
<td>Response to report/Talked to bully’s parents</td>
<td></td>
<td></td>
<td>.860</td>
<td></td>
</tr>
<tr>
<td>Response to report/ Intervened with victim</td>
<td></td>
<td></td>
<td>.860</td>
<td></td>
</tr>
<tr>
<td>Response to report/ Other</td>
<td></td>
<td></td>
<td>.812</td>
<td></td>
</tr>
</tbody>
</table>
Correlations

In order to help establish if perceptions between teachers, parents, and/or students were similar among the groups of surveyed participants the totals for the bullying and school climate factors, established during the factor analysis, were calculated for each participant. These factor scores from each survey type were combined at the group and school level and then correlated.

Correlations at the individual factor level. First the data was analyzed for correlations between groups at the individual factor level. When the data was correlated by group only one correlation existed. Parent and student perceptions of bullying correlated at the .01 level.

Table 3. Confirmatory Factor Analysis Student Survey Factors

<table>
<thead>
<tr>
<th>Direct bullying</th>
<th>Observed bullying</th>
<th>School climate</th>
<th>Adult response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feel welcome</td>
<td></td>
<td></td>
<td>.686</td>
</tr>
<tr>
<td>Like going</td>
<td></td>
<td></td>
<td>.727</td>
</tr>
<tr>
<td>Adults help when asked</td>
<td></td>
<td>.487</td>
<td>- .434</td>
</tr>
<tr>
<td>Caring teacher</td>
<td></td>
<td>.664</td>
<td>- .449</td>
</tr>
<tr>
<td>Adults help handle peer conflict</td>
<td></td>
<td>.554</td>
<td>- .494</td>
</tr>
<tr>
<td>Dislike of school</td>
<td></td>
<td></td>
<td>- .781</td>
</tr>
<tr>
<td>Witness to bullying</td>
<td></td>
<td>.939</td>
<td></td>
</tr>
<tr>
<td>Response to another/Joined in</td>
<td></td>
<td>.928</td>
<td></td>
</tr>
<tr>
<td>Response to another/ignored it</td>
<td></td>
<td>.849</td>
<td></td>
</tr>
<tr>
<td>Response to another/Told an adult</td>
<td></td>
<td>.907</td>
<td></td>
</tr>
<tr>
<td>Response to another/Other</td>
<td></td>
<td>.885</td>
<td></td>
</tr>
<tr>
<td>Response to another/I tried to stop it</td>
<td></td>
<td>.872</td>
<td></td>
</tr>
<tr>
<td>Sense of safety</td>
<td></td>
<td></td>
<td>.728</td>
</tr>
<tr>
<td>Sense of belonging</td>
<td></td>
<td></td>
<td>.821</td>
</tr>
<tr>
<td>Location of incidences/Classroom</td>
<td></td>
<td>.912</td>
<td></td>
</tr>
<tr>
<td>Location of incidences/Hallway</td>
<td></td>
<td>.888</td>
<td></td>
</tr>
<tr>
<td>Location of incidences/Cafeteria</td>
<td></td>
<td>.907</td>
<td></td>
</tr>
<tr>
<td>Location of incidences/Other</td>
<td></td>
<td>.917</td>
<td></td>
</tr>
<tr>
<td>Reason/Student’s race or skin color</td>
<td></td>
<td>.905</td>
<td></td>
</tr>
<tr>
<td>Reason/Student’s gender</td>
<td></td>
<td>.933</td>
<td></td>
</tr>
<tr>
<td>Reason/Looks</td>
<td></td>
<td></td>
<td>.908</td>
</tr>
<tr>
<td>Reason/Other</td>
<td></td>
<td></td>
<td>.916</td>
</tr>
<tr>
<td>Own response/Bullied that person back</td>
<td></td>
<td>.916</td>
<td></td>
</tr>
<tr>
<td>Own response/Did nothing</td>
<td></td>
<td>.922</td>
<td></td>
</tr>
<tr>
<td>Own response/Told an Adult</td>
<td></td>
<td>.886</td>
<td></td>
</tr>
<tr>
<td>Own response/Other</td>
<td></td>
<td>.878</td>
<td></td>
</tr>
<tr>
<td>Proactive approach to bullying</td>
<td></td>
<td></td>
<td>-.558</td>
</tr>
<tr>
<td>Adults ignoring bullying</td>
<td></td>
<td></td>
<td>.688</td>
</tr>
<tr>
<td>Bullying worsened with adult intervention</td>
<td></td>
<td></td>
<td>.593</td>
</tr>
<tr>
<td>No response to request for help</td>
<td></td>
<td></td>
<td>.606</td>
</tr>
</tbody>
</table>
Table 4. Confirmatory Factor Analysis Parent Survey Factors

<table>
<thead>
<tr>
<th></th>
<th>Bullying</th>
<th>Reasons for bullying</th>
<th>School climate</th>
<th>Parent contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Involved with PTO</td>
<td></td>
<td></td>
<td></td>
<td>.751</td>
</tr>
<tr>
<td>School Council membership</td>
<td></td>
<td></td>
<td></td>
<td>.450</td>
</tr>
<tr>
<td>Contact with administration</td>
<td></td>
<td></td>
<td></td>
<td>.800</td>
</tr>
<tr>
<td>Child reported seeing bullying</td>
<td>.690</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child self reported bullying</td>
<td>.544</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sense of safety</td>
<td></td>
<td></td>
<td></td>
<td>.848</td>
</tr>
<tr>
<td>Sense of belonging</td>
<td></td>
<td></td>
<td></td>
<td>.905</td>
</tr>
<tr>
<td>Location of incidences/Classroom</td>
<td>.904</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location of incidences/Hallway</td>
<td>.892</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location of incidences/Cafeteria</td>
<td>.887</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location of incidences/Other</td>
<td>.886</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of bullying/Name calling</td>
<td>.834</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of bullying/Pushing</td>
<td>.827</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of bullying/Verbal Threats</td>
<td>.863</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of bullying/Other</td>
<td>.874</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reason/Student's race or skin color</td>
<td>.887</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reason/Student's gender</td>
<td>.877</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reason/Physical differences</td>
<td>.888</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reason/Other</td>
<td>.864</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response/Talked to administrator</td>
<td>.473</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response/Talked to bully's parents</td>
<td>.519</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response/Talked to school psychologist, behavior specialist or guidance counselor</td>
<td>.516</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response/Talked to my child's teacher</td>
<td>.459</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response/Talked to my child</td>
<td>.595</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response/Other</td>
<td>.545</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Correlations for Bullying Factor by Group

<table>
<thead>
<tr>
<th></th>
<th>Student</th>
<th>Parent</th>
<th>Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>1.000</td>
<td>.615**</td>
<td>.156</td>
</tr>
<tr>
<td>Parent</td>
<td>.615**</td>
<td>1.000</td>
<td>.124</td>
</tr>
<tr>
<td>Teacher</td>
<td>.156</td>
<td>.124</td>
<td>1.000</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed).

Table 6. Correlations for School Climate Factor by Group

<table>
<thead>
<tr>
<th></th>
<th>Student</th>
<th>Parent</th>
<th>Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>1.000</td>
<td>.039*</td>
<td>-.106</td>
</tr>
<tr>
<td>Parent</td>
<td>.039</td>
<td>1.000</td>
<td>-.088</td>
</tr>
<tr>
<td>Teacher</td>
<td>-.106</td>
<td>-.088</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Correlations between the school climate and bullying factors by school. The next part of the data analysis sought to establish any correlations between perceptions of bullying and school climate at the individual school level. A correlation between perceptions of school climate and bullying existed at both the student (.01 level) and teacher (.05) levels for School 5 only. No other significant correlations existed among the groups.
A correlation between parents’ perceptions of school climate and bullying was identified only for School 5. This correlation is significant at the .05 level. Correlations between the school climate and bullying factors by group. When the bullying and school climate factors were correlated by group, only two correlations existed. There was a correlation between the bullying and school climate factors for students (.01 level) and parents (.05 level).
Table 10. Correlation Between Bullying and School Climate Factors for Teachers

<table>
<thead>
<tr>
<th>Bullying</th>
<th>School climate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bullying</td>
<td>1.000</td>
</tr>
<tr>
<td>School climate</td>
<td>-0.082</td>
</tr>
</tbody>
</table>

Table 11. Correlation Between Bullying and School Climate Factors for Students

<table>
<thead>
<tr>
<th>Bullying</th>
<th>School climate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bullying</td>
<td>1.000</td>
</tr>
<tr>
<td>School climate</td>
<td>-0.267**</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed).

Table 12. Correlation Between Bullying and School Climate Factors for Parents

<table>
<thead>
<tr>
<th>Bullying</th>
<th>School climate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bullying</td>
<td>1.000</td>
</tr>
<tr>
<td>School climate</td>
<td>-0.246*</td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 level (2-tailed).

Conclusion. Perceptions about school climate and incidences of bullying are not similar when teacher, student, and parent responses are correlated. Only isolated correlations existed. The data does not support the earlier hypothesis that schools which are perceived to have more positive school climates are also perceived to have less bullying incidences.

Means

Mean scores by group. Teacher/specialist, student, and parent perceptions of school climate and bullying differ across groups and schools. When the mean scores for each group were examined, teachers identified the climate as more positive than the other groups and identified bullying as more of an issue than did students and parents. In each school, the groups had differing views of the levels of bullying and school climate when their responses were compared to the others schools. For example, in School One, parents saw bullying as more of an issue than students and teachers. Students saw school climate as more negative than did their parents and teachers/specialists. In Schools Two and Three, parent responses indicated that they felt that their child's schools had a more positive school climate and less incidences of bullying. In School Three, however, teachers/specialists viewed the school climate as more negative than did the parents and students. In School Four, parents saw the school as having a poorer school climate with higher amounts of bullying. Teachers saw bullying as an issue but not the school climate, whereas for students the reverse was true. (Note: Incidence reports indicated that this school reported the highest number of bullying incidences per school year compared to the other four schools.) In School Five, parents saw bullying as less of an issue than did teachers/specialists and students. However, students saw the school climate as being better than did their teachers/specialists and parents. Therefore, it is important for administrators to be cognizant of the fact that how staff views the school climate and level of bullying is not necessarily the same as how students and parents perceive them. Attempting to change either would require input from all three groups.

Mean scores by individual items. When means were collapsed by group (i.e., teachers/specialists, students, and parents), information was learned that could help administrators address school climate and bullying issues. When analyzed, individual group perceptions of each may impact how bullying and school climate issues are addressed.

The teacher/specialist data revealed that the majority of the surveyed educators have witnessed bullying. Bullying was more likely observed in the hallways than in other school locations. Also the most frequent types of bullying were name calling and exclusion. Teachers were likely to intervene during an observable bullying situation as well as provide assistance to the victim when a situation was reported.

The student data revealed that most students feel welcome at their respective schools; however, that did not correspond with whether they like going to school. Most students felt that adults are supportive and help when asked. Most students either tried to stop bullying when they witnessed it or did not get involved in the situation.
Table 13. Means for Teachers/Specialists

<table>
<thead>
<tr>
<th>Reason for contacting parents/Social</th>
<th>Highest point value per item</th>
<th>Statistic</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasons for contacting parents/Social</td>
<td>1.0</td>
<td>.4020</td>
<td>.04879</td>
</tr>
<tr>
<td>Reasons for parents contacting teachers/Behavior</td>
<td>1.0</td>
<td>.4118</td>
<td>.04897</td>
</tr>
<tr>
<td>Reasons for parents contacting teachers/Social</td>
<td>1.0</td>
<td>.4412</td>
<td>.04941</td>
</tr>
<tr>
<td>Witnessed bullying</td>
<td>1.0</td>
<td>.8137</td>
<td>.03874</td>
</tr>
<tr>
<td>Sense of safety</td>
<td>3.0</td>
<td>2.6765</td>
<td>.07262</td>
</tr>
<tr>
<td>Sense of belonging</td>
<td>3.0</td>
<td>2.6139</td>
<td>.07171</td>
</tr>
<tr>
<td>Location of incidences/Classroom</td>
<td>2.0</td>
<td>1.2529</td>
<td>.07167</td>
</tr>
<tr>
<td>Location of incidences/Hallway</td>
<td>2.0</td>
<td>1.6782</td>
<td>.07402</td>
</tr>
<tr>
<td>Location of incidences/Cafeteria</td>
<td>2.0</td>
<td>1.3103</td>
<td>.07368</td>
</tr>
<tr>
<td>Type of bullying/Name calling</td>
<td>2.0</td>
<td>1.0114</td>
<td>.05480</td>
</tr>
<tr>
<td>Type of bullying/Pushing</td>
<td>2.0</td>
<td>.9796</td>
<td>.03547</td>
</tr>
<tr>
<td>Type of bullying/Other</td>
<td>2.0</td>
<td>1.6837</td>
<td>.06085</td>
</tr>
<tr>
<td>Type of bullying/Verbal Threats</td>
<td>2.0</td>
<td>1.1477</td>
<td>.06371</td>
</tr>
<tr>
<td>Type of bullying/Exclusion</td>
<td>2.0</td>
<td>1.0000</td>
<td>.05423</td>
</tr>
<tr>
<td>Reason/Students' race or skin color</td>
<td>2.0</td>
<td>1.0102</td>
<td>.03971</td>
</tr>
<tr>
<td>Reason/Individual differences</td>
<td>2.0</td>
<td>1.6556</td>
<td>.07605</td>
</tr>
<tr>
<td>Response to observed incidence/Intervened with bully</td>
<td>2.0</td>
<td>1.1477</td>
<td>.06371</td>
</tr>
<tr>
<td>Response to observed incidence/Ignored it</td>
<td>2.0</td>
<td>.8889</td>
<td>.04311</td>
</tr>
<tr>
<td>Response to observed incidence/Other</td>
<td>2.0</td>
<td>.8889</td>
<td>.04311</td>
</tr>
<tr>
<td>Response to observed incidence/Referred to school psychologist/guidance</td>
<td>2.0</td>
<td>1.1889</td>
<td>.07223</td>
</tr>
<tr>
<td>Response to observed incidence/Referred to an administrator</td>
<td>2.0</td>
<td>1.1556</td>
<td>.06872</td>
</tr>
<tr>
<td>Response to report/Talked to administrator</td>
<td>2.0</td>
<td>1.2323</td>
<td>.08351</td>
</tr>
<tr>
<td>Response to report/Talked to bully's parents</td>
<td>2.0</td>
<td>.8485</td>
<td>.05608</td>
</tr>
<tr>
<td>Response to report/Intervened with victim</td>
<td>2.0</td>
<td>1.3200</td>
<td>.08514</td>
</tr>
<tr>
<td>Response to report/Other</td>
<td>2.0</td>
<td>.9100</td>
<td>.06371</td>
</tr>
</tbody>
</table>

Table 14. Means for Students

<table>
<thead>
<tr>
<th>Reason</th>
<th>Highest point value per item</th>
<th>Statistic</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feel welcome</td>
<td>3.0</td>
<td>2.4239</td>
<td>.04459</td>
</tr>
<tr>
<td>Like going</td>
<td>3.0</td>
<td>1.8478</td>
<td>.05551</td>
</tr>
<tr>
<td>Adults help when asked</td>
<td>3.0</td>
<td>2.3333</td>
<td>.04921</td>
</tr>
<tr>
<td>Caring teacher</td>
<td>3.0</td>
<td>2.2787</td>
<td>.05163</td>
</tr>
<tr>
<td>Adults help handle peer conflict</td>
<td>3.0</td>
<td>2.3261</td>
<td>.05008</td>
</tr>
<tr>
<td>Dislike of school</td>
<td>3.0</td>
<td>.6831</td>
<td>.05737</td>
</tr>
<tr>
<td>Witness to bullying</td>
<td>1.0</td>
<td>.5055</td>
<td>.03797</td>
</tr>
<tr>
<td>Response to another/Joined in</td>
<td>2.0</td>
<td>.5394</td>
<td>.04078</td>
</tr>
<tr>
<td>Response to another/Ignored it</td>
<td>2.0</td>
<td>.7879</td>
<td>.06535</td>
</tr>
<tr>
<td>Response to another/Told an adult</td>
<td>2.0</td>
<td>.5879</td>
<td>.04786</td>
</tr>
<tr>
<td>Response to another/Other</td>
<td>2.0</td>
<td>.5697</td>
<td>.04567</td>
</tr>
<tr>
<td>Response to another/I tried to stop it</td>
<td>2.0</td>
<td>.7333</td>
<td>.06146</td>
</tr>
<tr>
<td>Sense of safety</td>
<td>3.0</td>
<td>2.3405</td>
<td>.04834</td>
</tr>
<tr>
<td>Sense of belonging</td>
<td>3.0</td>
<td>2.3027</td>
<td>.05057</td>
</tr>
</tbody>
</table>
More parents reported that their child had witnessed bullying than was involved in a bullying incidence. Parents felt that their children felt safe and welcome at school. They reported the cafeteria as being the location of most incidences of bullying. Name calling was the most frequent form of bullying. More parents responded to bullying by talking to their child rather than contacting the school.

**DISCUSSION**

**Limitations**

In reviewing the survey development, process for school selection, participants, survey administration, and item responses, certain limitations were identified. Identifying these limitations prior to discussing the implications of the knowledge learned from the survey is important because different conclusions may be drawn from future studies when these research limitations are addressed.

**School selection.** The first issue that may have affected the research results was school selection. The schools chosen to participate in the survey were not randomly selected. Rather, the Massachusetts Department of Education website and professional connections were used as resources to recruit the schools. Over 30 school districts were contacted to participate in the study and only five actually took part in the study. Thus, the principals and superintendents who chose to have their schools participate in the study may have had positive preconceptions about their school climate and rates of bullying and as a result felt more comfortable participating. Also the school administrators that were involved in the study each had personal knowledge of the University of Massachusetts Lowell doctorate program and/or this particular research study, and the school administrators may have had more confidence regarding the confidentiality that would be afforded to the survey results which may have made them more likely to participate than other schools with negative school climates and higher incidences of bullying.

**Timelines.** The second limitation identified with the research methodology was the timeline in which the schools received the surveys. Difficulty with identifying
sites and securing agreement from the schools to participate impacted overall survey distribution. The surveys were not all distributed to the schools at the same time. Students and parents who took the surveys at a later date, therefore, had more opportunities for experiences, positive or negative, with the school. The results, therefore, do not capture a snapshot of bullying and school climate at the same time within all five schools.

Respondents. Only five schools participated in the study. Of those five, four were suburban and one urban. No rural schools agreed to participate. The majority of the schools were located in what is considered the Metro-West portion of Massachusetts. Massachusetts as an entire commonwealth is not well represented in the study. Additionally, the sample size is small. There are 328 middle/junior high schools in the commonwealth (Massachusetts Department of Education, 2008) and only five participated in this study. However, it is important to note that, 102 teachers/specialists along with 186 7th-grade students and their parents did participate in the study for a total of 474 participants. Each group’s (i.e., all teachers/specialists, students, and parents) responses offered valuable information regarding perceptions of school climate and incidence of bullying for those groups. School climate and bullying may be linked when a wider representation of the population is sampled.

Some parents also may have not received the surveys as the students were given the survey at school and asked to take it home to their parents. If the parents did not receive the surveys then their children did not complete them either as written consent was required from all respondents.

Race/Ethnicity. Based on the population of the schools, the surveys could have been translated into Spanish, Portuguese, and Asian languages in order to ensure that the survey was accessible to and representative of all ethnicities. No data were collected regarding race of respondents so it is unknown if most respondents were white or of other ethnic backgrounds. This gap in data collection may have skewed the results when examining the impact school climate has on incidences of bullying.

Survey development. One limitation of the survey development had to do with different items across surveys. In some cases, like worded questions had differing number of responses. For example, in the parents and student surveys, the question that asked if bullying occurred in the cafeteria had three possible answers on the student survey (i.e., “I have not been bullied,” “No,” and “Yes”) yet four possible answers on the parent survey (i.e., “Not applicable,” “I don’t know,” “Yes,” and “No). Also, the consistency of questions across the three surveys had discrepancies. That is, questions that could have applied to all groups were not asked on all surveys. For example, the item, “I feel welcome at my school,” could have been asked of all three groups but was only asked on the student survey.

Survey administration. Envelopes were included with each survey packet in order to return the surveys and permission slips as the parents’ permission was required for the students to participate. By returning both documents in the same envelope, the author would be able to link the respondent with the responses although such links were to be held confidential and would not be revealed as part of the published results. However, a number of parents complained to their respective principal that the survey was not truly anonymous as their and their children’s responses could be linked to their names despite assurances that this would not happen. Response numbers were low and may have been due to this inability to create a completely anonymous survey because of the age of the students and the need to get the parents’ permission.

Also it is unclear how many parents actually received the survey. As mentioned above, students were given the survey and asked to take it to their parents; as the students were the messengers, they may not have delivered the envelopes to their parents. Likewise, some parents may have completed the survey and their child did not return the envelope to school. Thus, the actual number of parents that received the survey is unknown and the number of parents who did receive the survey but chose not to participate is also not known.

Data Sharing at the Individual School Level

It is important that school administrators have an understanding of how school climate is perceived by staff, parents, and students along with how the issue of the bullying is perceived within the community. Ultimately, the frequency of responses will be shared with the individual school principals that participated in the surveys used for this research so that they have access to this information. The expectation is that by knowing how teachers/specialists, students, and parents responded to the survey items, then principals can further investigate practices in their school buildings that might lead to change regarding the prevalence of bullying. Below are two examples:

For example, in School One, 55.3 percent of the students reported seeing someone bullied in the previous month. Additionally, when asked if the students had seen adults observe bullying without intervening, 38.3 percent answered “yes.” Over 10% of the same stu-
dents polled shared that they had reported bullying to an adult and received no assistance. The survey results from School One also revealed that more students reported being bullied while in a classroom than in the cafeteria and hallway.

In School Three, 19.1 percent of the parents surveyed would not report an incidence to school staff if he or she heard about his or her child being bullied. When asked if their child had reported adults watching and not responding to a bullying incidence, 12.8% responded “yes.” When asked about bullying incidences, 59.6% of the School 3 parents responded that their children had seen another student being bullied and 30.4% responded that their children had been bullied. While these same parents reported positive school climate at School 3, the fact that parents were aware of bullying occurring at their children’s school and yet would not report bullying to school staff is, while not statistically correlated, still another interesting observed response on why and how bullying occurs.

**Instruments for Schools to Assess Perceptions of School Climate and Bullying**

The three surveys used for this research could be administered in any middle/junior school in order to assess perceptions of school climate and bullying. As a school psychologist it is important to develop an understanding of how members view their school. For example, if a high percentage of students and their parents feel that staff do not readily respond to bullying incidences then how can that be addressed? If students do not feel welcome at school, what practices may be implemented in order that the students develop a sense of belonging? By further developing an understanding of perceptions of school climate and bullying, then incidences of bullying may be able to be lessened.

**CONCLUSION**

Although this study did not indicate a link between perceptions of school climate and incidence of bullying, bullying is clearly an important issue in these schools and thus is a topic that must continue to be taken seriously. Educators need to be committed to ensuring that all students in a classroom/school are treated respectfully. Professional development regarding conflict resolution skills must also be provided to educators. Teachers need to be taught how to teach conflict resolution skills to their students so that the students have a skill set to use when responding to a bully. Teacher and administrator attitudes towards students also impact how a student is received within the school setting. Students sense if the adults in the school do not regard a peer in a favorable light. Training to assist teachers in recognizing their own reactions to/bullying towards their students may also be needed. The provision of support to teachers and students by administration and student support teams is critical in aiding students. Teachers should not feel that they have to deal with a challenging student alone. Students should be encouraged to tell if they are being bullied. All students are entitled to an education and an educational environment in which they are free from harm.

Consideration should be given to making the changes discussed in the Limitations section of this chapter and rerunning the study in order to further understand potential links between school climate and incidences of bullying. More respondents from a wider pool (i.e., ethnicity, socio-economic background, geographic location) may yield different results. Additionally, data sharing with the principals from each school that participated in the survey will also hopefully assist with identifying and eliminating the bullying that is occurring in the surveyed schools. While this study may ultimately impact students in five schools, until bullying is addressed and controlled within schools nationwide, all angles of bullying must continue to be examined.

**References**


Qualifying Papers

Increasing the Effectiveness of Mathematics Instruction Through Lesson Planning

Noreen C. Flanagan Johnson
Greater Lawrence Regional Vocational Technical High School, Andover, MA

ABSTRACT

It is imperative that mathematics teachers make their instruction more effective. This paper argues that comprehensive lesson planning is crucial for effective teaching. A review of the literature identifies three traits of effective teaching: (1) teachers' level of mathematical content knowledge, (2) class instruction developed explicitly on student learning objectives, and (3) focus on student prior knowledge. A definition of comprehensive lesson planning is provided, and research is reviewed into teachers planning habits, the decisions teachers make when planning, and planning in countries outside the United States. A brief summary of common lesson planning methods is provided. Finally, research is identified that shows a possible link between one specific method of lesson planning and increased effectiveness of mathematics instruction.

INTRODUCTION TO THE PROBLEM

Mathematics is a fundamental part of every student's education, with many hours devoted specifically to the development of mathematical knowledge. Yet, a great number of students in the United States leave school with only a superficial understanding of mathematics concepts (Stigler & Heibert, 1999). That so much effort and time on the part of the teacher and the student is used to little or no avail should be a concern to all educators. Not only will a mathematically illiterate population be at a disadvantage in our increasingly technologically-advanced society, students' failure to learn mathematics signals a greater failure in our nations' schools.

Data published by the Massachusetts Department of Education for the school year 2004–2005 (Massachusetts Department of Education, 2006) show that a significant number of sixth- and eighth-grade students achieve a score of Warning, essentially a failing grade, on the mathematics portion of the Massachusetts Comprehensive Assessment System (MCAS). Of these students, those who go on to repeat at that grade level do not improve enough to achieve a level of Needs Improvement when tested again at the end of the repeated year. The extra year of instruction given to students fails to provide the desired outcome of improved student achievement. This may, in part, be a direct result of the ineffectiveness of the classroom instruction provided to the students. It is imperative that classroom teachers find a way to make their mathematics instruction more effective.

Since the early 1980s, there has been a persistent shortage in the number of qualified mathematics and science teachers. Because of this shortage, many efforts have been made to bring more qualified mathematics and science teachers into the profession. The No Child Left Behind federal legislation includes several provisions for the recruitment of mathematically competent individuals to the teaching profession through alternative licensure routes (No Child Left Behind [NCLB], 2002). In Massachusetts, the Massachusetts Department of Education has provisions in its licensure procedures for those teachers who do not take a traditional path to become mathematics educators. As Laczkó-Kerr and Berlinger (2003) note, “research provides convincing evidence that subject-matter knowledge is necessary but not sufficient for teaching well” (p. 35). While it is imperative that all mathematics classrooms include a mathematically competent teacher, this rush to certify nontraditional applicants has created a situation where many teachers enter the classroom with little or no training in education. As a consequence, new math teachers may have little knowledge of how to plan lessons, let alone plan effective lessons that focus on the use of inquiry methods as recommended by NCTM (2000). Perplexingly, those teachers who do complete traditional teacher preparation programs are reported to cease planning in the way that they were instructed after leaving those programs (Searcy & Maroney, 1996). It is a common assumption that the longer one stays in the teaching profession, the less time one needs to devote to lesson planning. However, this paper will argue that consistent and comprehensive lesson planning is crucial for effective teaching.
THE PROBLEM

The problem that this paper will address is the failure of many middle and high school students to leave the classroom with more than a superficial understanding of mathematical concepts; in effect, they fail to learn mathematics. It is suggested here that the primary factor that influences student learning in the classroom is the effectiveness of the instruction provided to the student. It is an oft-stated fact that most teacher-preparation programs include methods classes in which preservice teachers are expected to plan lessons (for example, see Reiser & Mory, 1991). However, research shows that teachers either stop formal lesson planning altogether or do not continue to plan in the manner which they were taught (Searcy & Maroney, 1996). Alternately, those teachers who have not completed a traditional teacher preparation program may have little to no knowledge of lesson planning methods when they enter the classroom.

If the state of instruction in secondary mathematics classrooms is to be systematically improved, then it is imperative to assess what teaching activities and/or techniques are most effective at improving mathematics instruction. Whether or not comprehensive lesson planning can be linked to behaviors consistent with effective teaching is the purpose of this review. This topic is important because of the plethora of factors that can affect the ability of students to become proficient in mathematics. Aspects such as the students’ home lives, socioeconomic statuses, societal influences, and cognitive abilities, however, are nearly impossible for teachers to change. Therefore, these factors will not be discussed in this paper. It is primarily through the improvement of classroom instruction that teachers, administrators, and teacher educators can have the most impact on the teaching and, hopefully, the learning that takes place in the classroom.

PERTINENT AREAS OF RESEARCH

This paper consists of two strands. The first details several characteristics of instruction identified in the literature to be effective. The second strand reviews the relevant research on lesson planning. Specifically, it searches for evidence of ways in which teachers might improve their instruction through comprehensive lesson planning.

REVIEW OF RESEARCH LITERATURE

This review of the literature will focus on the argument that instructional effectiveness in the mathematics classroom can be improved through comprehensive lesson planning. The first part of this review will focus on three characteristics of effective teaching identified by the literature. The importance of teachers’ content knowledge, instruction based on student learning objectives, and students’ prior knowledge will be outlined. Research will be presented that links these characteristics to increased instructional effectiveness. A discussion on lesson planning will be presented in the second part of the review. The analysis will focus on the decisions teachers make when they plan, any common themes in the types or amounts of planning teachers engage in, and several common lesson planning formats currently available to teachers. Finally, research will be reviewed that points to a possible relationship between a specific model of planning for mathematics lessons and increased instructional effectiveness in the classroom.

CHARACTERISTICS OF EFFECTIVE MATHEMATICS TEACHING

“The teaching profession does not have enough knowledge about what constitutes effective teaching” (Stigler & Hiebert, 1999). Many different factors have been put forth to explain what constitutes effective teaching. Those that will be addressed in this paper are teachers’ level of mathematical content knowledge, class instruction developed explicitly on student objectives for learning, and teachers taking student prior knowledge into account during instruction. For the purpose of this paper, the personality traits of teachers that have been linked to levels of teacher effectiveness will not be discussed.

According to Mager (1975), “instruction is successful or effective to the degree that it accomplishes what it sets out to accomplish” (p. 1). This definition, though vague, does illustrate the ongoing argument over the notion of what constitutes effective teaching. The current NCLB legislation has placed a significant emphasis on student scores on standardized exams, therefore equating effective teaching with student achievement scores. Morine-Dershimer (1977) and Good and Grouws (1979), among others, have conducted research that attempts to link teacher effectiveness with an increase in students’ scores on various types of tests. Other educators may wish to define student learning in other ways, taking, for example, student participation and attitudes towards learning into account. For this paper, the precise measurement of student learning is not as important as is identifying the elements that many different researchers and educators can agree are shared by effective teachers.
Cruickshank, Bainer, and Metcalf (1999) argue that, in order to teach effectively, teachers must “organize their lesson content and activities logically” (p. 347). This requires teachers to have strong content knowledge of the subject they teach. Cruickshank et al. continue that effective teaching must “inform students of the objectives of the lesson,” and also “determine the most logical way to introduce content based on their students’ abilities, previous learning, and the natural structure of the content” (p. 347). As such, two additional characteristics of effective teaching have been established—student learning objectives must be identified, as well as the prior knowledge students need in order to achieve those objectives. It is feasible to conclude that classroom instruction which includes all three of these characteristics will be more effective than instruction that is lacking one or more of the identified traits. It is the literature on these three characteristics of effective teaching that will make up the first part of this paper.

Teachers’ Content Knowledge

“To be effective, teachers must know and understand deeply the mathematics they are teaching” (NCTM, 2000, p. 17). It seems obvious that to teach mathematics one needs to possess the requisite content knowledge. Yet, research indicates that approximately half of eighth grade math classes in the United States are taught by teachers who did not major in mathematics (National Center for Educational Statistics, United States Department of Education, 2000) and one-third of high school mathematics teachers do not possess a degree in mathematics, math education, or a related field (Ingersoll, 2001). While all teachers are now required to be highly qualified under the No Child Left Behind legislation (NCLB, 2002), this is far from reality. The Massachusetts Department of Education allows teachers to teach a maximum of two classes per day outside of their area of licensure, and in extreme cases provides a one-year waiver that allows uncertified, and possibly unqualified, teachers to teach a majority of their daily schedule in a subject for which they do not hold a valid teaching license. This process is in direct contradiction to some researchers’ findings (Fowler & Poetter, 2004; Manouchehri & Goodman, 2000; Wenglinsky, 2002), which indicate that a teacher’s level of content knowledge is a primary determinant of effective teaching.

The link between content knowledge and effective teaching was explored by Wenglinsky (2002) using data from the 1996 National Assessment of Educational Progress (NAEP) in mathematics. A nationally representative sample, consisting of 7,146 eighth-grade students from the United States, was investigated to determine which aspects of classroom instruction impact student achievement. It was found that, when controlling for class size and socioeconomic status, a teacher’s major in college was “modestly associated with academic achievement” (p. 18). Further support for the link between content knowledge and effective teaching comes from a case study of two seventh-grade mathematics teachers conducted over seven months. Manouchehri and Goodman (2000) looked at the teacher characteristics that affected the successful implementation of a new, reform-based mathematics textbook. The researchers concluded that a teacher’s content knowledge, and by extension, the teacher’s ability to see connections between mathematical concepts, is at the heart of effective instruction.

The Trends in International Mathematics and Science Studies (TIMSS) alerted educators in the United States to the problems with mathematics education in this country (National Center for Education Statistics, 2000). As a result, researchers have examined why other countries appear to be succeeding where we are failing. Fowler and Poetter (2004) examined why students in France scored much higher than their U.S. counterparts. They collected 60 hours of observations in primary schools, conducted formal and informal interviews with educators, and examined school documents. The researchers noted that a strong mathematical content knowledge is required of teachers in France before they are allowed to enter teacher preparation programs. The authors conclude that this is the primary contributor to the high achievement of French mathematics students.

While few researchers doubt the effect of teachers’ mathematics content knowledge on effective classroom instruction, a survey sent to New York principals did not reveal content knowledge in general to be one of their major concerns. Torff and Sessions (2005) mailed 300 surveys to secondary school principals in New York State, of which 242 were completed and returned. Of the principals who responded, the least likely cause of teacher ineffectiveness to be indicated was the content-knowledge deficiency of teachers. Instead, these principals listed classroom-management skills, lesson implementation skills, rapport with students, and lesson planning skills as more important for teacher effectiveness than content knowledge. However, had this study been focused on mathematics teachers and sent to department heads rather than principals, the results might have been quite different. Only a small portion of school principals start their careers as mathematics teachers, and as such they may not be competent.
enough in the secondary mathematics curriculum to evaluate whether or not a teacher has adequate mathematical content knowledge.

Similar to the principals in the Torff and Sessions (2005) study, preservice teachers have been shown to believe that content knowledge is less important than teacher enthusiasm and teaching skills when it comes to instructional effectiveness. Minor et al. (2002) surveyed 134 preservice teachers during their first week of an introductory-level education class, using the Preservice Teachers’ Perceptions of Characteristics of Effective Teachers Survey. The participants were asked to identify, rank order, and define what they perceived as characteristics of effective teachers. Eighty percent of the participants did not indicate that being knowledgeable about subject-matter was a trait of effective teachers, while more than one-half of the participants identified student-centered themes relating to teacher effectiveness. However, since these preservice teachers were enrolled in an introductory-level education class, it is probable that they had not yet had any significant experience as teachers in the classroom. It would be interesting to follow-up with these same teachers once they had experienced student teaching to see if their perceptions of effective teaching had changed. It seems reasonable to hypothesize that, once exposed to real-life teaching experiences, the importance of content-knowledge, particularly in mathematics, may become a larger component of what teachers perceive as a characteristic of effective teaching.

For the purposes of this review, one major limitation of the work done by both Minor et al (2002) and Torff and Sessions (2005) is that it did not focus explicitly on effective mathematics instruction. More conclusive evidence is needed on the link between teachers’ mathematical content knowledge and effective mathematics teaching.

As Laczko-Kerr and Berlinger (2003) note, “research provides convincing evidence that subject-matter knowledge is necessary but not sufficient for teaching well” (p. 35). Teachers must also possess instructional strategies that have been shown to positively influence student learning in the classroom. Literature that identifies two of these possible strategies will be reviewed next.

Objectives-Based Instruction

Over six decades ago, Ralph Tyler (1949) stated the following:

If an educational program is to be planned and if efforts for continued improvement are to be made, it is very necessary to have some conception of the goals that are being aimed at. These educational objectives become the criteria by which materials are selected, content is outlined, instructional procedures are developed and tests and examinations are prepared. (p. 3)

This demand for objectives-based instruction has been repeated by many other authors and researchers throughout the years (Mager, 1975; Gagne & Briggs, 1979; Airasian, 2001; Panasuk & Todd, 2005). Prior to writing lesson plans for instruction, it is “important to be able to state clearly just what you intend the results of that instruction to be” (Mager, 1975, preface). Panasuk and Todd (2005) note that the purpose of writing student objectives is “to guide the lesson planning process” (p. 219).

Student learning objectives should be the starting point for planning the instruction and assessment of the learner. This idea is concurrent to the statement made by Gagne and Briggs (1979) that “objectives should guide the instruction and the evaluation, not the other way around……the objectives should be determined before the lesson plans or evaluation instruments” (p. 32). According to Mager (1975), there are a number of reasons that objectives are important. If there is no clear objective, there is no basis for selecting materials or methods for instruction. If there is no clear objective, it is not possible to identify when the end of instruction has been reached. Finally, if there is no clear objective, students are less likely to be able to “organize their own efforts toward accomplishment of those objectives” (p. 6).

Mager posits that “an objective is a description of a performance you want learners to be able to exhibit before you consider them competent. An objective describes an intended result of instruction, rather than the process of instruction itself” (p. 5). According to Tyler (1949), there are three basic rules for writing objectives. The first is that objectives should not be stated as actions that the teachers will complete; these are not educational objectives. The purpose of an objective does not lie in what the teacher does, but in what the students learn. “It becomes important to recognize that any statement of the objectives …… should be a statement of changes to take place in students” (p. 44). With these statements, it is possible to determine which activities the teacher should perform so as to facilitate the students’ attainment of the stated objectives. The second rule is that objectives should not be stated in the form of content topics or generalizations that students are to encounter. “These are not satisfactory objectives since they do not specify what the students are expect-
ed to do” (p. 45). Merely stating what topics are to be covered in a lesson does not indicate the change in students learning that is expected to have occurred at the conclusion of the lesson. The final rule is that objectives should not be stated in “the form of generalized patterns of behavior” (p. 46) such as ‘to learn to be a better citizen’ or ‘to become a critical thinker’, since they do not indicate the area of “content to which this behavior applies, or the area in which such behavior is to be used” (p. 46).

“It is best that objectives be explicit, clear, and measurable” (Airasian, 2001, p. 74). Tyler notes that “the most useful form for stating objectives is to express them in terms which identify both the kind of behavior to be developed in the students and the content or area of life which this behavior is to operate” (p. 46–7). This coincides with Gagne and Briggs (1979) definitions of a well-written student objective. “An objective is precisely described when it communicates to another person what he would have to do to observe that a stated lesson purpose has in fact been accomplished” (p. 118).

Most teacher preparatory programs contain direct instruction for preservice teachers in how to construct student learning objectives (McCutchon, 1980; Kennedy, 2004; Stangis, Pringle, & Knopf, 2006). However, researchers have reported that, although preservice teachers acknowledged the importance of standards and lesson objectives when planning lessons, the preservice teachers were more concerned with first finding the instructional activities to take place during the lesson (Stangis, Pringle & Knopf, 2006). In this study, a majority of the preservice teachers reported having difficulty writing student learning objectives. Student teachers reported that writing objectives “would become easier over time” (p. 78) and noted that deficits in their own content knowledge made it difficult to begin planning lessons by writing objectives.

In research performed in Newfoundland, many in-service teachers rated themselves as adequate at instructional design strategies, but most could not “develop a specific objective on request” (Kennedy, 1994, p. 20). Primary and elementary teachers studied showed a familiarity with student learning objectives (Tobin, 1990, as cited in Kennedy, 1994), and secondary teachers reported “extensive use” of student learning objectives (Graham, 1992 and Thomey, 1991, as cited in Kennedy, 1994, pp. 20–1), yet most used those objectives that were provided in their curriculum materials. Two-thirds of the secondary teachers studied did not use objectives as the basis for assessment of student learning. In a study by Bol & Strage (as cited in Delong, Winter, & Yackel 2005), identified a distinct disconnect between high school biology teachers’ stated lesson goals and the goals that were suggested by their lesson assessments. This study demonstrates that even when teachers begin their lesson planning with student-learning objectives, they do not necessarily base their subsequent classroom activities and assessments on those objectives.

Strangis, Pringle, and Knopf (2006) concluded that, although preservice teachers do not always start planning with student objectives, identifying student learning objectives at some point in the planning process is key to good teaching. This is in direct opposition to research conducted by Panasuk and Todd (2005), who found that when teachers planned using a specific objectives-first method, writing and explicitly stating student learning objectives was closely linked to a logical flow of the lesson through multiple phases, as well as the homework correctly assessing the lesson’s objectives. “The core proposition [of objectives-based instruction] is that well-structured lessons flow from well-specified objectives” (p. 228). In a study performed by Morine-Dershimer (1977), it was noted that teachers who stated behavioral objectives in their written plans for a reading lesson were more frequently those teachers whose students had high gain scores than those with low gain scores. In this study, teachers were allowed to choose prewritten student objectives or to construct their own. She noted that “when original [objectives] were developed, the only teachers to state [these] were teachers with high pupil gain scores” (p. 19). In light of the conclusions drawn by Panasuk and Todd (2005) and Morine-Dershimer (1977), it is reasonable to conclude that objectives-based lessons can lead to more effective classroom instruction than lessons that are not explicitly and directly based on student learning objectives.

**Students’ Prior Knowledge**

In order for classroom instruction to be effective, teachers must first know what tasks and skills students have already mastered; they must take into consideration the students’ prior knowledge. Robert Gagne (1970/1973) stated, “the most dependable condition for the insurance of learning is the prior learning of prerequisite capabilities” (p. 111). During his research, results suggested that the performance of students in a self-paced mathematics program that was “otherwise effective” (p. 163) was highly dependent on whether or not they had previously learned the necessary subordinate knowledge to complete the program and was independent of other factors (Gagne, 1963/1973). Based on Gagne’s ideas of prior knowledge, Panasuk and Todd
(2005) advocate for teachers to determine the necessary prior knowledge, or “constituent parts” (p. 222) of the mathematics concept to be learned by the students. This allows teachers to present a gradual development of the concept from what the students are already able to do on their own to what they are desired to learn. It also provides teachers a lens through which to preview the types of student errors or misconceptions that they may develop as the lesson progresses. The teacher can then plan to introduce examples in the lesson that will challenge students’ misconceptions and increase the likelihood of learning for understanding.

The method of identifying the subordinate skills that are required by students to learn a new concept is called task analysis. “Task analysis refers to several different, though interrelated, procedures which are carried out to yield the systematic information needed to plan and specify the conditions for instruction.” (Gagne & Briggs, 1979, p. 114). Two types of task analysis that Gagne details are the learning task analysis and the information processing task analysis. A learning task analysis is the act of identifying the prerequisite knowledge and skills necessary for the objective or task that is to be learned. An information processing task analysis identifies each individual step which, when taken as a whole, makes up the totality of the task of the primary objective to be completed.

One means of performing a task analysis is for teachers to work out all of the homework and classwork problems that students will be required to complete during a class lesson. The purpose of this is for teachers to analyze the underlying concepts and skills to define the prerequisite skills and knowledge needed by the students to learn the new material (Panasuk & Todd, 2005). The act of working out homework before developing the class activities allows teachers to determine which sub-skills must be reinforced, as well as to predict where students may have misconceptions during the instruction. This allows the teacher to tailor the lesson accordingly to the maximum benefit of the students.

During the Missouri Mathematics Effectiveness Project, Good and Grouws (1979) instructed teachers in a treatment group to spend some time at the beginning of each class focusing on the recall of prior knowledge that students would need to use during the new lesson. While the posttest means of both the treatment and control groups were higher than the pretest means, students of teachers in the treatment group had larger gains on standardized achievement tests than students of the teachers in the control group. For the purpose of this paper, a major limitation of this study is that the focus on prior knowledge was only one of the teaching behaviors explicitly requested of the treatment teachers. As such, it is not possible to isolate the effects of this one specific treatment behavior from the entire treatment. However, the results are supported by other research. Wiegand (as cited in Gagne, 1970/1973) asked students to perform a task with which they were unfamiliar. At first, nearly all of the students could not complete the task. Wiegand then tested the students at successively simpler levels of the prerequisite skills necessary to complete the task until a base-level was found where all students possessed the same basic knowledge. The students were then taught, beginning at that level of commonality, all of the necessary sub-skills required to complete the task. However, they were not instructed in how to complete the task itself. Nine out of ten students were able to complete the final task without direct instruction once they possessed all of the necessary prior knowledge.

These findings are substantiated by those of Carnine (1980), who studied the effects of preteaching versus concurrent teaching of component skills of a new concept. Fifteen lower-achieving first-graders, identified by teachers, were assigned to either a treatment or control group. The two groups were both taught a multiplication algorithm. However, the preteaching group first taught the component skills of the algorithm and the concurrent group was taught the component skills along with the algorithm at the same time. Time measures, as well as correct responses to the final task, were recorded. “Children in the pre-teaching group reached criterion in significantly less time than children in the concurrent group,” (p. 378) and also scored significantly better on the final task. The results of this study seem to indicate that the order of the teaching of components skills of a new concept is a relevant factor to be considered by the teacher. Students in this study learned in less time and provided more correct responses to transfer facts when pre-taught the component skills of multiplication before being taught the multiplication algorithm itself. While these students were in first grade, it would seem likely that the more complicated the mathematical task, the more likely this method of insuring that students have the necessary prerequisite knowledge and skills needed to master a new concept prior to instruction would be beneficial. More research into whether or not these results would be transferrable to other levels of mathematics instruction, such as middle and high school, is necessary.

Summary
There is no one single method to produce effective mathematics teaching; instructional strategies that work
for one class of students may not work for another. There are, however, some elements which can be shown to be common to effective teaching. These are teachers’ level content knowledge, teachers’ use of objectives-based instruction, and teachers’ recognition of the prior knowledge and skills required of the students. Classroom instruction that includes these characteristics is more likely to be effective than instruction that is lacking in one or more of these traits. However, there must be some method of structuring and implementing these characteristics so that they may be more useful to teachers in the classroom. In order to facilitate effective classroom instruction, teachers should incorporate these three elements into their lessons during the planning process. Therefore, lesson planning will be the focus of the second part of this review.

**LESSON PLANNING**

From a search of online databases, it appears that research on lesson planning was at its peak during the late 1970s to the early 1980s, which coincided with the spread of instructional design strategies for education being taught in teacher training programs. This research identifies many different definitions of lesson planning. Peterson, Marx, and Clark (1978) state that planning is the “preactive decisions the teacher makes prior to the act of teaching” (p. 418). Clark and Yinger (1979) “define planning as a process of preparing a framework for guiding teacher action……[which] involves teacher thinking, decision making, and judgment” (pp. 8-9). More general definitions involve any and all thoughts or decisions that teachers make before they begin to teach (Fernandez & Cannon, 2005; Panusuk, Stone, & Todd, 2002). Panusuk and Todd (2005) state that “planning a lesson involves teachers’ purposeful efforts in developing a coherent system of activities that facilitates the evolution of students’ cognitive structures” (p. 215). For the purposes of this paper, comprehensive lesson planning will be defined as the act of planning for classroom instruction that:

- is written explicitly based on student learning objectives,
- includes assessments based on the stated objectives,
- requires teachers to take into consideration the prior knowledge of students as well as the prerequisite skills and knowledge necessary for completion of the stated objectives, and
- has consistency throughout all phases of the lesson.

The questions remain, however, to what extent do teachers plan comprehensively, and does comprehensive planning improve the effectiveness of their classroom instruction? It is these topics that will make up the second part of this paper.

**Teachers’ Planning Activities**

Research suggests that the majority of teachers do not consistently and thoroughly plan for instruction (Clark & Yinger, 1979; Morine-Dershimer, 1977; McCutcheon, 1980; Stigler & Hiebert, 1989; Kagan & Tippins, 1992; Searcy & Maroney, 1996; Reiser & Mory, 1991). Much research on teacher lesson planning seems to have been done ex post facto, with in-service teachers interviewed or surveyed about previous planning activities (Clark & Yinger, 1979; Sardo-Brown, 1990; Reiser, 1994; Searcy & Maroney, 1996; Sanchez & Valcarcel, 1999; Yildirim, 2001). Other studies have asked teachers to participate in a staged lesson-planning activity in which they plan, and sometimes subsequently teach, a concept outside the scope of their usual curriculum (Morine-Dershimer, 1977; Peterson, Marx, & Clark, 1978). Still other studies are built around the planning that occurs throughout the school year (Sardo-Brown, 1988, 1993; Reiser & Mory, 1991; Reiser, 1994; Panusuk, Stone, & Todd, 2002; Kagan & Tippins, 1992), observing and collecting artifacts from in-service teachers as they go about their regular teaching duties.

Several studies have relied on teachers’ self-reporting of their own planning habits and decisions through the use of survey instruments (Clark & Yinger, 1979; Searcy & Maroney, 1996). In a survey of special education teachers in Iowa, Searcy and Maroney attempted to collect information about how teachers plan lessons. The majority of the teachers who returned the survey stated that they did not write lesson plans for every lesson. However, this may be a function of the special education teacher not being the primary instructor of the class, deferring planning responsibilities to the regular education teacher in the room. In an attempt to answer the question “how do teachers plan” (p. 11), Clark and Yinger surveyed 300 elementary school teachers in Michigan. They found that many teachers self-reported that they did not write lesson plans, but rather relied on mental plans. This work is supported by previous research. The process of staged lesson planning was used by Morine-Dershimer (1977) whose goal was to document “the kinds of planning that teachers engage in before a lesson begins” (p. 1). In this study, 40 second- and fifth-grade teachers were given a topic and materials to create one lesson each in mathematics and reading. In the follow-up interviews, two-thirds of the
teachers acknowledged that the lesson plans that they had written for this study were done in much greater detail than their usual plans. These teachers also noted that, for the most part, the regular planning that they did was not of a written form; rather, it was ‘in their heads’ (p. 5). Taken together, these studies indicate a widespread lack of written planning being performed by classroom teachers. However, one limitation of the literature reviewed here is that none of these studies focused explicitly on the planning activities of mathematics teachers.

Research has shown that, when teachers do create written lesson plans, the most common form of these plans are outlines or lists, followed by narratives. This was noted in the work of Clark and Yinger (1979), who found that “the most common form of written plan was an outline or list of topics to be covered” (p. 15). Additionally, Sardo-Brown (1988) noted that at the yearly, unit, and weekly planning stage, teachers’ plans were ‘sketchy outlines’ or “lists and notes” (p. 77). At the daily planning level, teachers’ plans were “detailed notes” (p. 77) about what was to happen in the course of the class. Furthermore, McCutcheon (1980) noted that teachers planning resembled a grocery list, just a jotted list of reminders, as opposed to a guide for instruction. According to research, teachers often indicate that they do not use the lesson plan format that they were taught in their teacher preparation programs or those that were created and marketed by experts (Searcy & Moroney, 1996). Rather, they develop their own method and system for lesson planning that is highly individualized and not based on research.

Taking into consideration the focus most teacher-preparation programs put on the development of lesson plans, it seems that once teachers strike out on their own and are no longer required by instructors to write plans, many alter the form of their planning or stop writing plans altogether. However, this doesn’t mean that teachers aren’t planning for instruction before entering the classroom. Nevertheless, it does make it extremely difficult to study the planning habits of teachers if they are not in written form. In order to study the effectiveness of planning, the researcher needs first to know what decisions teachers entertain, dismiss, and accept as they think about each lesson.

Planning for Student Learning

Over the course of a twenty-year period, multiple studies have been conducted that indicate a pervasive trend in the lesson planning habits of teachers. These studies show that a majority of teachers focus their lesson planning on classroom activities rather than on student learning objectives (Peterson, Marx & Clark, 1978; Clark & Yinger, 1979; Mintz & Yarger, 1980; McCutcheon, 1980; Sardo-Brown, 1988). While each study has its own set of limitations, taken as a whole they provide evidence that teachers may unknowingly undermining their own best efforts at effective mathematics teaching by placing the focus of instruction on what is to be done by the students and teacher during the lesson rather than what must be learned by the student as a result of the lesson.

Peterson, Marx, and Clark (1978) recruited twelve experienced in-service teachers to plan and teach a social studies lesson to three staged classes, each day for three days, outside of the regular school schedule. “Each teacher was given a list of objectives and social studies text material, and was told to teach the material in any way he or she wished” (p. 419). Each teacher had exactly 90 minutes to plan at the beginning of each day and was asked to use the ‘think aloud’ method so that investigators could analyze their thinking and decision making processes while planning for instruction. The results were stunning. “Even though the teachers were provided with a list of desired cognitive and affective student objectives, they did not refer to them in their planning, nor did they relate their choice of instructional processes to learning objectives’ (p. 424). It was found that teachers focused most of their planning on the content and activities, and the least amount of thought was given to student learning objectives.

Clark and Yinger (1979) conducted both a survey of elementary school teachers as well as a field study of teacher planning activities. In the teacher survey, 78 of the 300 teachers who were mailed a survey completed and returned it, for a response rate of only 26%. Those teachers returning the survey self-reported that student learning objectives are not the basis of developing lesson plans; activities lead what happens in the classroom during any given lesson. In the field-study, six volunteer elementary school teachers (four individual teachers and one teacher-pair) were followed over the course of five weeks while they planned and implemented a two-week reading lesson that they had never taught before. Teachers were asked to keep a journal of their planning and were also interviewed on a bi-weekly basis. Clark and Yinger found that “rather than moving from well specified and carefully stated objectives and proceeding to designing activities to meet these objectives, our teachers more commonly began with a general idea” (p. 18).

Supporting this work by Clark and Yinger is a study of 68 elementary school teachers conducted by Mintz and Yarger (1980). They attempted to analyze written
lesson plans by asking in-service teachers to plan reading instruction for a simulated class of students. However, the researchers did not expect the teachers to implement the plans as part of this study and, therefore, were able to have the participants plan for one of three time periods: an entire school year, a span of several weeks, or a single class period. The researchers concluded that teachers focus their efforts during lesson planning on what content is to be taught by the teacher, rather than on what content is to be learned by the student. “There was little mention of objectives” (p. 8). In addition, there is a study completed by McCutcheon (1980), focusing on the nature of the planning done by elementary school teachers. During this research, it was reported that “teachers usually listed objectives for a lesson in their planbooks only if required to do so by the principal.” (p. 6). The teachers commented that objectives were provided in the teachers’ manuals, thus they felt that they didn’t need to indicate in their plans which objectives they were going to attend to in each lesson. Also, teachers in the study noted that their planning was impeded by the fact that they lacked sufficient coursework on planning in both preservice education and in-service professional development.

In Sardo-Brown’s (1988) research study on the planning habits and decisions of 12 in-service middle school teachers, participants were required to document their planning throughout the course of a school year in an effort to describe the real decisions about actual course curricula teachers implement. This research involved multiple collections of different types of data over a period of several months. Each of the 12 middle-school teachers who participated were interviewed a total of four times, once each at the yearly, unit, weekly, and daily levels of planning. The teachers also completed a written questionnaire, were asked to ‘think-aloud’ while planning their daily lesson and have the planning audio-taped, and submitted copies of their yearly, unit, weekly, and daily lesson plans to the researcher. Over the course of the four-month study, the researcher gleaned from the data that teachers did not write student learning objectives while planning daily lessons. Instead, the teachers’ “first considerations in planning were what instructional strategies to use” (p. 78).

In contrast, Searcy and Maroney (1996) conducted a survey of special education teachers in Iowa, asking 556 teachers to self-report on their own lesson planning methods. Of the 207 teachers who completed and returned a survey, approximately one-quarter noted that they did not write down student objectives when creating lesson plans. On the other hand, over half of the teachers claimed that they consciously plan student learning objectives; however, they just don’t write them out. As this study encompassed teachers from preschool through twelfth grade, Searcy and Maroney were able to note that, for this study, planning practices were not significantly different at the different grade levels. Once again, this result may be because of the special education teacher not always being the primary instructor in the classroom.

In addition, during a survey of 33 K–12 teachers, many of the teachers reported making decisions based on student learning objectives (Sardo-Brown, 1990). A multiple choice survey, a questionnaire, and an interview were all completed by the participants. However, several factors may have influenced these findings. These teachers were concurrently enrolled in a graduate degree program, with half of the participants of the study enrolled in a class taught by the researcher. These teachers were not volunteers; rather, they were required to participate as part of their degree program. These teachers also were all employed by a single school district which required teachers to use a specific lesson-plan template. It was noted that this template included the setting of student objectives. It is, therefore, plausible that the data collected from these teachers does not represent the teachers’ preferred approach to lesson planning, but instead is a response either to the school district’s lesson planning requirements or the desire to make themselves look better in the eyes of the researcher.

Limitations aside, some research does support Sardo-Brown’s assertion that many teachers do refer to student learning objectives when planning for instruction. During a case-study of two experienced elementary school teachers, Reiser and Mory (1991) noted that, of the two teachers, the one who had completed a graduate-level course in instructional design tended to use the systematic lesson planning model that she was taught as she planned her own lessons. The teacher who had not received formal training in instructional design did not use any systematic method for lesson planning. However, for the purpose of this study, the researchers only observed the planning activities of these two teachers as they related to the teaching of science lessons. As elementary school teachers, these teachers were responsible for teaching most of the subjects that their students studied. While it is plausible that the teachers used a consistent method of planning for all subjects that they taught, it is possible that the teacher who did not use a systematic method for planning her science lessons may have been influenced by factors other than a lack of formal training, such as a deficiency in her
content knowledge or a general dislike of the subject of science. Finally, if the focus of this study had been on planning for mathematics instruction rather than science, the results may have been significantly different.

It should be noted that, although there have been many studies conducted in an effort to describe the decisions that teachers make during the lesson planning process, none of the literature reviewed here has focused on mathematics teaching. While it can be implied that research into the planning practices of elementary school (Clark & Yinger, 1979; Mintz & Yarger, 1980; Reiser & Mory, 1991), middle school (Sardo-Brown, 1988), and special education (Searcy & Maroney, 1996) teachers has implicitly included planning for mathematics classes, any effort to improve mathematics instruction should include research conducted explicitly on the lesson planning practices and decisions of mathematics teachers.

Given the overwhelming disregard for writing student objectives, it is plausible that there is little congruency between what students learn and what is intended to be implemented in the classroom. If teachers do not consider student learning outcomes and subsequent assessments as they plan, they will not be able to recognize whether or not they are teaching effectively. Few studies have examined the link between comprehensive lesson planning and instructional effectiveness but, from those which have, there are indications of a close association. This connection will be discussed later in the paper.

**Planning in Countries Other Than the United States**

Research from other countries shows a similarly uneven approach to lesson planning. The international literature does not reveal any one universally accepted system of lesson planning. In Spain, Sanchez and Valcarcel (1999) performed research that involved interviewing 27 science teachers. These teachers taught either 11–13 year-olds or 13–17 year-olds and held either a diploma in education or a science degree. According to the researchers, there was no reported difference between the groups as pertains to planning. Analysis of the interviews showed that 92% of teachers wrote some type of document to go along with their lessons, even if the entire lesson plan was not written out. Of these teachers, 76% stated that what they wrote “consisted of notes and guidelines for the teacher’s own use on the content to be taught and how the lessons were to run” (p. 500). However, this number may be artificially high based on the context in which the interviews were conducted. The interviewed teachers were attending “an in-service training course on teaching unit design……following the latest reforms in……education” (p. 496). Asking teachers to describe their own planning decisions in the midst of teaching them a reform planning method may have influenced the teachers to over-report their own use of the desired attributes in an effort to please the researchers.

Research has shown that, even when teachers are required to plan around student learning objectives, they may not hold much faith in the usefulness of this practice. Yildirim (2001) studied the responses of almost 1,200 Turkish school teachers to a survey “designed to explore teachers’ perceptions of their planning at the primary school level” (p. 7). He noted that the Turkish education system had strict rules for the writing of lesson plans at the yearly, unit, and daily levels, and that all of these planning levels included student learning objectives. In his analysis, it was found that when teachers were asked to rate what components of daily lesson plans they perceived as most important, writing student learning objectives rated lowest. “Many teachers did not find writing behavioral objectives useful in planning” (p. 18) even though they were required to do it at all levels of instructional planning. Teachers were more concerned with how and what to teach, the content of the lessons, than with writing student learning objectives for the lesson or the evaluation of students based on those objectives. While Yildirim notes that the planning of teachers in Turkey is strictly monitored, he does not mention whether or not the classroom teachers are required to have had formal training in the components of lesson planning, such as development of student objectives for instruction and assessment.

Preservice teachers at Oxford University in England were studied by John (1991) in an effort to identify both how they plan and what they take into consideration while planning. Five preservice teachers were chosen at random; they participated in open-ended interviews and think-aloud activities, as well as provided all lesson-planning materials to the researcher. John noted that the writing of objectives was not seen as having great importance to the student teachers. Similar to research previously identified, these teachers tended to focus on the topics, resources, and activities of a lesson when planning. John notes that the Tyler (1949) method of lesson planning does not resemble the thinking and actions of teachers in practice. He notes that “teacher educators should begin to question the Tylerian or overtly theoretical approach to planning in the light of the evidence cited in this paper and elsewhere” (For Teacher Education section, T11). What it seems he is suggesting is that teacher preparation programs and
schools of education stop teaching lesson planning as
an objectives-first process and begin teaching it so as to
more closely resemble what teachers have been
observed to already do while planning. However, John
does not take into account in this research two very
important issues, specifically, whether or not these five
preservice teachers were properly taught how to plan
following an objectives-first method, and whether or
not they had successfully mastered this skill. Also, there
is no evidence that these teachers, who did not correct-
ly use the Tylerian method of lesson planning, would
find more success with other methods.

In summary, the Turkish teachers were required to
follow a strict formal structure for their lesson plans
while the Spanish teachers had a much more flexible
system. The English teachers were found to engage in
lesson planning activities that were quite different from
those prescribed by educational theorists. The diversity
of lesson planning formats and the complexity of indi-
vidual teachers’ planning activities point to some of
the difficulties in studying lesson planning. While obvious-
ly not an exhaustive review of the literature of interna-
tional lesson planning, this small sample does identify
some themes that are common to teaching in class-
rooms across the globe. Teachers in other countries
seem to be comparable to those in the United States in
terms the structure of their written plans, their use of
student learning objectives to guide instruction and
assessment, and their perception of the usefulness of
those objectives.

Common Lesson Planning Models

As stated previously, any effort to improve mathe-
matics instruction should include research conducted
explicitly on the lesson planning practices and decisions
of mathematics teachers. The term lesson planning, how-
ever, is very broad and can include any number of dif-
ferent activities that teachers carry out before instruction
occurs. Therefore, it is necessary to briefly summarize a
number of different methods for lesson planning provid-
ed in the literature as well as a range of formats commer-
cially available to teachers. Cruickshank, Bainer, and
Metcalf (1999) identified a sample of six different types
of lesson planning formats including those by Callahan,
the Center for Vocational Education, the Air Force, Eby,
Gange and Briggs, and Hunter, as well as their own “re-
commended lesson plan format” (p. 147). While there are
several components that many of these models have in
common, such as resources and assessments, there is
only one that they all have in common. Each of these
lesson planning formats includes an objectives compo-
nent. It is presumed that each of the methods referenced
by Cruickshank, Bainer, and Metcalf have tailored their
planning method to a specific audience, for example,
vocational school instructors, Air Force instructors, or
academic teachers and, therefore, each method will
include different lesson planning components specific to
that audience’s content. For the purpose of this paper,
planning models not specifically designed for the aca-
demic classroom will not be considered.

Much has been written about daily lesson planning
for academic school teachers. An early book published
on the topic is Tyler’s Basic Principles of Curriculum and
Instruction (1949). In it, he states that there is a four-
step process to planning before instruction. First, state
the objectives. Second, select learning experiences for
the students that are likely to promote the attainment of
these objectives. Third, organize the learning experi-
ences so as to best facilitate the attainment of the stated
objectives. Fourth, evaluate the learning experiences to
determine whether or not the stated objectives were
met. Decades later, this process for lesson planning was
renewed and reinforced by the writings of Mager
(1975), Gagne and Briggs (1979), and Airasian (2001).
These are all authors that have included approximately
the same steps in their respective recommendations for
teachers to attend to while planning daily lessons.

Many of the various formats include a section on
materials or resources, and/or include a section for eval-
uation or assessment. The act of stating lesson objec-
tives seems to be the only common element in many
different methods for lesson planning that have been
described here. However, each planning procedure
places the writing of objectives at a different ‘step’ in
their plan. Many of the lesson plan formats do not begin
with stating students learning objectives. Interestingly,
each lesson plan format mentioned is presented in the
manner that the plan will be enacted in the classroom.
Since no notation is made to the contrary, it can be
assumed that the steps presented in each form are in the
order that the authors also recommend the lesson be
planned.

In an effort to focus explicitly on improving the
instruction provided to students in mathematics classes,
Panasuk, Stone, and Todd (2002) present a lesson plan-
ing format that is tailored specifically for the daily
planning of mathematics lessons. The first, and most
important, difference to note between this method and
the other previously mentioned is that the order for
planning instruction is different than the order for
delivery. The components for developing a lesson plan
as stated in the Four Stages of Lesson Planning (FSLP)
model are (in order) objectives, homework, develop-
mental activities, and mental mathematics. This order
for development proceeds directly from the literature on objectives-based instruction and students’ prior knowledge. The literature reviewed previously shows possible connections between teachers’ use of objectives-based instruction and task analysis, along with teachers’ awareness of students’ prior knowledge, and more effective classroom instruction.

During a statewide Middle School Mathematics Initiative, middle-school teachers were trained in the FSLP model. Panasuk, Stone, and Todd (2002) chose a purposive sample of three teachers from the larger population to complete a questionnaire and be interviewed about the implementation of the FSLP strategy in their own classrooms. The researchers found that “thoughtfully constructed…lesson plans…helped the teachers to become effective classroom leaders and facilitators of efficient use of classroom time” (p. 825). They also noted that “the teachers accomplished instruction based upon established [student learning] objectives” (p. 825). One of the teachers noted that, as she planned, she began to recognize the gaps in her own content knowledge. Also, using the FSLP caused her to “think more often about connections among mathematics concepts” (p. 819).

In further research conducted into the FSLP strategy by Panasuk and Todd (2005), a link was shown between the explicit writing of student learning objectives and effective teaching as determined by direct observation of researchers. Thirty-nine teachers from the original Middle School Mathematics Initiative participated in a smaller study in which they engaged in a pre-observation conference, an observation by a trained specialist teacher, and a post-observation conference. In addition, a total of 261 written lesson plans were developed and collected for the study; each lesson plan and observation was then individually scored using a rubric. Panasuk and Todd found that when written lesson plans included explicitly stated learning objectives for students, lessons also included multiple representations of concepts, teaching was in line with state curriculum frameworks, and possible or common student misconceptions had been considered by the teacher before beginning instruction. Thorough lesson planning in this study led to lessons that were considered by the researchers to be highly effective in achieving the intended learning outcomes.

Summary

Each of the lesson planning formats described here recognizes the necessity of teachers identifying student learning objectives prior to instruction. These formats also rely on teachers having a strong content knowledge background, a large supply of possible instructional activities, and the creativity to design appropriate lessons for their students. However, only the FSLP method put forth by Panasuk, Stone, and Todd (2002) meets the previously stated definition of comprehensive lesson planning. Lessons developed using the FSLP method:

1. are written explicitly based on student learning objectives;
2. include assessments of the students based on the stated objectives;
3. require teachers to perform task analyses to identify the prior knowledge of students, as well as the prerequisite skills and knowledge necessary to complete the stated objectives; and
4. have consistency throughout all phases of the lesson.

The case is therefore made that, all other things being equal, lessons planned using the FSLP method will lead to more effective mathematics instruction than those that are not.

CONCLUSIONS AND POTENTIAL RESEARCH QUESTIONS

In order for teachers to provide the best possible education to their students, classroom instruction needs to be effective. The research has provided several characteristics that are common to effective mathematics instruction. The first is the need for teachers to possess a high level of content knowledge. The second is that instruction in mathematics must be based explicitly on student learning objectives. Lastly, the prior knowledge of the students and the necessary subskills of the content must be taken into account. Instruction that contains these three characteristics is likely to be more effective than instruction that is lacking in one or more of these.

Mathematics lessons planned in a comprehensive manner may lead to more effective classroom instruction. This is because this method of planning forces teachers to recognize and attend to their own content-knowledge deficiencies. It requires that teachers base all instructional activities and assessments in the lesson on explicitly stated student learning objectives. Also, this method of planning leads teachers to analyze the stated objectives in order to determine the best methods to bridge the gap between student prior knowledge and the new learning. Because of the focus on these three characteristics identified by researchers as being effective at improving student achievement, it is possible to hypothesize that teachers may improve their classroom
instruction by implementing comprehensive planning methods in their classrooms.

Colleges of teacher education expect pre-service teachers to plan lessons and assume that these novice teachers will continue this practice as they begin their teaching careers. Yet, the literature indicates that they do not. In addition, teachers who have entered the classroom through alternative methods of licensure may have little to no training or knowledge of lesson planning methods. Unfortunately, once in the classroom, these teachers are often left to their own devices when it comes to planning for instruction. Research into professional development programs for in-service teachers aimed at improving lesson planning in mathematics is warranted due to the potential for positive effects on both classroom instruction and student learning. The following research questions are suggested:

1. Does the particular method that teachers use to plan lessons impact the effectiveness of their classroom instruction?
2. What attitudes or habits held by teachers may aid or impede their adoption of comprehensive lesson planning methods?
3. Is there a significant difference in achievement on a specific unit of content between the students of teachers who have been taught to plan using the Four Stages of Lesson Planning method and those teachers that have not?

References


Interactive-Engagement and Reflective Writing

Michael Doherty
Andover High School

ABSTRACT

This paper discusses the use of pedagogical approaches (called interactive-engagement (IE) strategies) that have been found to be more effective at facilitating conceptual understanding than traditional instructional techniques. Commonalities between IE approaches are identified. The claims regarding differences between IE and traditional instructional strategies have been justified by the analysis of Force Concept Inventory (FCI) pretest and posttest data through the computation of average normalized gain \( \langle g \rangle \). The average normalized gain for a course is the ratio of the average actual gain \( \langle G \rangle \) to the maximum possible average gain. The FCI is a formative assessment that has been used as an instrument to assess the effectiveness of instructional methodologies at facilitating conceptual understanding among high school students taking introductory physics courses that cover Newtonian Mechanics. It is acknowledged that interactive-engagement techniques require more time than traditional teaching techniques and therefore result in less content coverage for equal student-teacher contact time.

Reflective journal writing, an activity shown to facilitate conceptual understanding in other learning environments, is also reviewed. The evidence reported regarding graded reflective writing by introductory level physics students suggests that reflective journal writing can facilitate understanding, promote metacognitive engagement leading to the learner making connections between learning events and prior knowledge, and create a space for dialogue between the teacher and student. These are also attributes of IE techniques. It is suggested that student engagement in daily reflective journal writing, coupled with a weekly reflective summary, each of which is a graded component of the course, may facilitate a more time-efficient use of IE techniques. This could effect a reduction in the content coverage deficit resulting from implementing IE strategies, while facilitating the increased conceptual understanding associated with IE methods.

INTRODUCTION TO THE PROBLEM

According to the American Institute of Physics (AIP) the trend in physics education is toward increased enrollment. There has been an increase in the percentage of high school graduates that have completed a physics course, from less than 18% in 1986 to approximately 33% in 2001 (AIP Statistical Research Center, as cited in Hehn & Neuschatz, 2006). The development of the conceptual approach to teaching physics, which allows for new instructional techniques involving more active student engagement in the learning process, is a factor contributing to the increased high school physics enrollment. Historically, high school physics classes were likely to be populated by students that had exhibited the highest academic performance (Hehn & Neuschatz, 2006). The implication here is that the increased enrollment comes from the segment of the student population that is more dependent upon the effectiveness of instructional technique to develop conceptual understanding than has historically been the case. Other trends identified were an increase in emphasis on what is learned, rather than what is taught, and increased use of formative assessments to guide instruction (Hehn & Neuschatz, 2006).

Newton’s laws of motion, a central topic of introductory physics courses, are mediated through the concept of force. Many students enter into a study of Newton’s laws with an existing schema that interacts with the instruction, and thereby may modify the intended learning outcome (Watts & Zylberztajn, 1981; Halloun & Hestenes, 1985a). These commonly held and frequently scientifically inaccurate beliefs are the result of personal observations and intuitive thought processes that have not been exercised in a rigorous and thorough manner aimed at deriving a comprehensive understanding of the underlying principles responsible for producing the observed phenomena (DiSessa, 1988; Driver, 1989; Driver, Guesne, & Tiberghien, 1985; West & Pines, 1985). In both science education research literature and physics education research literature, one will encounter many similar terms, each referring to the existence of notions regarding the concept of force prior to instruction. The abundance of essentially synonymous terms (see Appendix A for a partial listing) referring to the same phenomenon is indicative of the prevalence of its occurrence.

Research indicates that the pre-instruction understandings derived from a common-sense approach to operating within the demands of daily living can contain internal inconsistencies, which do not seem to be recognized by the individual (Clement, 1982; Halloun & Hestenes, 1985a; Redish, 1994). Traditional introductory physics instruction produces minimal change in the beliefs derived from common-sense (Clement, 1982), even when students have successfully learned problem-solving algorithms (Halloun & Hestenes,
Traditional instruction refers to a classroom in which the teacher presents the facts, and students utilize drill and practice activities leading to end of the chapter assessments (Russell, 1993).

The recognition that existing student understandings can mediate the effect of instruction, and that the traditional model of introductory physics instruction is ineffective at replacing student preconceptions with Newtonian alternatives for the majority of students (McDermott, 1984; Halloun & Hestenes, 1985b), led to the development of the FCI (Hestenes, Wells, & Swackhamer, 1992). The FCI (in Appendix B) is a thirty item multiple choice assessment that offers four Aristotelian (common-sense) choices, and one Newtonian response to each question. The FCI has become the standard formative assessment and instructional technique evaluation instrument within the physics education research community, when an evaluation of learning with respect to Newtonian Mechanics is conducted. It is commonly utilized in a pre-instruction, post-instruction pairing that is analyzed to measure student gain with respect to learning Newtonian concepts. Gain \( g \) (the average normalized gain) is calculated by dividing the actual average gain (posttest - % pretest) by the maximum possible average gain (100 - % pretest) (Hake, 1998a). The purpose of a formative assessment is to probe the state of student learning; then the teacher can adjust instruction and learning activities to meet student needs (Dufresne & Gerace, 2004). A summative assessment exists to provide a measure for a person outside of the classroom (Starkman, 2006). This could be a report card grade for a parent or an SAT score for college admission consideration.

Interactive teaching strategies can produce greater gains in conceptual understanding than traditional methods of instruction (Hake, 1998a; Mazur, 1997). Interactive teaching strategies are defined as “those designed at least in part to promote conceptual understanding through interactive engagement of students in heads-on (always) and hands-on (usually) activities which yield immediate feedback through discussion with peers and/or instructors” (Hake, 1998a, p. 65). Other advantages associated with interactive-engagement pedagogies are these: student-teacher and student-student discussions probe student conceptualizations, allowing for the recognition of and reflection on student prior notions (Mestre, 1994); the construction of explanations of ideas by students promotes conceptual change (Hewson & Beeth, 1993); the instructor gains insight into what students are thinking because they express it in words (Redish, 1994); more effective development of complex reasoning skills occurs (Johnson, Johnson & Smith, 1991). In addition, there is improved grasp of material (Crouch & Mazur, 2001); improved quantitative problem solving (Hake, 1998a; Thacker et al., 1994); significant reduction of the performance gap associated with gender (Lorenzo, Crouch, & Mazur, 2006); and higher success rates, particularly for females and minorities (Beichner, et al., 2007).

If all other factors were equal then there would exist little incentive for teachers to use traditional instructional techniques. But all other factors are not equal. Non-traditional instructional techniques require more time than traditional instruction and therefore result in less content coverage within a physics course (Crouch et al, 2007; Laws, 1991; Meltzer & Manivannan, 2002). In the high school environment there are physics courses that are often followed by nationally administered standardized tests. Examples include Advance Placement (AP) Physics leading to the AP Examination, and honors physics leading to the SAT II Physics Examination. These assessments draw upon a broad survey curriculum, and those students that are not exposed to the range of potential question topics may be disadvantaged compared to students that have received instruction in the entirety of the prescribed curriculum. Performance on these assessments may have consequence within the college admission process (Starkman, 2006).

In summary, research indicates that interactive engagement practices result in greater gains in conceptual understanding of Newtonian mechanics than traditional instructional techniques. The trend of increasing enrollment in high school physics courses brings with it an increased burden upon the pedagogical practices used within the classroom. The interactive engagement strategies require a slower pace of instruction than traditional instructional techniques. The competing yet reasonable demands of choosing instructional strategies that favor breadth of coverage to maximize the chance of a high score on the high stakes, end-of-course, assessment versus teaching less content to facilitate conceptual understanding of fewer topics through interactive-engagement methods is a decision that the high school physics teacher encounters.

THE PROBLEM

Some high school introductory physics courses have an end-of-course, national, high-stakes assessment associated with them that may influence the college acceptance process. These assessments are based on a broad range of curriculum topics that require a rapid
rate of progress through the curriculum if all content topics are to be covered. Traditional teaching strategies allow for more content coverage in a given time interval, making this instructional approach desirable in an environment where breadth of content coverage is of paramount importance. However, traditional instruction encourages rote memorization of facts and fails to connect classroom lessons with the students’ existing understandings (Mestre, 1994).

Changes in the post-secondary employment market have resulted in a greater percentage of the high school student population taking introductory physics courses. The broader range of student capabilities within introductory courses places an even greater burden on the method of instruction in the classroom with respect to facilitating student understanding. If the desired outcome of instruction is to go beyond rote memorization and include conceptual understanding for a greater percentage of the students within the classroom, then non-traditional instructional methodologies are required. The new pedagogical practices inspired by physics education research are called interactive-engagement strategies. IE teaching strategies result in greater conceptual understanding for students learning Newtonian mechanics in introductory physics courses (Hake, 1998a). The competing demands of a faster rate of content coverage to maximize the broad range of subject matter tested by high stakes national assessments and a slower rate of content coverage to maximize student conceptual understanding through interactive-engagement techniques create a dilemma for the instructor. This paper will consider the possibility that instruction that incorporates reflective journal writing by students may be able to reduce the content coverage differential between traditional and interactive-engagement instructional methodologies in introductory physics courses, while sustaining the types of gains associated with interactive-engagement instructional methodologies.

PERTINENT AREAS OF RESEARCH

Interactive-engagement strategies used during introductory physics instruction are discussed. The Force Concept Inventory, a formative assessment developed within the physics instructional reform movement as a tool to give feedback to the instructor upon completing instruction regarding Newtonian mechanics, is reviewed with respect to its more generalized role within the physics education research community. That role is as an instrument used to assess the effectiveness of an instructional methodology. Pedagogical reforms (called IE strategies) inspired by the same research data that resulted in the creation of the FCI are reviewed in relationship to their relative success at affecting gain in student understanding of Newtonian Mechanics. Commonalities between these interactive-engagement strategies are discussed.

Reflective journal writing, a student learning activity shown to contribute to conceptual understanding in other learning environments, such as pre-service training of teachers (Bell, 2001), business management, clinical psychology and higher education (Bolton, 2005), but currently underutilized in physics instruction, is reviewed. Interactive-engagement methodologies are not restricted to verbal interaction but may also include writing that reflects upon interpersonal as well as intrapersonal dialogue. The attributes associated with reflective journal writing are compared to those identified within interactive-engagement instructional methodologies. A research project involving the use of graded reflective journal entries, to be evaluated with respect to conceptual gain through the use of the FCI as a pre-instruction, post-instruction assessment, is proposed. The effect of reflective journal writing with regards to minimizing loss of curricular content coverage would also be evaluated.

INTERACTIVE-ENGAGEMENT STRATEGIES FACILITATE CONCEPTUAL UNDERSTANDING

The Force Concept Inventory Assesses Conceptual Understanding

Within the science education community physicists have led the way in measuring how well students learn the fundamental concepts in introductory courses (Stokstad, 2001). Within the physics education research (PER) community the most commonly utilized formative assessment following instruction on Newtonian mechanics is the Force Concept Inventory (FCI), (Hestenes, Wells, & Swackhammer, 1992). The instrument was developed to probe student beliefs regarding force, the central concept in Newtonian mechanics. The FCI consists of thirty multiple choice questions that cover the concept of force with respect to kinematics, Newton’s Laws of motion, gravitation, contact forces, and the superposition principle. Each question requires a choice from four answers associated with common sense understandings and one choice based on the Newtonian conceptualization. The four distractors were chosen from a compilation of the most common misconceptions articulated by students in exit interviews following instruction in introductory courses, over a number of years, at Arizona State University (Halloun &
Hestenes, 1985b; Hestenes, Wells, & Swackhamer, 1992). The conceptual nature of the questions, which are based on easily understood situations, allows the inventory to be administered prior to formal instruction in mechanics.

The FCI is published with a table that correlates the five previously identified Newtonian concepts with the specific questions that test for understanding regarding these concepts. Item analysis permits a view into whether Newtonian or non-Newtonian thinking regarding these concepts is exhibited across questions testing for understanding of the same concept. When the FCI is used by the instructor as a pre-instruction assessment of students’ existing beliefs regarding force and motion, the existence of common sense notions regarding force and motion is revealed, and a baseline is established against which instructional effectiveness can be calibrated through the comparison of pretest/posttest scores. When used exclusively following instruction, the class data can be analyzed to inform the instructor regarding the effectiveness of instruction comparatively between each of the five Newtonian concepts by which the questions are grouped. The FCI has been used in both instructional and research settings (Hestenes, Wells, & Swackhamer, 1992).

Upon publication, the instrument immediately went into use as a metric to measure the effectiveness of instruction; in fact, the majority of the original journal publication (Hestenes, Wells, & Swackhamer, 1992) is devoted to the analysis of the pretest/posttest data for eight high schools (N > 1500), and three universities (N > 500) that were the subjects of the study. The effectiveness of instruction was assessed through the analysis of the difference between the pretest scores and posttest scores, with the class average difference being considered the measure of instructional effectiveness. The authors contend that a score of 80% indicates the threshold of mastery of Newtonian concepts (regularly engaging in Newtonian thinking). A score of 60% can be considered to be the threshold indicator of Newtonian thinking (beginning to engage in Newtonian thinking) for students in an introductory physics course, and that without achieving a minimum understanding of Newtonian mechanics, quantitative problem-solving activities are meaningless exercises involving rote memorization of the algorithms associated with force and motion (Hestenes, Wells, Swackhamer, 1992). These percentages were arrived at by correlating the posttest FCI scores with a diagnostic mechanics test (the Mechanics Baseline Test; Hestenes & Wells, 1992) developed to test quantitative problem-solving skills following introductory instruction in Newtonian mechanics. Each participant in the study took both the FCI and the mechanics test upon completion of instruction in mechanics. The fact that the class average of correct responses on the post-instruction administration of the FCI for more than 40% of the participants in this study was below the minimum Newtonian threshold (a score of 60%), and that no class scored a class average of 80% correct on this test, caught the attention of the physics education research community (Mazur, 1997).

The Hake Gain Analysis Assesses Effectiveness of Instructional Methodologies

After the initial publication of the FCI, Richard Hake, Physics Professor at Indiana University, collected data over the next six years from high school, college, and university teachers of introductory physics courses, that had used the FCI in the prescribed pretest/posttest mode to gauge student learning regarding force. The publication of his meta-analysis of the 62 data sets (N = 6542) reports on the average effectiveness of a course at promoting conceptual understanding, which is defined as the average normalized gain <g> (Hake, 1998a). Average normalized gain for a course is defined as the ratio of the posttest class average % – pretest class average % to 100% – pretest class average. In reporting the results of the meta-analysis, Hake (1998a) grouped the data by instructional methodology for these courses. The range of scores for course averages was categorized as: high gain (≥ .7); medium gain (.7 > g ≥ .3); and low gain (< .3). The fourteen traditionally taught courses yielded an average gain of .23 while the forty-eight courses incorporating some form of interactive-engagement (IE) strategies yielded an average gain of .48, close to two standard deviations above the traditionally taught courses. Other researchers reporting similar average gain score differences between courses utilizing traditional and interactive-engagement methodologies include: Redish & Steinberg (1999); Meltzer & Manivannan (2002); Fagen, Crouch, & Mazur (2002); and Royuk & Brooks, (2003). These findings are credited by a number of introductory physics instructors with influencing curricular reforms regarding instructional methodologies (see Appendix C for a partial listing).

There were fourteen high school data sets reviewed by Hake (1998a, 1998b). Four of these courses were categorized as Traditional. The average normalized gain for these courses was .255. The ten IE data sets had an average normalized gain of .545 (which exceeds the gain average for all IE courses because a few of the community college and university IE courses yielded gain scores in the low category). Two of the traditionally
taught courses administered the Mechanics Baseline (MB) test (Hestenes & Wells, 1992) following instruction in mechanics; the average of the two class averages was 34.5%. Four of the IE taught courses administered the MB at the completion of mechanics instruction. The average of the class averages was 54.25%. These results are similar to those reported by Hestenes, Wells, and Swackhamer (1992) in that higher scores on the FCI correlated with higher scores of the MB, and that the MB scores ran about 15% on average lower than the scores on the FCI. Hake (1998a) reports a positive correlation coefficient \( r = +0.91 \) between the course average MB score and course average FCI score. In summarizing the comparative results for the entire meta-analysis, Hake (1998a) states that “The conceptual and problem-solving test results strongly suggest that the use of IE strategies can increase mechanics-course effectiveness well beyond that obtained with traditional methods” (pg 71).

Examples if IE Strategies Used in High School Physics Instruction

Two IE instructional approaches were used within all ten IE introductory high school physics courses analyzed in the meta-analysis (Hake, 1998a): peer instruction and microcomputer-based labs (MBL) (Thornton & Sokoloff, 1990). Peer instruction, as described by Mazur (1997), involves both requiring students to bring a written response to the assigned reading when reporting to class, and group discussion in lecture. A conceptual question based upon a common misconception follows each topic described during lecture. Approximately five minutes of small group discussion terminates when the teacher polls the class for answers and explanations, then moves on. The discussions and answers are not graded. The active engagement of the students with each other and with the teacher, as well as the students’ construction of explanations, has been credited with promoting conceptual understanding (Mazur, 1997). It was not reported if the collaborative peer instruction, as utilized in the ten high school classes, exactly mirrored the implementation described by Mazur, or some form of personalized adaptation by individual instructors that included small group discussion and whole class discussion around questions related to common student misconceptions. An advantage of group discussion which is focused on a question designed to draw out the existence of a common sense understanding, if the student holds such a notion, is that the conflict that exists between the Newtonian conceptualization and the student belief is brought to the forefront, and optimally to resolution. Requiring students to articulate and defend an understanding, in a mutually respectful manner, within first the group and then the class, striving for consensus, engages students simultaneously with both the concept being discussed and their personal belief regarding that concept (Crouch et al., 2007). A disadvantage of using the peer instruction technique in conjunction with traditional lecture presentation and conceptual tests, as practiced at Harvard University, is that the slower pace that allows for student discussion results in a 10-15% loss of content covered within the course. The Harvard physics courses reported on do have a high-stakes, end-of-course assessment associated with them, a somewhat analogous condition to many high school introductory physics courses. The high-stakes assessment for the Harvard students is the Medical College Admission Test. At Harvard the instructor’s solution to this dilemma, not covering the entire range of topics from which test questions can be selected, is advising the students at the beginning of the course that they will have to learn the other 10-15% of the physics on their own (Crouch, et al., 2007). The combination of the caliber of student matriculating at Harvard University and the use of traditional lecture within the course results in less content coverage loss (10-15%) than might be anticipated in many other instructional environments. Asking high school students to learn some part of the physics course on their own prior to sitting for a nationally administered examination may not yield the same result that occurs at Harvard University.

The micro-computer based laboratory (MBL) is the product of Tufts University’s Center for Mathematics and Science Education and was developed by Thornton and Sokoloff (1990) as a means of using technology to crystallize the conflict between student understanding and real-world outcomes for specific cases of motion.

These investigations employ the capacity of computers interfaced with motion detectors to display graphs of motion in real time. The laboratory investigations, which are used in conjunction with small group discussion surrounding results, and predict and compare activities, were found to be effective at reducing student misconceptions and improving students’ conceptual understanding of force and motion (Laws, 1991; Thornton & Sokoloff, 1990). Again Hake (1998b) does not report on the degree of alignment of the high school implementations with respect to how these materials were described in educational research publications (Laws, 1991; Thornton & Sokoloff, 1990; Sokoloff & Thornton, 1997). Hake (1998a) does report that he collected the original data by soliciting FCI pretest/posttest data sets at colloquia and meetings as
well as through e-mail postings on the PHYS_L and PhysLrnR nets between 1992 and 1998. It seems fair to assume that most if not all of the high school instructors responding to these requests had probably become aware of the MBL program through involvement with PER literature or by participating in National Science Foundation funded summer workshops during that time period that trained teachers to use the MBL materials.

There are multiple advantages associated with using the MBL with its capacity to graph, in real time, the complete set of motion graphs (displacement, velocity, and acceleration, each with respect to time) stacked for easy comparison. These include elimination of student error that results from poor transposition of data recorded in table form to graphical representation, and instantaneous results which allow for immediate discussion between lab team members rather than the students at some later time constructing the graphs (a time-involved activity) without lab team members present and available for discussion. The time gained through the computer program’s instantaneous graphing capability is used in discussion both informally within the groups and formally by responding to the structured questions associated with these labs. Discussion of and reflection on the three aspects of motion represented graphically present the opportunity for the student to experience a discrepant event. A discrepant event is an outcome that contradicts the outcome the student expects based upon the existing understanding (Sokoloff & Thornton, 1997). Placing students in the position of resolving discrepancies between the common-sense based predictions and a witnessed outcome is an objective of physics reform materials in general, and the MBL activities are an example of this approach to promoting conceptual understanding. In relating the content coverage for a university introductory course taught traditionally one year to the same course taught by the same instructor the following year using the MBL activities and other PER inspired reforms, a 25% reduction in content coverage was noted (Laws, 1991).

Half of the high school IE courses in this study employed the modeling theory approach to teaching physics which has become the focus of the physics education research group at Arizona State University. The modeling theory of teaching physics states that physics is based on a small set of models that represent the structure seen in the world. Models are explicitly identified as models of reality, not physical realities (Hestenes, 1987). Modeling was developed to address identified conceptual deficiencies commonly present in students that have completed traditional physics courses. These deficiencies include fragmented physics knowledge and the persistence of alternative conceptions which are attributable to the tendency of students to focus on the memorization of facts and algorithms rather than understanding the processes involved (Hestenes, 1987). Through group discussion students are required to identify a question that might be asked of nature by a scientist. Then, in small groups, experiments are designed and conducted. The experiment is run, data are collected and analyzed, and students make claims based on their findings. The instructor is expected to be knowledgeable with respect to common misconceptions, and able to employ Socratic questioning to prompt and guide the inquiry. The instructor must also take care in choosing as few paradigms as possible to investigate so that students will have enough time to learn and understand the concepts as well as have the time to resolve conflicts between their misconceptions and the accepted scientific conceptualizations (Hestenes, 1998). As all of the IE high school courses reviewed in the study (Hake, 1998a) used the MBL materials, modeling is a good pedagogical fit that should interface seamlessly between the classroom and lab setting. Once again this specific reform is based on acknowledging the prevalence of common sense notions of force and motion and using group discussion to mediate between those understandings and the accepted scientific conceptualization. The requirement of more time per topic to allow for the discussion and reflection associated with promoting conceptual change results, however, in less content coverage.

Just after the publication of Hake’s meta-analysis in January of 1998, a detailed comparison between traditional introductory physics instruction and introductory physics instruction utilizing active learning strategies was undertaken. This study occurred during the fall semester of 1998 at California Polytechnic State University (Knight, 2002). California Polytechnic contributed four data sets to the Hake (1998a) meta-analysis. Three of the sets reported were categorized as IE, and each of these courses was taught by Randall Knight, who conducted but did not teach any of the sections for the comparative study (Knight, 2002). The other data set from Cal Poly reported on by Hake (1998b) was categorized as traditional in format. For the comparative study (Knight, 2002) the students were randomly assigned to one of six introductory sections (N = 225). For this investigation three professors each taught one traditional section and one studio classroom session. The studio classroom is designed to eliminate the boundary between lecture and laboratory and promote
an active learning environment. This is facilitated by the elimination of lecture and having lab and other activities occur in the same classroom. The active learning strategies included MBL (Thornton & Sokoloff, 1990), group problem solving involving worksheets based on PER findings, and group discussion in the think/pair/share format. The studio class used a research inspired text for outside of class readings. The traditionally taught class used a traditional introductory physics text. As a historical aside, the elimination of lecture from introductory physics courses as a reform aimed at improving students’ conceptual understanding cannot be called a recent innovation. In an autobiography, Robert Millikan, the man credited with determining the quantity of electrical charge on an electron, describes elimination of lecture from the introductory courses he taught, because he found lecture ineffective at helping students understand physics. Millikan instituted this reform at the University of Chicago in 1896 (Millikan, 1950).

For all students in the Knight (2002) study the FCI was administered as a pretest on the first day of class, and no differences were found between the traditional and non-traditional sections prior to instruction. At the end of the quarter all students were given the FCI as a posttest and all students took the same final exam. The final exam, described as a typical quantitative examination, was written by a professor not involved in the comparison. Student examinations were identified by a number and each problem on the exam was corrected by only one faculty member to assure fairness in correcting. The findings of this comparison were similar to Hake’s meta-analysis (1998a). The students who learned physics in the active-learning environment did significantly better on both the FCI posttest (11% average difference) and on the quantitative final exam (7.5% average difference). A comparison of overall instructor rating was also tabulated through the use of the standard course evaluation questionnaire, which used a 1-5 scale. The same three instructors taught one section in each format yet received an average evaluation of 3.51 for the traditional course and an average of 3.92 for the studio class. The same instructors were evaluated higher in the active-learning environment. The results of this study support the contention of Hake (1998a) that IE methods facilitate greater gains than traditional instruction in conceptual understanding related to the Newtonian concept of force, and also provide evidence for the claim that greater conceptual understanding of Newtonian mechanics results in improved quantitative problem-solving skills within this part of an introductory physics course.

This study did not report on the comparative time-rate of content coverage, or on the number of topics studied during the semester. The fact that all six sections took the same final examination is indirect evidence that the content coverage was essentially equivalent. The elimination of lecture by the professors in the active learning courses may account for some of the reduction of time differential reported in other comparisons of content coverage between the discussion-rich IE courses and the lecture-based traditional instruction method. In another university setting the elimination of lecture and a move to MBL based investigations supported by student discussion resulted in a 25% reduction of content coverage (Laws, 1991), so merely eliminating lecture doesn’t always result in a pace comparable to that associated with traditional introductory physics instruction.

**Summary**

The three studies reviewed (N= 9000) have each concluded, through analysis of assessment data, that students who develop a better conceptual understanding of force as measured by the FCI (Hestenes, Wells, & Swackhamer, 1992) will also exhibit superior quantitative problem-solving skills regarding Newtonian mechanics. Each of the three instructional reforms reviewed is designed to facilitate conceptual understanding by addressing existing common-sense based notions that were derived from life experiences prior to formal instruction in physics. The mechanism for addressing existing misconceptions that each of these interactive-engagement strategies employ involves students actively discussing their reasoning. The process includes evaluating discussion as well as constructing knowledge, and developing the ability to construct arguments. The give and take that occurs between students during discussion requires that students reflect upon the reasoning they have used in constructing the understanding that they present and defend during the discussion process. The role of the teacher is to act as a facilitator rather than as a transmitter of information. A large percentage of class-time is utilized to develop conceptual understanding. This focus on negotiating meaning through the articulation of student understandings results in less content being covered than in courses taught in a traditional manner.

The physics instructional reform movement grew out of the recognition that students enter into instructions with existing understandings that do not align with scientific conceptualizations regarding the topics studied. Newtonian mechanics is one of the central concepts of introductory physics courses. Research has
established that traditional physics instruction does little to change the existing understandings that students hold, both in general and specifically in the case of the concepts of force and motion. The instructional reforms aimed at improving conceptual understanding, which have been found to be highly correlated with computational problem solving ability, have been called interactive-engagement methods. These reforms are rooted in the constructivist view of learning that posits that students construct their understandings through social interactions with each other and the instructor, as well as through reflection on the relationship between their understanding and the views of others being presented (Piaget, 1964, 1970; Vygotsky, 1962, 1978).

In describing the role of instruction within interactive-engagement pedagogies, Mestre (2005) relates that the teacher must probe student understanding to check that instruction is leading to students constructing knowledge that aligns with accepted scientific conceptualizations. It has been argued (Mazur, 1997) that without probing student understanding the instructor is left with illusions regarding student learning rather than insight that can inform instruction. The discussion component of the specific interactive-engagement methodologies previously described is designed in part to probe student understandings. Verbal dialogue is not the only mode through which the instructor can probe student conceptualizations. In a publication relating findings from cognitive studies to the teaching of physics, Redish (1994) states that “to find out what our students know we have to give them the opportunity to explain what they are thinking in words” (p. 798).

**REFLECTIVE WRITING FACILITATES CONCEPTUAL UNDERSTANDING**

Student reflective writing offers the instructor the opportunity to appraise student understanding. Reflective writing is informal writing that aids students in acquiring ownership of ideas (Connally & Vilardi, 1989). Researchers reporting on reflective writing in the secondary science classroom (Keys, et al., 1999) have concluded that when teachers encourage students to write, the students engage in metacognition and that writing facilitates the generation of meaning. Metacognition has been defined by Baird as “knowledge awareness and control of one’s own learning” (as cited in Etkina, 2000, p. 604). Metacognition also includes the ability to decide when one’s own level of understanding is inadequate (Mestre, 2005).

**Graded Reflective Writing**

Research has reported that unless something is graded students are not likely to buy into it (Resnick & Resnick, 1992) and that writing for credit yields a better quality of writing, which correlates with higher test scores than writing that receives credit intermittently, or not at all (Hautau et al., 2006). The benefits associated with students engaging in reflective writing combined with the motivational advantage of grading activities that students engage in suggest that graded reflective writing may be a productive interactive-engagement strategy. One specific activity that combines reflective writing with grading is the graded reflective journal. The National Research Council (1995) describes reflective journal writing as an activity in which students assess their learning. In a later publication (National Research Council, 2001) reflective journaling is described as a formative assessment as well as an opportunity for students to make connections between the topic of reflection and real world understandings. A review of the literature has revealed that reflective writing is an under-utilized strategy in introductory physics courses. The few studies reporting on student reflective writing in introductory physics courses all involved the university setting.

Research regarding the use of a structured journal in an introductory university physics course has been reported (Etkina, 2000). Students were required to complete a weekly reflection. Students were guided in terms of what to reflect upon by responding to three questions within their reflections. The questions were suggested by Arons (1990) in a book that summarized research, observations, and conclusions gleaned from more than thirty years of teaching introductory physics to college and university students. The questions were: What did you learn? What remains unclear? If you were the professor, what questions would you ask to determine if students understand the material? The maximum score on each weekly report was six points, which was equal in value to the homework component of the course during each week. The final grade for the course was based on a points system. Extra credit was given for interesting questions.

The reports allowed for communication with each student on an individual basis through written commentary. The findings include a 90% match between what students reported having learned and the teacher expectations regarding these statements. Frequently students reported some concept as learned and followed that with an indication that it had not really been learned because the same idea reported on as learned was identified in the response to the question on what
remains unclear. In the conclusion Etkina (2000) reports on how the weekly reports impacted her teaching. Specific instances cited include: developing an understanding of topics that presented conceptual difficulties for the students; being allowed insight into the misconceptions held by those students; and reducing the amount of lecture in the course to allow for more discussion around questions culled from the weekly reports.

It was also stated that the weekly reports made student thinking visible by allowing insight into how students were thinking about their learning as well as detailed exploration of student’s knowledge creation. The contention is made that weekly reports are a practical system to facilitate the development of metacognitive skills. There is no evaluation given to the weekly reporting requirement’s impact on course content coverage, but it was noted that the grading of the reports increased the correcting time for the instructor. It was also reported that the mismatch between what the instructor assumed the students had learned and what the reports revealed as remaining unclear prompted the reduction in the lecture component of the instruction (Etikina, 2002). There was no end of course, high-stakes examination of the kind previously identified associated with other courses. This study was the first of three that occurred during a three year period that utilized the same question set, responded to in a weekly, graded report. Slightly different introductory physics university student populations were reported on in each study.

In the second study of the series (May & Etkina, 2002) the subjects were honors engineering students (SAT mathematics test average score = 680) in an introductory physics course that spent ten weeks studying mechanics. This is 50% of the course’s allotted time, which is between 10% and 20% longer than would be expected in a high school course teaching students of similar ability. In addition to requiring reflective writing, other interactive-engagement strategies were extensively employed. A fourth question, “What did you learn in lab this week, and how did you learn it?” was added to previously used guiding questions. A coding scheme of fourteen items grouped into three categories was developed as a grading rubric. The reports were evaluated regarding the quantity and quality of writing. The weekly reflective reports accounted for 10% of the course grade. All students were given the FCI in the pretest/posttest manner and the Hake gain analysis was employed. From the half of the class that had the lowest pretest scores twelve students were selected, since these students had the greater opportunity to exhibit gain in conceptual understanding, as measured by the FCI. The twelve subjects selected for evaluation were the six highest and six lowest gain values achieved within this half of the group. The average gain of the six lowest was .228, a result comparable to students receiving traditional introductory instruction in the studies that compared gain to method of instruction (Hake, 1998a, 1998b; Knight, 2002). The average gain of the six highest was .71 which Hake (1998a) would have classified as high gain. The average gain for the entire group was not reported but if the gain for these twelve students was averaged the .469 result would mask a significant difference between the two groups.

For these twelve students, the scores achieved on the weekly reports were compared to the gain in conceptual understanding as indicated by the Hake gain score from the FCI. The results were interpreted to suggest a possible correlation between the conceptual gain and a student’s ability to reflect on learning. Specific findings reported that the high gain students wrote more about how they learned, they referred to the reasoning involved in developing understanding more frequently and they also showed a greater concern for the coherence of knowledge than the lower gain students. In contrast, it was observed that the lower gain students tended to focus on knowing formulas without indicating a comprehension of their meaning as evidence of new knowledge, and relied upon teacher authority as the source of understanding. More extensive research is warranted if these claims are to be accepted. A follow-up study was begun to investigate possible corroboration of these findings.

The third study in this chronological sequence involved two introductory physics courses for engineering students designated as at-risk (Etkina & Eisner, 2003). This study was designed to determine if the research finding of the previous study (May & Etkina, 2002) could be generalized. The populations of the two studies were of dissimilar abilities as exhibited by the at-risk designation as opposed to the honors designation and the difference in mathematical performance as indicated by the mathematics average score of 480 on the SAT for the at risk group. The research investigated how students of lower math and reading skills would report on their learning, as well as investigating the result of reflection upon learning for this population. The required graded weekly report as previously described (Etikina, 2000; May & Etkina, 2002) and the coding rubric developed during the previous study (May & Etkina, 2002) were employed. The introductory courses for this study differed from previous studies in that the courses took place over two semesters covering the full year of introductory physics rather than only the
first semester, which includes mechanics, and the course focused on lab investigations as the basis for claims of knowledge. The FCI was used as pretest, but as concepts covered by the FCI accounted for only 10% of the course time, a novel method of calculating gain was employed. Gain was calculated as the quotient of the difference between the decimal average of the final exams for each semester and the decimal score of the FCI to the difference between one and the FCI decimal score. This interpretation supposes that the pre-instruction understanding regarding mechanics is an accurate indicator of a generalized baseline of understanding for all concepts studied in an introductory course. There is no discussion within the report regarding assigning the FCI pretest result as the baseline of understanding for all understanding of topics studied in an introductory physics course. This is certainly a debatable supposition. The range of scores that were categorized as low and the range of scores that were categorized as high were not reported. No indication of the relationship of the low gain versus high gain to those categorizations as identified by Hake (1998a) with numerical score was reported either.

It was reported that both categories of students indicated concepts were more important than equations for developing understanding, and that reflection on the reasoning used to understand experimental resulted in improved conceptual understanding. Differences reported were that high gainers more frequently cited learning from authority (lecture and formulas) more often than low gainers (which stands in contrast to findings for these categories within the study of honors students), and that low gainers tended to focus on what they learned rather than how they learned it. Etkina and Eisner (2003) claim that this focus is evidence that low gain students lack metacognitive skills, which may contribute to the lower gain for this group. The student population that is the focus of this study may be similar to the segment of the overall student population from which the increased introductory physics enrollment referred to by Hehn and Neuschatz (2006) is being drawn. If that is so then the findings regarding this study are of particular interest to instructors in academic environments where introductory physics courses are experiencing increasing enrollment.

Each of the studies that have been referenced to this point in the paper comes from the PER community. This group of researchers is familiar with the same body of research findings reported on over the past forty years relating to physics instruction and learning. Within this community a great deal of credence is placed on the findings reported through the use of the FCI. The acceptance of Hake's (1998a) contentions regarding the comparative effectiveness of traditional instruction versus interactive-engagement methodologies is near universal. Evidence for this contention includes the citation of the same core of researchers in the reference section of the published findings, and the experiments conducted all attempt to increase conceptual understanding with respect to Newtonian mechanics through instructional reform that includes attempting to address common-sense understandings through some form of external or internal dialogue. The last study reviewed comes out of a different research tradition, the writing across the curriculum movement.

A study investigating how students perceived and accomplished reflective writing focused on assigned introductory physics text readings was conducted (Kalman et al., 2008). The readings and reflections occurred prior to instruction, which was traditional lecture format mechanics course. At the beginning of the course one class period was dedicated to instruction regarding the reflective writing. The goal of the reflective writing was to capture metacognitive reflection occurring during the reading activity. Each week the writing assignments, which were not graded, and three sentences, which were graded, were collected. The three sentences were based on important ideas identified in the jottings recording the reflections upon the readings. The sentences were graded and the best nine sentences submitted during the semester accounted for 15% of the course grade.

From the half of the class (class total = 100) that volunteered to participate in the study five students were chosen. The reflective writing was evaluated with respect to Bereiter and Scardamalia's knowledge transformation model of writing (1987). This model purports that students increase their conceptual understanding through engaging in reflective writing. Reflective writing will involve knowledge transformation which is characterized by structures that include comparison, argument, and explanation that includes the identification of new ideas. Writing that is based on a recitation of facts from the reading or existing knowledge is categorized as knowledge-telling and is not considered evidence of metacognitive reflection. Coding of reflective writing samples was done for only 25% of the student samples. The five subjects were interviewed about their reflective writing at the third, seventh, and thirteenth (final) week of the course.

All students claimed that engaging in reflective writing helped them prepare for class. Analysis of the writing samples led to the claim that three of the five students showed evidence of recognizing conceptual
difficulties through the act of reflective writing and the same number showed evidence of a positive outcome, evidence of an ability to inductively solve new problems while engaging in reflective writing. It is claimed (Kalman et al., 2008) that reflective writing promotes self-dialogue about existing knowledge and new concepts encountered in textbook reading.

This is a very small sample size. The report does not offer insight into why only five of the fifty students that volunteered to participate in the investigation were selected. The time required evaluating writing samples, grading the submitted sentences, and interviewing participants may be involved in the restriction of the research to only five participants. No reporting regarding course or test grades, the relationship of graded writing to course grade, or the relationship of student perceptions, as revealed through interviews, to course grade was included. The omission of the potential link between metacognition as indicated by reflective writing and the overall course performance, or the relationship of writing indicating a lack of metacognition and overall course performance suggests an area for further research. This paper makes no reference to any research reporting on differences between traditional instruction and interactive-engagement methods of instruction. As the course involved what has been described as traditional instruction, other than the use of the reflective writing, no reporting on the content coverage aspect of the course was expected or reported.

**Summary**

Researchers have reported that graded reflective writing, structured by questions or instructions that guide students with respect to focus, has multiple benefits for students when learning Newtonian mechanics. A particular benefit described is the facilitation of the development of metacognitive skill, a critical capability for learning in this topical area because of the prevalence of common-sense derived understandings that resist traditional instruction and oppose the accepted Newtonian conceptualizations. Evidence has been presented that reflection can and does lead students to the recognition of a discrepancy between the understanding held and the scientifically supported conceptualization. Recognition of the discrepancy is the first step toward reconciliation of conflicting notions. It is also reported that reflective writing facilitates the resolution of conflicting notions by promoting self-dialogue about the existing knowledge and the new concept. Another benefit claimed is that graded reflective writing facilitates the development of inductive problem solving capabilities. It was reported that all students benefit from graded reflective writing focused on learning, and that many students perceive reflective writing as contributing to knowledge transformation.

Differences were reported regarding the type of writing exhibited as well as the focus of the writing within each study. When the focus of the reflective writing was correlated to gains in conceptual understanding as quantified through the use of the FCI and the Hake gain analysis, it was found that the high gain students within a class of high academic achievers tended to write about the reasoning they used to develop understanding. This apparent correlation between gain in conceptual understanding and reflection on learning may indicate that concern for coherence of knowledge is a factor involved in achieving larger Hake gains. The lower gain students within this same class often cited authority (formulas or teacher statements) as the basis for developing new understandings.

In a parallel study of much lower academic achievers it was found that the high gain students of this group often wrote about authority as the source of understanding. This result which contradicts the findings for students with higher mathematical and verbal skills, as measured by the SAT, led to the conjecture that students of different academic abilities may have differences in the way that they reflect on learning and that instruction should take this into account.

It is also reported that reflective writing creates a space for dialogue between the teacher and student. The writing can make students’ thinking visible which permits insight and comment by the teacher that may not occur otherwise. Changes in student perception can be captured, and if a particular learning activity is extremely effective the instructor could receive feedback through examination of student reflections. Another attribute of corrected reflective writing is that it allows the instructor to have insight into difficulties being experienced by a large number of class participants which can result in adjustment to the pace of instruction to accommodate resolution of the problem before moving on. Due to the informative capability of the reflective journal it is clear that a student reflective journal is a formative assessment, as well as a learning activity. In summary, graded reflective writing can benefit both the learner and the instructor.

**CONCLUSION AND POTENTIAL RESEARCH QUESTIONS**

Interactive-engagement strategies were developed to respond to the less than optimal effectiveness of traditional instruction at facilitating conceptual understand-
ing regarding Newtonian mechanics for students in introductory courses, as measured by the FCI. Interactive-engagement strategies seem to be more effective at moving students from the common-sense based understandings regarding force and motion that they entered into instruction believing, to the Newtonian understandings of force and motion. Physics education researchers have presented evidence that interactive-engagement strategies are more effective at facilitating conceptual understanding and quantitative problem solving capabilities in students than are traditional instructional strategies, for students studying Newtonian mechanics in an introductory course. Interactive-engagement strategies can use a number of events to generate discussion among students. Interactive-engagement strategies use student discussion as the medium through which a discrepancy between an existing understanding held by a student and the understanding being presented, discussed, and defended by other students is crystallized. In reflecting on this discrepancy, the holder of the misconception is forced to negotiate between the competing notions through the construction of arguments to defend a position. The construction of the arguments parallels the construction of knowledge. This is a time involved process that requires external arguments to become internalized, and optimally the discrepancy is resolved through achieving a Newtonian understanding. The discussion process takes up a significant portion of the time that the students and the instructor share. Metacognitive reflection, during and after discussion, is required if the student is to gain the Newtonian understanding.

Metacognitive reflection is also the goal of graded student reflective writing. Research indicates that when the students are guided to reflect on the learning process, metacognitive skills can be developed. When the students engages in self-dialogue regarding the discrepancy between what they have held as true and the new information being presented, and reflected upon, knowledge transformation can occur. Graded reflective writing moves the intrapersonal discussion outside of the classroom and conserves some of the precious student-teacher contact time for progress through the curriculum. Both interactive-engagement strategies and graded reflective writing require the student to engage in metacognitive reflection in order to transition from a common-sense based understanding to the Newtonian conceptualization. Because the reflective process is engaged in outside of the classroom for the reflective writing strategy, it may be worth examining the effect that grading student reflective journals has on improving conceptual understanding and upon reducing the time differential that exists between traditional and interactive-engagement teaching strategies.

Research indicates that the trend of increasing enrollment in introductory physics courses will continue. Researchers have found that all students benefit by engaging in graded reflective writing. It has also been suggested that students’ metacognitive skills, skills which are required to move from common-sense based understanding to Newtonian thinking, are enhanced through the practice of reflective writing focused on learning. Research reporting on the effect of graded reflective writing on students that have academic capabilities spanning a wide range suggests that students whose mathematics and verbal skills, as indicated by SAT scores, are toward the opposite ends of the range may learn differently. Allowing the instructor insight into the students’ perceptions upon learning, as captured in the reflections recorded, may permit the instructor to adjust instruction to the advantage of the learner. Because of the expanding enrollment within introductory physics courses the effectiveness of the instruction at accommodating students of all capabilities becomes of even greater importance than has historically been the case.

Further research on the potential of reflective writing to facilitate the development of conceptual understanding for students studying Newtonian mechanics in an introductory course is suggested. Based upon the research findings cited, a proposed study using graded student reflective writing is proposed. Reflective writing engaged in on a daily basis and summarized through a weekly summative reflective paragraph may lead to insight regarding the following questions:

1. How will graded reflective journals compare to interactive-engagement strategies, with respect to conceptual gain for students in an introductory physics course, as measured with the FCI using the Hake gain analysis?
2. What effect will using graded reflective journals have upon the rate of content coverage within introductory physics courses using interactive-engagement strategies?
3. What effect will grading reflective journals have upon the instructor regarding pace of instruction?
4. What will be found regarding the type of reflection (knowledge building or knowledge telling) and focus of reflection (internal dialogue or external authority) for high gain versus low gain students as grouped by the Hake gain analysis of pre/posttest FCI scores.
References


*American Journal of Physics, 70* (6), 639-654.


---

**APPENDIX A**

**TERMS USED TO DESCRIBE “COMMON-SENSE” UNDERSTANDINGS**

<table>
<thead>
<tr>
<th>Term</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative conception</td>
<td>Driver &amp; Easley, 1978</td>
</tr>
<tr>
<td>Alternative framework</td>
<td>Driver &amp; Easley, 1978</td>
</tr>
<tr>
<td>Intuitive science</td>
<td>Preece, 1984</td>
</tr>
<tr>
<td>Naïve belief</td>
<td>Carramazza, McCloskey, &amp; Green, 1981</td>
</tr>
<tr>
<td>Naïve conception</td>
<td>Champagne, Klopfer, &amp; Anderson, 1980</td>
</tr>
<tr>
<td>Naïve theory</td>
<td>Resnick, 1983</td>
</tr>
<tr>
<td>Misconception</td>
<td>Posner, Strike, Hewson, &amp; Gerzog, 1982</td>
</tr>
<tr>
<td>Phenomenological primitive (p-prim)</td>
<td>DiSessa, 1988</td>
</tr>
<tr>
<td>Preconception</td>
<td>Clement, 1982</td>
</tr>
</tbody>
</table>
APPENDIX B
THE FORCE CONCEPT INVENTORY

1. Two metal balls are the same size but one weighs twice as much as the other. The balls are dropped from the roof of a single story building at the same instant of time. The time it takes the balls to reach the ground below will be:
   (A) about half as long for the heavier ball as for the lighter one.
   (B) about half as long for the lighter ball as for the heavier one.
   (C) about the same for both balls.
   (D) considerably less for the heavier ball, but not necessarily half as long.
   (E) considerably less for the lighter ball, but not necessarily half as long.

2. The two metal balls of the previous problem roll off a horizontal table with the same speed. In this situation:
   (A) both balls hit the floor at approximately the same horizontal distance from the base of the table.
   (B) the heavier ball hits the floor at about half the horizontal distance from the base of the table than does the lighter ball.
   (C) the lighter ball hits the floor at about half the horizontal distance from the base of the table than does the heavier ball.
   (D) the heavier ball hits the floor considerably closer to the base of the table than the lighter ball, but not necessarily at half the horizontal distance.
   (E) the lighter ball hits the floor considerably closer to the base of the table than the heavier ball, but not necessarily at half the horizontal distance.

3. A stone dropped from the roof of a single story building to the surface of the earth:
   (A) reaches a maximum speed quite soon after release and then falls at a constant speed thereafter.
   (B) speeds up as it falls because the gravitational attraction gets considerably stronger as the stone gets closer to the earth.
   (C) speeds up because of an almost constant force of gravity acting upon it.
   (D) falls because of the natural tendency of all objects to rest on the surface of the earth.
   (E) falls because of the combined effects of the force of gravity pushing it downward and the force of the air pushing it downward.

4. A large truck collides head-on with a small compact car. During the collision:
   (A) the truck exerts a greater amount of force on the car than the car exerts on the truck.
   (B) the car exerts a greater amount of force on the truck than the truck exerts on the car.
   (C) neither exerts a force on the other, the car gets smashed simply because it gets in the way of the truck.
   (D) the truck exerts a force on the car but the car does not exert a force on the truck.
   (E) the truck exerts the same amount of force on the car as the car exerts on the truck.
USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT TWO QUESTIONS (5 and 6).

The accompanying figure shows a frictionless channel in the shape of a segment of a circle with center at "O". The channel has been anchored to a frictionless horizontal table top. You are looking down at the table. Forces exerted by the air are negligible. A ball is shot at high speed into the channel at "p" and exits at "r."

5. Consider the following distinct forces:

1. A downward force of gravity.
2. A force exerted by the channel pointing from q to O.
3. A force in the direction of motion.
4. A force pointing from O to q.

Which of the above forces is (are) acting on the ball when it is within the frictionless channel at position "q"?

(A) 1 only.
(B) 1 and 2.
(C) 1 and 3.
(D) 1, 2, and 3.
(E) 1, 3, and 4.

6. Which path in the figure at right would the ball most closely follow after it exits the channel at "r" and moves across the frictionless table top?

7. A steel ball is attached to a string and is swung in a circular path in a horizontal plane as illustrated in the accompanying figure.

At the point P indicated in the figure, the string suddenly breaks near the ball.

If these events are observed from directly above as in the figure, which path would the ball most closely follow after the string breaks?
USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT FOUR QUESTIONS (8 through 11).

The figure depicts a hockey puck sliding with constant speed $v_0$ in a straight line from point "a" to point "b" on a frictionless horizontal surface. Forces exerted by the air are negligible. You are looking down on the puck. When the puck reaches point "b," it receives a swift horizontal kick in the direction of the heavy print arrow. Had the puck been at rest at point "b," then the kick would have set the puck in horizontal motion with a speed $v_k$ in the direction of the kick.

8. Which of the paths below would the puck most closely follow after receiving the kick?

9. The speed of the puck just after it receives the kick is:
   (A) equal to the speed "$v_0$" it had before it received the kick.
   (B) equal to the speed "$v_k$" resulting from the kick and independent of the speed "$v_0$".
   (C) equal to the arithmetic sum of the speeds "$v_0$" and "$v_k$".
   (D) smaller than either of the speeds "$v_0$" or "$v_k$".
   (E) greater than either of the speeds "$v_0$" or "$v_k$", but less than the arithmetic sum of these two speeds.

10. Along the frictionless path you have chosen in question 8, the speed of the puck after receiving the kick:
    (A) is constant.
    (B) continuously increases.
    (C) continuously decreases.
    (D) increases for a while and decreases thereafter.
    (E) is constant for a while and decreases thereafter.

11. Along the frictionless path you have chosen in question 8, the main force(s) acting on the puck after receiving the kick is (are):
    (A) a downward force of gravity.
    (B) a downward force of gravity, and a horizontal force in the direction of motion.
    (C) a downward force of gravity, an upward force exerted by the surface, and a horizontal force in the direction of motion.
    (D) a downward force of gravity and an upward force exerted by the surface.
    (E) none. (No forces act on the puck.)
12. A ball is fired by a cannon from the top of a cliff as shown in the figure below. Which of the paths would the cannon ball most closely follow?

13. A boy throws a steel ball straight up. Consider the motion of the ball only after it has left the boy's hand but before it touches the ground, and assume that forces exerted by the air are negligible. For these conditions, the force(s) acting on the ball is (are):

(A) a downward force of gravity along with a steadily decreasing upward force.

(B) a steadily decreasing upward force from the moment it leaves the boy’s hand until it reaches its highest point; on the way down there is a steadily increasing downward force of gravity as the object gets closer to the earth.

(C) an almost constant downward force of gravity along with an upward force that steadily decreases until the ball reaches its highest point; on the way down there is only a constant downward force of gravity.

(D) an almost constant downward force of gravity only.

(E) none of the above. The ball falls back to ground because of its natural tendency to rest on the surface of the earth.

14. A bowling ball accidentally falls out of the cargo bay of an airliner as it flies along in a horizontal direction. As observed by a person standing on the ground and viewing the plane as in the figure at right, which path would the bowling ball most closely follow after leaving the airplane?
USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT TWO QUESTIONS (15 and 16).

A large truck breaks down out on the road and receives a push back into town by a small compact car as shown in the figure below.

![Diagram of a truck and a car pushing each other]

15. While the car, still pushing the truck, is speeding up to get up to cruising speed:
   (A) the amount of force with which the car pushes on the truck is equal to that with which the truck pushes back on the car.
   (B) the amount of force with which the car pushes on the truck is smaller than that with which the truck pushes back on the car.
   (C) the amount of force with which the car pushes on the truck is greater than that with which the truck pushes back on the car.
   (D) the car's engine is running so the car pushes against the truck, but the truck's engine is not running so the truck cannot push back against the car. The truck is pushed forward simply because it is in the way of the car.
   (E) neither the car nor the truck exert any force on the other. The truck is pushed forward simply because it is in the way of the car.

16. After the car reaches the constant cruising speed at which its driver wishes to push the truck:
   (A) the amount of force with which the car pushes on the truck is equal to that with which the truck pushes back on the car.
   (B) the amount of force with which the car pushes on the truck is smaller than that with which the truck pushes back on the car.
   (C) the amount of force with which the car pushes on the truck is greater than that with which the truck pushes back on the car.
   (D) the car's engine is running so the car pushes against the truck, but the truck's engine is not running so the truck cannot push back against the car. The truck is pushed forward simply because it is in the way of the car.
   (E) neither the car nor the truck exert any force on the other. The truck is pushed forward simply because it is in the way of the car.
17. An elevator is being lifted up an elevator shaft at a constant speed by a steel cable as shown in the figure below. All frictional effects are negligible. In this situation, forces on the elevator are such that:

(A) the upward force by the cable is greater than the downward force of gravity.
(B) the upward force by the cable is equal to the downward force of gravity.
(C) the upward force by the cable is smaller than the downward force of gravity.
(D) the upward force by the cable is greater than the sum of the downward force of gravity and a downward force due to the air.
(E) none of the above. (The elevator goes up because the cable is being shortened, not because an upward force is exerted on the elevator by the cable).

18. The figure below shows a boy swinging on a rope, starting at a point higher than A. Consider the following distinct forces:

1. A downward force of gravity.
2. A force exerted by the rope pointing from A to O.
3. A force in the direction of the boy’s motion.
4. A force pointing from O to A.

Which of the above forces is (are) acting on the boy when he is at position A?

(A) 1 only.
(B) 1 and 2.
(C) 1 and 3.
(D) 1, 2, and 3.
(E) 1, 3, and 4.
19. The positions of two blocks at successive 0.20-second time intervals are represented by the numbered squares in the figure below. The blocks are moving toward the right.

Do the blocks ever have the same speed?
(A) No.
(B) Yes, at instant 2.
(C) Yes, at instant 5.
(D) Yes, at instants 2 and 5.
(E) Yes, at some time during the interval 3 to 4.

20. The positions of two blocks at successive 0.20-second time intervals are represented by the numbered squares in the figure below. The blocks are moving toward the right.

The accelerations of the blocks are related as follows:
(A) The acceleration of "a" is greater than the acceleration of "b".
(B) The acceleration of "a" equals the acceleration of "b". Both accelerations are greater than zero.
(C) The acceleration of "b" is greater than the acceleration of "a".
(D) The acceleration of "a" equals the acceleration of "b". Both accelerations are zero.
(E) Not enough information is given to answer the question.
USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT FOUR QUESTIONS (21 through 24).

A rocket drifts sideways in outer space from point "a" to point "b" as shown below. The rocket is subject to no outside forces. Starting at position "b", the rocket's engine is turned on and produces a constant thrust (force on the rocket) at right angles to the line "ab". The constant thrust is maintained until the rocket reaches a point "c" in space.

21. Which of the paths below best represents the path of the rocket between points "b" and "c"?

22. As the rocket moves from position "b" to position "c" its speed is:
   (A) constant.
   (B) continuously increasing.
   (C) continuously decreasing.
   (D) increasing for a while and constant thereafter.
   (E) constant for a while and decreasing thereafter.

23. At point "c", the rocket's engine is turned off and the thrust immediately drops to zero. Which of the paths below will the rocket follow beyond point "c"?

24. Beyond position "c" the speed of the rocket is:
   (A) constant.
   (B) continuously increasing.
   (C) continuously decreasing.
   (D) increasing for a while and constant thereafter.
   (E) constant for a while and decreasing thereafter.
25. A woman exerts a constant horizontal force on a large box. As a result, the box moves across a horizontal floor at a constant speed \( v_0 \).

The constant horizontal force applied by the woman:
(A) has the same magnitude as the weight of the box.
(B) is greater than the weight of the box.
(C) has the same magnitude as the total force which resists the motion of the box.
(D) is greater than the total force which resists the motion of the box.
(E) is greater than either the weight of the box or the total force which resists its motion.

26. If the woman in the previous question doubles the constant horizontal force that she exerts on the box to push it on the same horizontal floor, the box then moves:
(A) with a constant speed that is double the speed \( v_0 \) in the previous question.
(B) with a constant speed that is greater than the speed \( v_0 \) in the previous question, but not necessarily twice as great.
(C) for a while with a speed that is constant and greater than the speed \( v_0 \) in the previous question, then with a speed that increases thereafter.
(D) for a while with an increasing speed, then with a constant speed thereafter.
(E) with a continuously increasing speed.

27. If the woman in question 25 suddenly stops applying a horizontal force to the block, then the block will:
(A) immediately come to a stop.
(B) continue moving at a constant speed for a while and then slow to a stop.
(C) immediately start slowing to a stop.
(D) continue at a constant speed.
(E) increase its speed for a while and then start slowing to a stop.
28. In the figure at right, student "a" has a mass of 95 kg and student "b" has a mass of 77 kg. They sit in identical office chairs facing each other.

Student "a" places his bare feet on the knees of student "b", as shown. Student "a" then suddenly pushes outward with his feet, causing both chairs to move.

During the push and while the students are still touching one another:

(A) neither student exerts a force on the other.
(B) student "a" exerts a force on student "b", but "b" does not exert any force on "a".
(C) each student exerts a force on the other, but "b" exerts the larger force.
(D) each student exerts a force on the other, but "a" exerts the larger force.
(E) each student exerts the same amount of force on the other.

29. An empty office chair is at rest on a floor. Consider the following forces:
1. A downward force of gravity.
2. An upward force exerted by the floor.
3. A net downward force exerted by the air.

Which of the forces is (are) acting on the office chair?
(A) 1 only.
(B) 1 and 2.
(C) 2 and 3.
(D) 1, 2, and 3.
(E) none of the forces. (Since the chair is at rest there are no forces acting upon it.)

30. Despite a very strong wind, a tennis player manages to hit a tennis ball with her racquet so that the ball passes over the net and lands in her opponent's court.

Consider the following forces:
1. A downward force of gravity.
2. A force by the "hit".
3. A force exerted by the air.

Which of the above forces is (are) acting on the tennis ball after it has left contact with the racquet and before it touches the ground?
(A) 1 only.
(B) 1 and 2.
(C) 1 and 3.
(D) 2 and 3.
(E) 1, 2, and 3.
APPENDIX C

RESEARCHERS REPORTING INTRODUCTORY PHYSICS CURRICULAR REFORMS AS RESPONSES INFORMED BY FCI & HAKE GAIN FINDINGS

<table>
<thead>
<tr>
<th>Institution Implementing Reform</th>
<th>Researcher(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>California Polytechnic State University</td>
<td>Hoellwarth, Moelter, &amp; Knight (2005)</td>
</tr>
<tr>
<td>Harvard University</td>
<td>Crouch &amp; Mazur, (2001)</td>
</tr>
<tr>
<td>Massachusetts Institute of Technology</td>
<td>Dori &amp; Belcher, (2004)</td>
</tr>
<tr>
<td>Rutgers, State University of New Jersey</td>
<td>Etkina &amp; Van Heuvelen, (2001)</td>
</tr>
<tr>
<td>Tufts University</td>
<td>Thornton (1999)</td>
</tr>
<tr>
<td>University of Colorado</td>
<td>Pollock (2004)</td>
</tr>
</tbody>
</table>
2010 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 Fifteenth Annual Colloquium on Research in Mathematics and Science Education and published in the Colloquium Journal, vol. XV. 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 as separate documents.

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. Papers must be received by December 1, 2009.

6. Authors will be notified about the status of their papers by January 15, 2010.

7. The Colloquium will be scheduled for April 2010.

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