Volume 11 | Number 1 | Fall 2006

Electronic Journal of Science Education

ISSN 1087-3430

Published by Southwestern University

Electronic Journal of Science Education

Electronic Journal of Science Education

Volume 11, Issue 1, Fall 2006

CONTENTS

Editorial:
Continuing the Vision

Michael Kamen
Southwestern University

This issue of the Electronic Journal of Science Education is the first under new editorial leadership. John Cannon and David Crowther demonstrated foresight and extraordinary commitment to science education by starting and publishing EJSE for ten years. As the founding editors of EJSE they were certainly a notable force in pushing the acceptance of electronic publications as credible scholarship by the academy. Charles Eick, Julie Luft, and Molly Weinburgh deserve recognition for their close reading of manuscripts and thoughtful recommendations as associate editors. In addition EJSE would not be possible without the hard work of the editorial review board, the Southwestern University Staff (including Connie Imhof, Ansa Copeland, and Laura Marquez), the advisory board, and especially all the authors who submit manuscripts. It is an honor to have the opportunity to serve as editor of the Electronic Journal of Science Education.

My vision for EJSE is to support the dissemination of well-written and substantive research, theoretical, and innovative perspective articles. We welcome science education manuscripts that report meaningful research, present research methodology, develop theory, and explore new perspectives. EJSE is an open access journal with a vigorous peer-review process and high standards for publication. While much of academia is calling for *rigor* and *rigorous* review, I prefer the adjective *vigorous*. Dictionary definitions of *rigorous* typically include words such as *rigid*, *harsh*, *inflexible*, and *severe*. EJSE strives to be open, insightful, energetic, and helpful while maintaining high scholarly standards. Within the limits of our resources, we are committed to as fast a review process as possible while providing constructive feedback to authors.

Manuscripts are first screened by the editor, and feedback is provided to the author(s) or the blind copy is e-mailed to the editorial review board. Reviewers agree to review articles that they are most qualified to evaluate. When three reviews are completed, an associate editor reads the article and the reviews and makes a recommendation to the editor. The editor and associate editor agree on a decision and the author(s) is notified. The turn around time is getting quicker, and the reviews are thoughtful and thorough. We believe we are selecting high-quality articles while being supportive and providing helpful feedback to developing authors.

We are at a time when we need to become more creative, timely, and responsive in science education scholarship. While there is certainly a need to build on established research agendas, there is an equal or greater need to push the envelope. It is disheartening to visit schools, read state mandates, and attend policy meetings. We need scholarship that will challenge what is assumed and advocate for those who are left out. While EJSE will continue to publish a variety of kinds of articles, a priority is to continue to be an innovative force in science education. Scholars in science education are invited to use EJSE as one of their venues to develop, explore, and evaluate innovative thought and connections that teaching science in a diverse and unequal world requires. It is our goal to continually push EJSE to be diverse, vigorous, and innovative. We believe that these three elements are required to serve the needs of both science education scholars and practitioners.

We are delighted with the first set of articles we have the privilege of publishing in this issue. The seven articles from five countries have diverse foci ranging from science education research methodology to connections between art and science. Looking at methodology Scharfenberg, Bogner, and Klautke provide recommendations for designing experimental research in educational settings. The latter, an innovative perspective article by Ashkenazi, explores metaphorical and cognitive similarities between art and science and postulates that this new perspective may provoke different approaches in science education inviting the interest of more students.

The international scope of this issue is demonstrated by Dal's study, which investigates and compares French students' and student teachers' understanding of volcanism. Meichtry and Smith study a professional development program preparing teachers to implement place-based education in the context of aquatic ecology to explore its impact on teachers' confidence and classroom practice. Another manuscript authored by Garrison and Amaral documents the development of an instrument to evaluate the impact of professional development on classroom practice.

Waldrip, Prain, and Carolan investigate the demands and benefits of multi-modal representation. This publication includes two PowerPoint animations created by the subjects of the case studies. This is an example of how electronic journals can exploit their own modality to enhance the sharing of research findings. And, finally, Mitchell and Hoff address an important equity issue by examining the role of assessment in contributing to the gender gap in science.

These articles represent diversity in country of origin, research paradigm, and content. It is our intention that, through articles of this caliber, EJSE will enhance the quality of research, add to our understanding of how to facilitate science teaching and learning, and provide data needed by schools and policy makers.

Metaphors in Science and Art: Enhancing Human Awareness and Perception

Guy Ashkenazi
The Hebrew University

Abstract

Science and art are commonly considered as two separate cultures, which differ in both tongue and value. However, while the material artifacts produced by science and art are markedly different, the creative cognitive process of their construction is closely related – both cultures use a metaphorical language, which sharpens perception of details and enhances awareness of structure. The strong societal association of science with technology and of art with aesthetics masks the cognitive similarities. By re-emphasizing these similarities, we hope to gain access to a student population who was previously alienated by the utilitarian, impersonal presentation of science.

Correspondence should be addressed to Guy Ashkenazi, The Hebrew University, Jerusalem 91904, Israel Email: guy@fh.huji.ac.il

Introduction

In the eyes of the public, science and art are considered as two separate cultures, which differ in both tongue and value (Snow, 1959). Scientific language is considered objective and factual, where art's language is subjective and metaphorical. Science is valued based on its utility and its future prospect of technological development. Art, on the other hand, is appreciated for its aesthetic value and influence on human emotion. Science benefits society, art enriches the individual.

However, that has not always been the case. For the great masters of the Renaissance, science and art were inseparable in their pursuit of expanding the bounds of human knowledge and experience – the realistic depiction of physical space and the human body were both informed by and contributing to the development of physics, anatomy and mathematics (Dauben, 1991). This paper will show that even today, the gap is not as insuperable as it may seem. While the material artifacts produced by science and art are markedly different, the creative cognitive process of their construction is closely related. Both cultures share a common metaphorical tongue, which deeply influences human perception and awareness. Both use images and metaphors which reveal the intangible fabric of tangible existence. Such metaphors allow us to see and feel things that are otherwise passed by unseen and unfelt, and thus enrich our experience of the natural world. Through these metaphors, we ultimately become aware of meaning and structure in the intricate complexity of the surrounding world.

Art, perception and awareness

Our five senses flood the brain with a constant stream of input, too rich and rapid to be processed in its entirety. To reduce this perceptual overload, the mind filters out most of the sensory input, and focuses attention on only a small fragment (Broadbent, 1958). Past experience plays a crucial role in this filtering process in two different ways. First, it sharpens perception by offering contextual cueing (Chun & Nakayama, 2000) – attention is primarily directed to patterns of details that proved to be significant in the past (when crossing a street, we are more likely to notice motion on the road and the sound of a car engine than movement on the sidewalk and the chirping of birds). Second, it facilitates awareness by "chunking" information (Miller, 1956) – grouping the diverse details of perception into familiar units that carry additional meanings, and then processing relations between generalized meaningful units rather than between specific details (the moving collection of metal, glass and rubber is identified collectively as a car, which means a vehicle with the purpose of transportation but also with the capacity to kill careless street-crossers). In this filtration and meaning construction process, past experience can take the shape of prior beliefs, expectations, conceptions, language and culture, all of which exert a strong influence on how we perceive the world (Swoyer, 2003). The more experience we have, the more details we can perceive, and the more comprehensive our awareness of the situation.

Art is a source of experience that is explicitly designed to shape our perception and awareness of the world. In many cases, art portrays nature and human condition from the viewpoint of the artist, which allows the observer to "step out" of his limited, egocentric point of view, and gain a broader perspective on the subject, making him aware of facets he hadn't considered and features he might have missed. Impressionist paintings, like Van Gogh's *Starry Night*, emphasize perception. The visible brushstrokes and bright colors accentuate lighting and overall composition while suppressing precise details, striving to recreate the painter's sensation of viewing the subject, rather than recreating the subject itself.

The Starry Night, 1889, by van Gogh

Cubist works of art, on the other hand, emphasize awareness. The distorted collage of multiple view angles presented simultaneously (like Picasso's *Guernica*), reflect the artist's inner conceptualization, rather than his direct perception, of the multi-faceted nature of the subject (Ortega y Gasset, 1913).

Guernica, 1937, by Picasso, copyright Succession Picasso 2006

In the literary arts, the author's personal experience is often communicated by the use of metaphors, which are employed to emphasize similarities between current and past experiences. Metaphors can be straightforward, like in this excerpt from Herman Melville's *Moby Dick:*

> *It was while gliding through these latter waters that one serene and moonlight night, when all the waves rolled by like scrolls of silver…*

In this perceptual metaphor, the author interprets new sensory input (moonlit waves) by casting it into the structure of past experience (the shape of scrolls and the glistening of silver). This provides contextual cueing for other people that read these words, who acquire a new and more affluent way to appreciate the natural world – a person taking his time to look for "scrolls of silver" in a moonlit sea values the moment more than someone who just notices a mundane scene of the ocean at night. Metaphors can also take a more complex form, in which the entire storyline stands as a metaphor for an actual human condition or event (allegory). This is done in order to simplify a multifaceted situation and raise awareness to the interactions of a few key concepts, which would otherwise be hidden in the complexity of details in the actual situation. For example, George Orwell's *Animal Farm* is an allegory of the socialist Russian revolution and its moral deterioration into the communist dictatorship under Stalin. In this book, the political and social processes characterizing the entire Soviet nation are described by reference to animals in a small farm in England. Each animal is attributed with just a single representative quality like idealism, greed, viciousness, or gullibility. The struggle between justice and equality vs. greed and power ultimately leads to the corruption of a potentially utopian society. Following the simple structure of interactions between a small set of meaningfully defined characters develops awareness and understanding of the more complicated real world situation.

Perceptual metaphors are also employed in wine tasting, where tangible visual and corporal metaphors are associated with the more elusive sensations of flavor and aroma. Where an ordinary person would only perceive a glass of red wine with a pleasant taste, a wine taster would notice that the wine has "good legs" by swirling it in the glass; discern a "full body" by sloshing the wine in his mouth; contemplate to determine its "round character"; and finally detect a "smooth aftertaste" that lingers in the mouth after swallowing. Undoubtedly, the art of wine tasting has provided the wine taster with a store of metaphors that allows him to savor a much richer experience than the uninitiated person.

Science, perception and awareness

Metaphors play an important role in science, as well. In science, metaphor is a tool of exploration and discovery, providing a way of imposing or discovering structure within novel or unfamiliar situations by relating them to familiar experiences. Metaphors such as "electricity is a fluid" or "atoms are hard spheres" are contextual cues that direct the scientist's attention to look for details associated with fluids or hard spheres. Fluids can be associated with flow and conservation; hard spheres with packing and random motion. Even if these metaphors are ultimately replaced by more elaborate mathematical models, they still guide the thoughts of practicing scientists when they try to make sense of a new experience (Brown, 2003).

However, the role of scientific metaphors does not end in scientific practice. With time and training, scientific metaphors extend beyond the boundaries of professional scientific activities and pervade the scientist's daily life as well. Scientific metaphors emphasize similarities between seemingly different situations encountered in the natural world, and therefore apply naturally to commonplace encounters with the world. They provide contextual cues that enhance perception, and increase sensitivity to salient features that would otherwise elude the senses. For example, synthetic organic chemists often use "boiling stones" (small beads of an inert, porous material) to facilitate gentle and efficient formation of bubbles and avoid sudden eruption of liquid and gas when boiling their esoteric brews. Just like the shape of scrolls and the glistening of silver, the convoluted surface of boiling stones can serve as a powerful metaphor. A chemist who drinks from a glass of soda would first look at the growing bubbles on the side of the glass; attracted by the resemblance of the process to the formation of bubbles on boiling stones, he would look for similar surfaces on the glass; seeing no visible sign for pours and bumps, he would imagine small specks of dust or dirt which harbor in their convoluted microscopic structure tiny bubbles, into which gas escapes from the over-saturated solution, inflating them like soap bubbles with a straw; he would focus his vision on a growing bubble, and confirm his conjecture by observing that once the bubble floats to the top, another bubble starts forming at the exact same spot; he would notice that some spots produce bubbles slowly, and others more rapidly; and he would be fascinated by the way some bubbles form in pairs, others in continuous streams, and yet others in short, rapid bursts. Undoubtedly, the chemist enjoys a much richer experience than just quenching his thirst[1]. For him, being a practicing scientist had turned the

---

[1] This experience can be further enhanced by dropping a grain of cooked rice into the soda, and watching how the bubbles formed on its surface send it floating ever upwards, like Icarus soaring towards the sun, and then bursting at the surface and letting it fall back down, the mythical tragedy repeating over and over again…

unattended act of drinking soda into a significant context of perception for a wealth of otherwise unperceivable details[2].

Science also uses extended allegories – models that simplify a multifaceted situation and raise awareness to the interactions of a few key concepts, which would otherwise be hidden in the complexity of details in the actual situation. Like literary allegories, scientific models reduce the number of accounted players and strip them of all but their most representative characteristics. For example, in the *information processing model* of human cognition the complex neural structure of the brain is replaced by a small number of computer-like hardware components; in the *ideal gas model* molecules are stripped of all structure-dependent attractions and repulsions save for hard-spheres-like elastic collisions. Following the simple structure of interactions between a small set of meaningfully defined characters develops awareness and understanding of the more complicated real world situation. Once again, this enhanced awareness is not limited to professional activities, but manifests itself in the apparently ordinary – where the scientific illiterate sees plain rock formations, the geologist catches a glimpse of the passing of eons; instead of merely looking at a "starry night", the astronomer witnesses an ever expanding universe; and the annoying insect infested shrubbery transforms into an intricate eco-system in the eyes of a biologist. Familiarity with scientific models transforms unadorned and frequently ignored experiences into awe inspiring appreciation of natural existence in the mind of a trained scientist.

Science, art and society

We have seen that science has a capacity, similar to that of art, to change the way we view the world, sharpen our perception of details and enhance our awareness of structure. Unfortunately, science often fails to "market" itself as sharing these values of art. In the last five decades, science has been mainly portrayed as a practical endeavor, a necessary tool for technological progress. The first science education revolution, which started in the 1960s in the US, was aimed at the mass production of professional scientists that will serve as a technical work force necessary for economical development. This revolution was unsuccessful because it failed to attract the interest of students, who turned away from science. In an attempt to correct this, the second science education revolution of the 1990s aimed at producing "scientifically literate citizens", who use science as a basis for making rational decisions in a technological society (DeBoer, 2000). Even though the two approaches seem very different, both share the same underlying premise – science is important because it supports society's technological needs, and students should learn science because of its utility, whether in promoting a future career or in making personal decisions.

Art, on the other hand, never had to answer such strict demands for utility. We don't show our students paintings of van Gogh in art classes because we hope they will grow up to be tormented artists, nor do we have them read *Romeo and Juliet* because we expect this will help them make educated decisions in their romantic lives. In our culture, art needs no practical justification. We actually value art *because* it has no practical value. We value art because it enriches the "soul" rather than the "body", or in other

[2] For more information about effervescence in soft drinks, see Liger-Belair, Tufaile, Robillard, Jeandet and Sartorelli (2005).

words, because it serves no basic survival needs. However, the appeal to the soul has emphasized the aesthetic and emotional qualities of art, at the expense of its cognitive capacity to extend perception and awareness beyond our individual experience. Art is more than just “pleasant” or “emotionally arousing” – it can be intellectually stimulating and inspire creativity through the act of employing metaphors to explore different worldviews and to construct representations for these newly acquired perspectives.

Science and art share a common cognitive core, but it is the difference in their societal value that dominates the public’s opinion. The clash between technology and emotion, between utility and aesthetics, is the reason for the widening gap between the two cultures. The dominance of societal values is so strong, that even attempts to reconcile the difference often focus on integrating the dissimilarities, rather than uncovering the similarities – discussions on the technological aspects of creating, preserving and exhibiting art, or the aesthetic aspects of presenting science, only serve to accentuate the different instead of revealing the parallel.

This paper exposes the common cognitive core of art and science, in an attempt to promote a culturally rich discourse based on a comparable semantic structure (Galili & Zinn, 2006). We must find a way to show that technological progress and personal fulfillment are not contradictory, but rather, complimentary; that both science and art are intellectually appealing and rewarding in much the same way. Maybe a different approach in science teaching, one that emphasizes the contribution of science to the enrichment of the individual, will help attract the interest of students that have been previously turned off by the utilitarian, problem-solving based presentation of science (Tobias, 1991). Maybe such an approach, which resonates science’s creative and metaphorical voice, will help alleviate some of the antagonism many people have (and sometimes even take pride in) towards the technical and impersonal language of science. And maybe, just maybe, this will re-open a path of communication, one of common tongue and value, between the two cultures, a path that will lead to a more integrated understanding of man and nature[3].

[3] For a detailed history of the interrelations between art and science in man’s investigation of the nature of reality, see Shlain (1991).

References

Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.

Brown, T. L. (2003). *Making Truth: Metaphor in Science*. Champaign, IL: University of Illinois Press.

Chun, M. M., & Nakayama, K. (2000). On the functional role of implicit visual memory for the adaptive deployment of attention across scenes, *Visual Cognition*, *7*, 65-81.

Dauben, J. W. (1991). *The Art of Renaissance Science: Galileo and Perspective* [Videotape]. New York: Science Television Company. Also retrieved October 3, 2006 from http://www.mcm.edu/academic/galileo/ars/arshtml/arstoc.html

DeBoer, G. E. (2000). Scientific Literacy: Another Look at Its Historical and Contemporary Meanings and Its Relationship to Science Education Reform. *Journal of Research in Science Teaching*, *37*, 582-601.

Galili, I., & Zinn, B. (2006). Physics and Art: A Cultural Symbiosis in Physics Education, *Science and Education* (in press).

Liger-Belair, G., Tufaile, A., Robillard, B., Jeandet, P., & Sartorelli, J.C. (2005). Period-adding route in sparkling bubbles. *Physical Review E, 72*, 037204.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

Ortega y Gasset, J. (1913). On Point of View in the Arts. Translated in *The Dehumanization of Art and Other Essays on Art, Culture and Literature*. Princeton, NJ: Princeton University Press, 1968.

Shlain, L. (1991). *Art and Physics: Parallel Visions in Space, Time, and Light*. New York: William Morrow & Co.

Snow C. P. (1959). *The two cultures and the scientific revolution*. New York: Cambridge University Press.

Swoyer, C. (2003). Relativism and the Constructive Aspects of Perception, Supplement to Relativism. In N. Zalta (Ed.). *The Stanford Encyclopedia of Philosophy* (Spring 2003 ed.). Retrieved October 3, 2006 from http://plato.stanford.edu/archives/spr2003/entries/relativism/supplement1.html

Tobias, S. (1991). *They're Not Dumb, They're Different: Stalking the Second Tier*. Washington, DC: Science News Books.

(Dis)Interest in Science: How Perceptions About Grades May Be Discouraging Girls

Sidney N. Mitchell
University of Maine

Dianne L. Hoff
University of Maine

Abstract

In spite of advances in many fields, women are still under-represented in the sciences. In this paper, we report the results of a study investigating the perceptions of high school girls enrolled in science classes on whether hard work leads to success, if they are receiving the scores they deserve, and if the assessment system used in class is unfair. Analyses indicated that girls received better grades than boys, but generally believed that hard work does not lead to success and that the grading system is not completely fair. The findings suggest subtle ways that classrooms may be discouraging girls, and recommendations for teaching practices in science education to address this problem are provided.

Correspondence should be addressed to Sidney N. Mitchell, PhD, College of Education and Human Development, 5749 Merrill Hall, Room 223, University of Maine, Orono ME 04469-5749 Telephone (207) 581-3435, Fax (207) 581-3120. Email: sid.mitchell@maine.edu

Introduction

The national spotlight again shines on the dwindling number of students entering the sciences as a profession, with international reports highlighting our lost standing in world rankings (OECD, 2003) and the popular press reporting the poor achievement of students in science classes (Lemonick, 2006; Science Scores, 2006 author unknown). Further, studies indicate that the large gender differences between the number of women and men entering the sciences have not receded. Girls continue to lose interest and abandon plans for careers in the field at a higher rate than boys (Kerr & Robinson Kurpius, 2004), and college females continue to opt for majors in science and mathematics at a lower rate than college males (National Center for Education Statistics, 2000). Women are still less likely to be employed in sciences than men, and for those women who are employed in the sciences, they earn, on average, 20% less than their male counterparts employed in the same job (Graham & Smith, 2005). If our nation needs more scientists, this under-representation of women must be examined and remediated, and it starts with girls in school.

Why are girls less likely to pursue careers in the sciences? The enduring under-representation of women in science has plagued educators for many years, despite concerted efforts to raise girls' academic success and interest in the field, such as mentoring programs (McLaughlin, 2005), and single sex classes (Ransome,1993). There is clear evidence that girls have virtually closed the previous achievement gap in science. In all fields, except physics, the extant literature shows that girls achieve academically at the same or higher levels than their male counterparts (Kleinfeld, 1999; National Center for Education Statistics, 1997; National Science Foundation, 1996). The answer to why fewer girls pursue science, therefore, does not seem to lie in underachievement.

Nor does the answer appear to lie in lack of access to science instruction. Title IX of the 1972 Education Amendments prohibits gender discrimination in all school programs, and since the early eighties, schools across the country have taken pro-active steps to overcome the representation gap in math and science with programs designed to encourage girls' enrollment in these disciplines (Gilbert, 2001; Sandler, Silverberg, & Hall 1996). Providing programs for girls, however, is just a small part of addressing the problem of under-representation of women in the sciences. Programs often do not address the beliefs that students, teachers, parents, or administrators bring to the learning environment – they just provide access opportunities for girls, with the belief that the access alone will produce the desire to pursue sciences. Although getting girls into science classes is important, many school programs fail to recognize or address some of the underlying assumptions about the place of women in the sciences (Gilbert, 2001).

Discriminatory treatment of females within science classrooms also does not explain the under-representation of women in sciences. For the most part, overt acts of sexism in the classroom have disappeared from our public schools and college campuses (Allan & Madden, 2003). However, even without blatant discrimination, girls may be treated differently than boys by science teachers (mostly males), who may interact differently with female students than they do with male students or who may not recognize that their teaching methods are not effectively reaching girls (Sandler, Silverberg, & Hall, 1996). Sandler, et al. (1996) noted that one way girls are treated differently is science classes is in the types of questions they are asked. For example, girls are often asked lower level factual questions which can be evaluated right or wrong and only permits specific instruction from the teacher. Whereas, boys are asked open-ended higher level questions which allows them to "display their talents," engage in critical thinking, and even guess at the answer (p. 10). Several other examples of how girls are treated differently than boys include: grouping women in ways that indicate they have less ability or status; making seemingly helpful comments which imply girls are not as competent as boys; doubting girls work and accomplishments; expecting less of girls in the future; or calling males "men" and women "girls" (for a complete review see Sandler et al., 1996 pp. 10-11).

The question of under-representation of women and girls in science, therefore, must include an examination of the more subtle aspects of classroom climate. If schools just "check off" that they have provided science opportunities, that girls are achieving satisfactorily, and that there is no overt discriminatory treatment toward girls, it may obfuscate classroom practices that are contributing to a chilly climate for girls and young women in science. Chilly climates are defined as the "... subtle ways women are treated differently - ways that communicate to women that they are not quite first-class citizens in the academic community" (Sandler & Hall, 1986, p. 1). Chilly climates can be created by overt behaviors directed at girls to make them feel unwanted, such as not calling on them when they raise their hands or praising males but not females (Morris, 2003). On the other hand, they can be more covert, such as deliberately grading girls' work more severely than boy's work. Or chilly climates can be created unintentionally, stemming from lack of knowledge or insensitivity to the different learning styles and needs of girls and boys (Salter, 2003).

Further, classroom climate is not defined solely by behaviors exhibited by others, but also by the perceptions girls hold about the learning environment. Believing that these perceptions can influence girls' motivation and participation in the class, we attempt to explore in this study a few of these subtleties by examining their perceptions on specific classroom practices. Although the literature indicates that girls are performing at equal or higher levels than boys, what is often not reported are more subtle factors that affect girls' efficacy in science, such as their perception of classroom grading practices. Specifically, this study is an effort to describe the perceptions high school girls have about the grading practices in their science classes.

Research Questions

The questions that guided this investigation are:

Question 1: Do students perceive that their hard work will lead to success in class?

Question 2: Do students perceive that they are accurately assessed?

Question 3: Do students perceive that the assessment system in class is unfair?

Methods

Participants

The participants for this study were 121 grade 10 and 11 high school students from a medium sized high school located outside Montreal, Canada. This included 46 females and 75 males, ranging in age from 14 to 18, with an average age of 15.72 years. The age range is due to the October data collection, which meant students with later birthdays were still 14 in 10th grade, and students with early birthdays, or those who started a year later, had turned 18 in the 11th grade. Students represent varied achievement levels, SES, cultures, and ethnicities. The breakdown of gender by grade indicates that 33 girls were enrolled in grade 10 and 13 in grade 11. The breakdown for boys indicates that 47 were enrolled in grade 10 and 28 in grade 11. Permission from the teacher, school, school board, and parents was obtained before administering the surveys. Additionally, we obtained informed consent from each student, and they were told that they could stop at any time or have any of their responses eliminated from the analysis.

All of the students had the same teacher during the school year being studied, although they entered the class with a wide variety of experiences with multiple teachers in previous years. The teacher is a male, award-winning science teacher with 17 years of teaching science experience and a Master's degree in his field. He received two district awards for outstanding teaching and his dedication to his profession. We limited the study to the students of one teacher, to provide greater power in interpreting the results, since by doing so, we did not have to disentangle teacher effects from the student responses. The teacher's instructional approach included a mixture of direct instruction, cooperative groupings (Abrami, et al., 1995), and lab work. At the start of a unit the teacher would begin with direct instruction and then move to cooperatively structured groups for lab work. Within the groups, students took on various roles (e.g., experimenter, recorder, materials coordinator), which they alternated after each experiment.

Measures

The Classroom Life Scale (CLS; Johnson, Johnson, & Anderson, 1983) was used to measure attitudes toward grading. The CLS is a 5-point scale ranging from 1=Completely False to 5=Completely True. The CLS consists of several subscales measuring students' attitudes on the fairness of grading, grading practices, cooperativeness, feelings of alienation, academic self-esteem, academic support, goal and resource interdependence, external motivation, cohesion, independent learning, competitive learning, controversy, valuing homogeneity and heterogeneity. This paper is focused on the grading practices in terms of fairness of grading and the beliefs that one is getting the scores they deserve. On the CLS, students were encouraged to provide written feedback on any aspect of their learning in science courses. Comments were analyzed for any patterns or themes that related to the study's questions. Achievement data was also examined, consisting of two mid-term exams held in mid-October and mid-May and weekly lab assignments.

Procedure

In early October, researchers administered the CLS to all students in their science classroom. We told students that we were interested in learning about their attitudes toward science and how these attitudes affected their learning and grades. Students were informed that we wanted their honest responses to the questions and that there was no right or wrong answers. The teacher was not present during the survey administration and was not shown the completed questionnaires; he only saw aggregated data after the study was completed. The study lasted for seven months (October to May). In late May, two weeks after their midterm, the CLS was administered a second time. We decided to wait two weeks before administering the CLS so that any positive or negative emotions surrounding their mid-term results would have lessened, and thus would not overly influence their responses.

Throughout the year, students completed weekly assignments and were graded on all assignments. By the time the students completed the first mid-term exam they had completed five weekly assignments; by May they had completed more than 20. The exams and the weekly assignments were evaluated by the teacher and factored into their overall grade.

Results and Discussion

The results of the study are presented by research question. However, before examining the results of the survey, we analyzed students' grades on weekly lab assignments, plus two mid-term exams to highlight any gender differences in achievement. On average, girls carried a 'B' average and boys carried a 'B-'average in the class. The averages were submitted to a one-way analysis of variance (ANOVA) and the test indicated that the differences were not significant $F(1,120) = .913$, $p>.34$. The perceptions of the students should be interpreted in light of the overall achievement levels in the class.

Research Question 1: Do students perceive that their hard work will lead to success in class?

Statistical evaluation of pretest scores measuring differences between girls and boys on their perceptions that hard work would lead to success in the course was performed using the Mann-Whitney U-Test. This nonparametric test was used because an ANOVA test of means did not meet the test of homogeneity of variance assumption necessary for using the parametric one way ANOVA. The results of the analysis showed that girls and boys likely entered the course with different perceptions about the value of hard work leading to success ($U = 1032$, $p <.001$, two-tailed). Was this gender difference in perception maintained throughout the course? To test this possibility we submitted posttest scores to a univariate ANOVA procedure. Tests of posttest score means showed that the assumption of homogeneity of variance underlying ANOVA was met and the results of this test indicated that girls and boys differed significantly on their perceptions that hard work would lead to success in the class $F(1,119) = 5.96$, $p <.01$, partial eta squared = .048; girls $M = 3.83$, $SD = 1.14$; boys $M = 4.29$, $SD = .94$.

To find out if differences in perceptions of hard work leading to success shifted from pretest to posttest within gender we conducted a correlated samples t-test separately for girls and boys. The results showed that girls exhibited a slight positive shift in overall mean scores in their perceptions from pretest to posttest (pretest $M = 3.52$, $SD = 1.16$; posttest $M = 3.83$, $SD = 1.14$); however, the change was not significant, $t(45) = -1.28$, $p>.20$. Likewise, for boys the results indicated that although there was slight downward shift in overall mean scores from pretest to posttest (pretest $M = 4.33$, $SD = .87$; posttest $M = 4.29$, $SD = .94$) the change was not significant $t(74) = .359$, $p>.70$. Although scores for both girls and boys were on the positive side of the scale, boys continued to agree with the statement that hard work would lead to success in class. Whereas, girls were closer to the midpoint of the scale which reflected the belief “sometimes true/sometimes false” indicating that girls were much more ambivalent about their beliefs in the value of hard work leading to success in these classes.

Discussion Question 1

It is clear from the data that girls entered the class with less confidence that their hard work would lead to success in science. We looked at the percentage of scores for girls and boys who chose “agree or strongly agree” and found that 50% of the entering girls agreed that hard work pays off, whereas 80% of the entering boys did. Even after the girls were in the class for almost a year with an award-winning teacher, their belief in the benefit of hard work in science did not change or approach the belief held by boys. Previous research and the qualitative data from the study provide insight into why girls may hold this belief. The pedagogy of independent and competitive learning inherent in science classes – lab experiments, forming hypotheses – has ramifications for girls that are linked to gender socialization and help explain why girls feel less sure about hard work leading to success (Ball, 2002). For example, Fennema and Peterson (1985) found that classes that require autonomous learning behaviors favor boys because they have many more opportunities both inside and outside of school to practice these skills. Boys are often given more freedoms than girls, participate in more competitive play and activities, and are more socialized for independence. In a 1996 study, Silverman and Pritchard observed that when students were working independently and must wait for the teacher's time and attention, boys benefit because they are more outspoken in demanding attention. They observed that teachers often ignored interactions between boys and girls that are negatively affecting girls' perceptions of the class. For instance, they noticed that boys rushed to get the supplies they needed, overrunning the girls in the process.

Girls, therefore, may be receiving reinforcement of the stereotype that "boys are better in science" by observing and being victimized by boys' more aggressive approach to learning, and the teacher's tacit acceptance – even approval – of this behavior. This suggests, and other research supports it, that girls would perform better in science and math classes if the classes were taught in a cooperative or individualized format rather than a competitive format (Eccles et al., 1998; Meece, 2002; Meece & Eccles, 1993).

The comments from girls also support the premise that girls may not believe they are thriving in independent learning environments. For example, the perception of many of the girls in the class was that the teacher left them alone "to struggle." As one 15 year old girl in grade 10 stated, *"The teacher does not teach! He leaves us to our own devices. I suspect he wants us to fail."* Another 15 year old girl in grade 10 shared a similar sentiment, but also referred to the teacher's willingness to answer questions *"Our teacher doesn't teach us enough. When we have a question he won't help us."* And, a 17 year old girl in grade 11 stated *"The teacher should do some teaching. Not just answer your questions with another question. Also he should help you when you don't understand."* Finally, one 16 year old girl in grade 11 stated *"The teacher shouldn't expect us to learn by ourselves because there are many times when I feel I cannot approach the teacher with a question because he will just ask me what I think, when I* <u>*don't know*</u> *[emphasis in the original] to start off with. It makes me feel dumb*." It is noteworthy that none of the boys shared similar comments. Girls, therefore, seem more likely to interpret autonomous learning time as a teacher ignoring them or not caring if they fail. Their statements are also reflective of the socio-emotional warmth available to students in the classroom. Research in the area of socio-emotional warmth indicates that teachers who respect and care for students provide environments that facilitate student engagement, persistence on academic tasks, and the development of positive achievement-related perceptions (Goodenow, 1993; Midgley, 2002; Midgley, et al., 1989). The comments from girls suggest that girls do not feel accepted or confident enough to approach the teacher when they are experiencing difficulty, indicating a low level of socio-emotional warmth, which some would also call a "chilly climate." It should be noted that these are the girls' perceptions; researchers did not note any behavior that would suggest gender bias on the part of the teacher.. However, the fact the girls reported it is important, because people act on what they believe to be true (Bandura, 1997).

Research Question 2: Do students perceive that they are accurately assessed?

To answer this question, we conducted a univariate ANOVA on pre-and-posttest scores. On pretest, the test revealed a significant main effect for gender on perceptions of whether they are receiving the scores they deserve, $F(1,119) = 33.06$, $p<.001$, partial eta squared .217. The data showed that girls perceived that they were not receiving the scores they deserved in the class. Whereas, boys perceived that, for the most part, they were receiving the grades they deserved (girls M = 2.63, SD = .951; boys M = 3.72, SD = 1.04). On posttest, girls and boys did not differ significantly ($p>.99$) and they exhibited the same mean scores (girls M = 3.11, SD = 1.14; boys M = 3.11, SD = 1.24). Within subjects tests reflected a positive shift in perceptions for girls (girls pretest M = 2.63, SD = .951; posttest M = 3.11, SD = 1.14, $p<.02$). However, the scores were still very close to the midpoint of the scale reflecting "sometime true/sometimes false." Boys, on the other hand indicated a significant negative shift on their perceptions of the accuracy of grading in the class (boys pretest M = 3.72, SD = 1.04; posttest M = 3.11, SD = 1.24, $p<.001$). Their responses also indicated greater neutrality in the accuracy of grading in the class. It's important to note that the perceptions here refer to the particular grade a student was attaining and whether or not he/she believed it was fair. The third research question deals with whether or not students perceive the grading system used in the class as a whole is fair.

Discussion Question 2

The perceptions students hold are important, because a student's agency, what he/she can and cannot control in the classroom, affects overall satisfaction in the course, and ultimately the subject. If boys are more consistently satisfied with the fairness of their science grades, they are more likely to continue study, resulting in a greater pool of potential students who might pursue this career path. Our data supports the premise that girls' negative perceptions affect their decision to continue in science and take higher level courses. For example, the class breakdown for grade 10 during the year the study was conducted was 41.25% girls and 58.75% boys. In the following year, the 11^{th} grade percentages were 31.70% girls and 68.30% boys, showing a much bigger drop in enrollment among girls.

Research Question 3: Do students perceive that the assessment system in class is unfair?

Entering perceptions on this question showed that gender significantly affected student perceptions of whether the assessment system in the class was unfair, $F(1,119) = 12.53$, $p<.001$, partial eta squared .095. The data showed that girls generally held the perception that the assessment system was not fair (girls M = 3.76, SD = .82). Whereas, their male counterparts in the class fell at the midpoint of the scale on whether the assessment system was fair or not (boys M = 3.09, SD = 1.11). However, the large standard deviation for males indicates greater variability in their responses on this question. On posttest, girls and boys did not differ significantly ($p>.63$) and they exhibited virtually the same mean scores (girls M = 3.40, SD = 1.01; boys M = 3.40, SD = 1.12).

To uncover whether there were differences in perceptions of fairness of grading within gender separate correlated samples t-tests were conducted for girls and boys. The results indicated a significant change for girls on perceptions of the fairness of grading from pretest to posttest two-tailed significance $t(45) = -2.84$, $p<.007$ (pretest $M = 3.76$, $SD = .82$; posttest $M = 3.30$, $SD = 1.00$). Likewise, for boys the results indicated a significant change from pretest to posttest two-tailed significance $t(74) = -2.08$, $p<.04$ (pretest $M = 3.09$, $SD = 1.11$; posttest $M = 3.40$, $SD = 1.13$). Over the course of the academic year girls showed a positive shift from pretest to posttest. Still, the shift indicates that girls became more ambivalent about the fairness of the scoring system. Boys, on the other hand, moved in the opposite direction from girls indicating a greater negative belief in the overall scoring system used in the class. Even with the shifts toward the middle of the scale, a large disparity still existed between the girls' and boys' perceptions of grading fairness.

Discussion Question 3

Although girls were achieving higher than boys in their science classes, they had the perception that the grading system was unfair. Why these perceptions? Even with their posttest shift toward the midpoint of the scale, it is still troubling. The shift was just to the midpoint of the scale reflecting "sometimes true/sometimes false," and this is hardly an endorsement about the fairness of grading.

Many studies point to the link between higher grades and student efficacy in the subject (Jinks & Morgan, 1996). For example, Bandura (1997) argues that students who achieve higher tend to display higher levels of self-efficacy, are more motivated, and show more interest in the subject. This study challenges that assumption for girls in science classes, or at least suggests that good grades may not be enough to affect their efficacy. In the case of these students, the girls were achieving at a higher rate, and even after a year with an outstanding teacher, did not believe the grading system was completely fair, suggesting that other, more subtle factors may have been in play that were negatively affecting the perceptions of girls. For example, girls may hear from the boys in the class that "boys are better at this" – whether through their actions or in actual teasing. For example, Sandler et al. (1996) reported that when girls do well on an assignment they are often questioned about whether or not they had help on the assignment. Boys, on the other hand, are asked this less frequently. Thus, girls construction of meaning and interpretations of their experiences with assessment, are complicated by the messages of teachers and peers within the classroom, which are powerful mediators of their developing beliefs. Other societal socializing may also be sending a negative message about girls in science, including signals from parents, the media, and even the depiction of men and women in their science books.

For teachers interested in raising the efficacy of girls in science, it means awareness that good, fair grades are not enough to offset prior and continuing socialization that science is "more for boys." It may mean using, as Gilbert (1996) asserts, "girl friendly" techniques. This means actively combating stereotypes and the negative reinforcement that girls may be internalizing from a variety of sources. Establishing an orderly system for distributing supplies and helping students during autonomous learning situations where boys cannot seize the spotlight is another step. Creating a climate of "zero tolerance" for put-downs, and giving encouragement for good work in a variety of ways beyond grades on tests or homework is also important. Finally, considering that so many science teachers are male (which again sends a gender message) the importance of bringing female role models to the classroom, emphasizing the achievements of women in science, and analyzing the depiction of women in texts and posters cannot be overstated.

Conclusions

The results of this study suggest that it is not enough to provide access to science for girls; nor do good grades ensure that girls will feel confident in their abilities in science. The entering and enduring perceptions the girls held about the assessment system is troubling because confidence in the assessment system is crucial for student participation and continuation in this discipline. When that confidence is violated, whether by overt teacher action or by teacher ambivalence to these perceptions, student ability to trust in their abilities is shattered, which can have significant negative consequences for learning and career pursuit.

Schools occupy an important place in the development of students' beliefs regarding their abilities and what is and is not an acceptable career path. Accordingly, students' beliefs can be a powerful predictor of the future action they will take. Bandura (1997) argued that the role of self-efficacy – the belief in one's ability to perform – is that "people's level of motivation, affective states, and actions are based more on what they believe than on what is objectively true" (p. 2). When we are assessed, we appraise our abilities and talents and plan courses of action on the basis of those beliefs. These beliefs, in turn, affect our expectancies for the future. When these perceptions are combined with the chilly climates reported by many girls in math and science classes, it is not difficult to imagine why girls do not pursue careers in these fields.

What can be done? Certainly schools need to alter the belief many girls hold about assessment in our science classrooms. The most obvious implication for teachers is that their grading system must not only *be* fair, but also as transparent as possible, within the limitations of confidentiality. Students who truly understand the basis of their grades (and the grades of their classmates) are more likely to believe in the fairness of the system. This study suggests that girls may need more feedback from science teachers, both in class and on tests, so they understand why their answers are right or wrong – as opposed to being left on their own to "feel stupid" or afraid to ask questions, as some girls in the study indicated.

Perhaps even more importantly, there are implications for climate – both within science classrooms, and extending throughout the school. Signs of a non-inclusive climate might include peer or teacher stereotyping that keeps students from exploring areas of interest, enrollment differentials in certain fields of study, or fewer girls seeking school leadership roles. Any of these could indicate a school-wide chilly climate for girls, and it is the responsibility of both teachers and school leaders to recognize these issues and take steps to establish a climate of inclusivity.

Limitations of the Present Study

Any self-reported information about perceptions among adolescents can be quite precarious, and their perceptions may be subject to influences not considered in this study. Further, the results of this study represent a snapshot at two particular points in time in once school district and may not apply to other samples or populations.

Future Directions

If we want more women in the sciences, we need first to investigate the ways that the sciences are taught in our classrooms. Researchers and curriculum designers should investigate ways to change the curricula to reflect various learning styles and gender differences. Science teachers, specifically, must explore and implement more girl-friendly methods, which will help girls succeed as well as help them see their place in science. As Ball (2002) argued, “curricula that represent the “voices,” images, and historical experiences of traditionally underrepresented groups are particularly important.” And school leaders must better understand the ways in which climate affects participation in traditionally male-dominated disciplines, and take steps to alter any cultural norms that may be subtly discouraging girls.

The under-representation of women in the sciences must also be viewed from a broader contextual viewpoint. Future studies should include measures on self-perceptions, perceptions of subject matter, and ability beliefs. Attempts should also be made to disentangle perceptions and beliefs that are based on school-effects (e.g., assignments, teachers, previous achievement) and societal-effects (e.g., socialization within the context of society, family and peer group). In our view, we cannot fully understand the under-representation of women in the sciences without understanding the forces that operate at various levels: societal, educational, familial, and personal. Only then can we hope to ameliorate the effects of prior beliefs and perceptions on girls’ willingness to pursue science as a viable career option.

References

Abrami, P., Chambers, B., Poulsen, C., De Simone, C., d'Apollonia, S., and Howden, J. (1995). *Classroom connections: Understanding and using cooperative learning*. Toronto, Ontario: Harcourt Brace & Company.

Allan, E. J., & Madden, M. (2003). *Chilly classrooms for female undergraduate students at a research university: A question of method?* Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 21-25).

Ball, A. F. (2002). Three decades of research on classroom life: Illuminating the classroom communication lives of America’s at-risk students. *Review of Research in Education, 26*, 71-112.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation. In N. Eisenberg (Ed.), *Handbook of child psychology* (Vol. 3, 5th ed., pp. 1017-1095). New York: Wiley.

Fennema, E., & Peterson, J. (1985). Autonomous learning behaviors: a possible explanation of gender-related differences in mathematics. In Wilkinson, L., & Marret, C. (Ed.) *Gender influences in classroom interaction*. New York: Academic Press.

Gilbert, J. (2001) . Science and its ‘Other’: looking underneath ‘woman’ and ‘science’ for new directions in research on gender and science education. Gender and Education, 13 (3), 291–305.

Goodenow, C. (1993). Classroom belonging among early adolescent students: Relationships to motivation and achievement. *Journal of Early Adolescence, 13(1),* 21-43.

Graham, J.W. & Smith, S.A. (2005). Gender differences in employment and earnings in science and engineering in the US. *Economics of Education Review, 24(3),* 341-354.

Johnson, D. W., Johnson, R., & Anderson, D. (1983). Social interdependence and classroom climate. *The Journal of Psychology, 114,* 135-142.

Kerr, Barbara, & Robinson Kurpius, Sharon E. (2004). Encouraging talented girls in math and science: Effects of a guidance intervention. *High Ability Studies, 15(1)*, 85-102.

Kleinfeld, Judith (1999). Student performance: Males versus females. *Public Interest,* 134, 3-20.

Lemonick, M. D. (2006, February). Are we losing our edge? *Time Magazine*, 167(7), 22-30.

McLaughlin, Robin (2005).Girls in science. *Science Scope, 28(7)*, 14-15.

Meece, J. L. (2002). *Child and adolescent development for educators* (2nd ed.). New York: McGraw-Hill.

Meece, J. L., & Eccles, J. S. (1993). Recent trends in research in gender education. *Educational Psychologist, 28*, 313-319.

Midgley, C. (2002). *Goals, goal structures, and patterns of adaptive learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Student/teacher relations and attitudes toward mathematics before and after the transition to junior high school. *Child Development, 60*, 981-992.

Morris, L. (2003). *The chilly climate for women: A literature review.* Paper presented at the Annual Meeting of the Mid-South Education al Research Association (Biloxi, MI, November 5-7).

National Center for Education Statistics. (1997). The conditions of education 1997: Women in mathematics and science. Washington, DC: Author.

National Center for Education Statistics. (2000). Findings from the condition of education 1997: Women in mathematics and science. Washington, DC: Author.

National Science Foundation. (1996). Women, minorities, and persons with disabilities in science and engineering: 1996 (NSF No. 96-311). Arlington, VA: Author.

Organisation for Economic Co-operation and Development (2003). Education at a glance. Retrieved February 18, 2006, at http://www.oecd.org/findDocument /0,2350,en_2649_37455 _1_119687_1_1_37455,00.html.

Packard, Becky Wai-Ling (2004). Student training promotes mentoring awareness and action. *Career Development Quarterly, 51(4),* 335-45.

Ransome, W. (1993). What every girl in school needs to know. *National Coalition of Girls' Schools,* Concord, MA. (BBB31886).

Salter, D.W. (2003). Exploring the "Chilly Classroom" phenomenon as interactions between psychological and environmental types. Journal of College Student Development, 44(1), 110-121.

Sandler, Bernice, and Hall, Roberta (1986). *The campus climate revisited: Chilly for women faculty, administrators, and graduate students.* Washington, DC: Project on the Status and Education of Women, Association of American Colleges.

Sandler, Bernice R., Silverberg, Lisa A. and Hall, Roberta M., (1996). *The chilly classroom climate: A guide to improve the education of women,* Washington, DC: National Association for Women in Education.

Science Scores Up in Grade 4, Stalled in Grades 8 and 12 (2006, May 24) USA Today, retrieved May 27, 2006, from www.usatoday.com/news/education/2006-05-24-science-scores_x.htm

Silverman, S., & Pritchard, A. M. (1996). Building their future: Girls and technology education in Connecticut. Journal of Technology Education, (7) 2. Retrieved Feb. 11, 2006, from http://scholar.lib.vt.edu/ejournals/JTE/jte-v7n2/silverman.jte-v7n2.html.

The Suitability of External Control-Groups for Empirical Control Purposes: a Cautionary Story in Science Education Research

Franz-Josef Scharfenberg
University of Bayreuth

Franz X. Bogner
University of Bayreuth

Siegfried Klautke
University of Bayreuth

Abstract

This article deals with a specific effect in one external control group incorporated to account for any pretest bias in a more comprehensive cognitive achievement study in a gene technology lab (as part of a modified Solomon's four-group plan). We monitored 12th graders (N = 117) in two external groups without any intervention: a one-test group (n = 55) and a three-test group (n = 62). Both samples participated in identical tests which quantified the relevant knowledge of the lesson unit applied in the main study. The three-test group yielded an unexpected increase in achievement scores. Subsequent analysis revealed two subsamples: one with no changes, the other with an increase (although without an intervention took place). A likely reason for the latter situation may lie in the role of the teacher(s) involved who might have wish to avoid potential negative results in his/her class. Consequently, we recommend the application of a modified Solomon's four group plan in science education research in order to prevent the influence of teacher intervention in future empirical analyses.

Correspondence should be addressed to Franz-Josef Scharfenberg, Centre of Math & Science Education, University of Bayreuth, Institute of Biology Didactics, Universitätsstr. 30, D-95445 Bayreuth, Germany. Phone: ++49-921-55-2590; Fax: ++49-921-55-2696. Email: franz-josef.scharfenberg@uni-bayreuth.de.

Introduction

A typical aspect of empirical analyses within science education is the desire to control as many variables as possible (Keeves, 1998). In particular, potential confounding variables, (e.g., maturation), possible external influences, or merely test repetition can threaten the internal validity of a study (Campbell, 1963; Campbell & Stanley, 1963). Almost 60 years ago, Solomon introduced a form of experimental design which is today typically referred as the Solomon four-group plan (Solomon, 1949). It is not the only four-group design (e.g., Huck & Chuang 1977; Marlatt, Demming & Reid, 1973) nor is it as commonly used as some other four-group plans (Rosenthal & Rosnow, 1997). However, this design is the only one known to assess adequately the confounding effect of pretesting with regard to the independent variables of interest (Walton Braver &

Braver, 1988). 'In this case, the process of measurement may change which has to be measured or repeating the measurement may enable subjects to perform more well' (Michel & Haight, 1996 [p. 367]). This effect is usually termed the 'pretest effect' (Bortz & Döring, 2001 [p. 505]), 'test reactivity' or 'pretest sensitization' (Huck & Chuang, 1977 [p. 409]), 'premeasurement sensitization' (Michel & Haight, 1996 [p. 367]), or 'memory carry-over' (McNemar, 1963 [p. 149]). The specific key element for the Solomon four-group design is that two groups (one experimental and one control group) perform a pre- and a posttest while the other two groups (again one experimental and one control group) receive no pretesting (Solomon, 1949). The comparison of the two control groups may unveil a potential pretest effect. Thus, such a design increases the degree of internal validity.

Although a number of studies report pretest effects, others do not (Rosenthal & Rosnow, 1997). Especially in the area of cognitive outcomes, Willson and Putnam (1982) described more (and additionally higher) pretest effects as given in studies of the affective domain. With regard to the delay between pre- and posttest larger effects were found for several days to two weeks. Additionally, these effects appeared to be larger for control groups than for treatment groups. Nevertheless, despite its advantages the Solomon four-group plan is underused (Walton Braver & Braver, 1988). Repeated calls for its application (e.g., Michel & Haight, 1996; Morgan 1997), especially in educational research (Cohen & Manion, 1994) have widely been ignored (e.g., Blanchard & Spence 1999). Walton Braver and Braver (1988) give four reasons for this: (i) the necessity of a higher number of subjects compared to simpler designs; (ii) the researcher's belief that pretest effects may not exist in his/her research arena; (iii) the greater difficulty of drawing conclusions due to the complexity of the design; and (iv) problems with regard to the statistical treatment of the results (e.g., Michel & Haight, 1996).

However, often in science education research as well as in our main study, it is impossible for investigators to perform experimental designs, because randomization is quite impracticable. Students in intact course groups allow only quasi-experimental designs (Cook & Campbell, 1979), and corresponding modifications of Solomons four-group plans have been described (e.g., Davies & Gould, 2000). With regard to the objective of our main study - monitoring the effectiveness of out-of-school laboratory work with regard to gene technology (Scharfenberg et al., in press) - many studies often lack the Solomon four-group design and indeed fail to include any special retest control (e.g., Killermann, 1998;Yager, Engen & Snider, 1969). Wilson and Putnam (1982) claimed that 'nonrandomized studies with pretests must be viewed with additional suspicion' (p. 256). They assume a potential bias due to pretesting likely be caused by the quasi-experimental selection of subjects per se. In order to counter this potential pretest effects we decided to incorporate in our quasi-experimental design two external control groups with no intervention: a three-test and a one-test group (Table 1). The specific objective of this present study thus is to investigate the suitability of such external control groups in science education research.

Methods

Design of the study

Our main study followed a quasi-experimental design (Cook & Campell, 1979) providing a modified Solomon four-group plan (Solomon, 1949). We combined three treatment groups within a comprehensive study and two external control groups without intervention (Table 1).

Table 1:
Design providing a modified Solomon four-group plan

Group		Test schedule						
		T-1	One week	Treatment		T-2	Six weeks	T-3
Treatment 1	Hands-on lab	O_1	→	X_1	→	O_2	→	O_3
Treatment 2	Nonexperimental lab	O_1	→	X_2	→	O_2	→	O_3
Treatment 3	Nonexperimental school	O_1	→	X_3	→	O_2	→	O_3
External control 1	Three-test	O_1	→			O_2	→	O_3
External control 2	One-test					O_2		

Note. O_n Outcome measure at test schedule T-n: T-1 pre-, T-2 posttest, T-3 retention test.

The objective of our main study was a quasi-experimental comparison of three instructional approaches. Our main method of instruction was a hands-on approach with a sequence of minds-on and hands-on phases in a dedicated out-of-school laboratory offered by us at the university. Two parallel methods covered identical contents but without experimenting (either in the laboratory or at school); in both cases, the content of the experimental lesson was taught in a problem-oriented learning modus (Reigeluth & Moore, 1999), but theoretically. We monitored cognitive achievement with respect to the upgrade of existing prior knowledge and to the acquisition of new knowledge. This was done in order to focus on the learning location effect (school vs. out-of-school lab without experiments) and of the experimentation itself (with experiments in the lab vs. non-experimental instruction in the lab or at school; for more details see, Scharfenberg et al., in press).

Students' sample

In all, 34 biology courses with 12th graders (N = 418; course size M = 12, SD = 3.7; age M = 18.0, SD = 0.68) participated in our main study. In order to establish similar courses in the different groups, we used only A-level ('Leistungskurs') students of the highest stratification level ('Gymnasium') in Bavaria (Germany). Additionally, all students have been enrolled in a regular half-year genetic course at school before participation in the study. This genetic education provided comparability of the courses: (i) The Bavarian Ministry of Education, Science, and Art (1991) obliges its content by the current syllabus; (ii) genetic education for all courses will be finished by a centralized formal exam at the end of high school. In general, the five groups did not differ in their prior achievement in biology (quantified as standard of written school work), Kruskal-Wallis-test $\chi^2(4, N = 394) = 3.65$, $p = .454$, and experiences with experimentation at school, Kruskal-Wallis-test $\chi^2(4, N = 404) = 9.17$, $p = .057$.

The hands-on group (n = 146) attended our teaching unit at the out-of-school laboratory. The day-long module "marker genes in bacteria" integrated four experiments into a lab lesson conformant with the syllabus. In general, the students worked in groups, mainly 3- or 4-person groups dependent of the course size actually given (M = 13, SD = 4.0). They transformed bacteria with a recombinant plasmid (coding for the Green Fluorescent Protein, Tsien, 1998), they isolated the plasmid and analyzed it with common restriction enzymes. At least, they visualized their results by agarose gel electrophoresis. All experiments followed the criteria of authentic inquiry (Chinn & Malhotra, 2002). The nonexperimental lab group (n = 72) followed the same themes at the lab-site but without hands-on experiments. The school group (n = 83) was taught the identical content at school (again without experimental activities). A single teacher previously unknown to all students taught all lessons. A consistent problem of studies comparing experimental and nonexperimental instruction lies in the students' different time exposures. This problem has often been ignored (e.g., Killermann, 1998); others (e.g., Saunders and Dickinson 1972, p. 461) used actions like "discussion of material presented in lecture" in order to achieve identical time schedules. Following this rationale, we included a nonexperimental "lab+time" group in our pilot study one year ago and provided a typical lab working environment in combination with printed information which allowed repetition of the themes taught. However, cognitive learning outcomes were similar (Scharfenberg, 2005), and we omitted this kind of treatment in our main study. The results of the three treatment groups have been described elsewhere (Scharfenberg et al., in press).

With regard to the control groups, we modified the Solomon four-group plan by omitting the control group with treatment and one test (posttest after treatment) because of the impossibility of organising three such groups with regard to each treatment. Altogether, 11 courses (N = 117) were assigned to the two external control groups: a three-test (n = 62) and a one-test group (n = 55). They received no corresponding instruction to permit us to examine potential pretest effects or other external influences (Hofstein & Lunetta, 1982; Keeves, 1998), but proceeded with the regular lessons being taught by their teachers.

Assessment of the questionnaire used

Generally, questionnaires were applied three times, as pretest (T-1) one week before participating, as posttest (T-2) immediately after and as retention test (T-3) six weeks later. In contrast, the one-test group responded only once to the test. The questionnaire covered cognitive achievement items dealing with the lesson content, and consisted of 15 multiple-choice and one open item (see, e.g., Table 2). In our main study we applied two levels of analysis: (i) one dealing with a student's expected task performance such as reproduction (rendering of facts from memory; seven items), reorganization (self-acting rearrangement of facts to a new knowledge structure; four items) and transfer of knowledge (self-acting application of known facts to an unknown example; five items; all definitions by Deutscher Bildungsrat, 1970); (ii) the other with content relation referring to testing updated prior knowledge (seven items) and newly attained knowledge (nine items) validated by a latent class analysis on students' individual response pattern (for details, see, Scharfenberg et al., in press).

Table 2:
Listing of five item examples providing the achievement survey

	Item	Item difficulty[a]	Item characterization[b]: Expected task performance	Item characterization[b]: Content relation
1	A plasmid for heterologous gene expression has to be constructed. Transformed bacteria will express the heterologous gene, if the following DNA segments are arranged in this way: a) antibiotic resistance gene → inserted gene → origin of replication. b) origin of replication → inserted gene → bacterial promotor sequence. c) bacterial promotor sequence → inserted gene → antibiotic resistance gene [correct]. d) inserted gene → origin of replication → bacterial promotor sequence.	30,1	Transfer	Updated prior knowledge
2	If an operon is positively controlled the 'switching molecule' starts a) translation of DNA. b) translation of m RNA. c) transcription of m RNA. d) transcription of DNA [correct].	49,7	Re-production	Updated prior knowledge
3	Green fluorescent protein (GFP) can be used in molecular biology in different ways because a) it is easy to detect its infrared fluorescent high emission. b) it is easy to generate fusion proteins with GFP and other proteins [correct]. c) it can diffuse in an organism form one cell to another. d) its luminescence component alone is useful.	52,1	Reorgani-sation	Newly attained knowledge
4	An electrophoresis apparatus consists of the following parts: a) one electrode, two buffer chambers, one gel carrier. b) two electrodes, two buffer chambers, two gel carriers. c) two electrodes, one buffer chamber, two gel carriers. d) two electrodes, two buffer chambers, one gel carrier [correct].	77,1	Re-production	Newly attained knowledge
5	A plasmid contains three recognition sites for the restriction enzyme *Bam* HI and one recognition site for the restriction enzyme *Eco* RI. How many fragments will result in a double digest with these two enzymes [four]?	46,9	Transfer	Newly attained knowledge

Note. [a]Item difficulty = % of correct answers (Bortz & Döring 2001).
[b] Two level of analysis: see text for details.

The consistency with the existing syllabus (Bavarian Ministry of Education, Science, and Art, 1991) provided content validity; in-service teachers (N = 12) provided an affirmative expert rating (similarity of the lesson to the syllabus as *good* or *excellent*). Positive correlations betweens students' test scores without any intervention with their prior achievement in school (quantified as their standard of written school work in biology) supported the convergent validity (as criterion-related validity type; see, e.g., Bortz & Döring, 2001) of the questionnaire for knowledge assessment at all (Spearman rank correlation coeffizient T-1 $r_s = .320$, $p = .011$, $n = 62$, T-2 $r_s = .231$, $p = .013$, $N = 117$, T-3 $r_s = .344$, $p = .006$, $n = 62$). Cronbach's alpha of the test scores was .68 (T-2, $N = 418$). The item difficulties were normally distributed (Kolmogorov-Smirnov test with Lilliefors modification $p = .200$, see Figure 1a) and the corrected item-total correlations were for 9 items > .3 ($p < .001$) and for 7 items > .2 ($p < .001$). We accepted the latter in spite of the low value (below .3) because of the complexity of the lesson content involved (Diehl & Kohr, 1999). Additionally, the corrected item-total correlations relate to item difficulties in a parabolic way (Lienert, 1969, see Figure 1b), a fact that item selection has to take into account (Bortz & Döring, 2001).

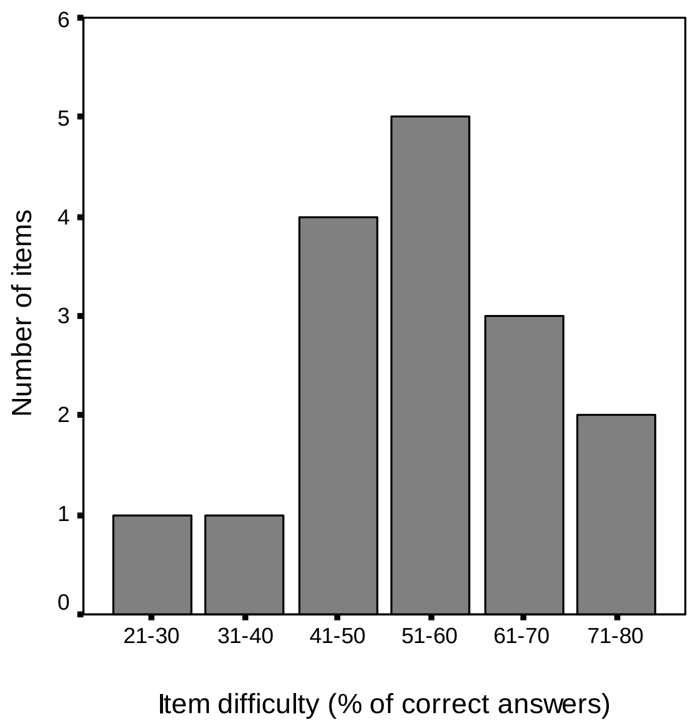

Figure 1a:
Distribution of questionnaire item difficulties with regard to patterns of ten units.

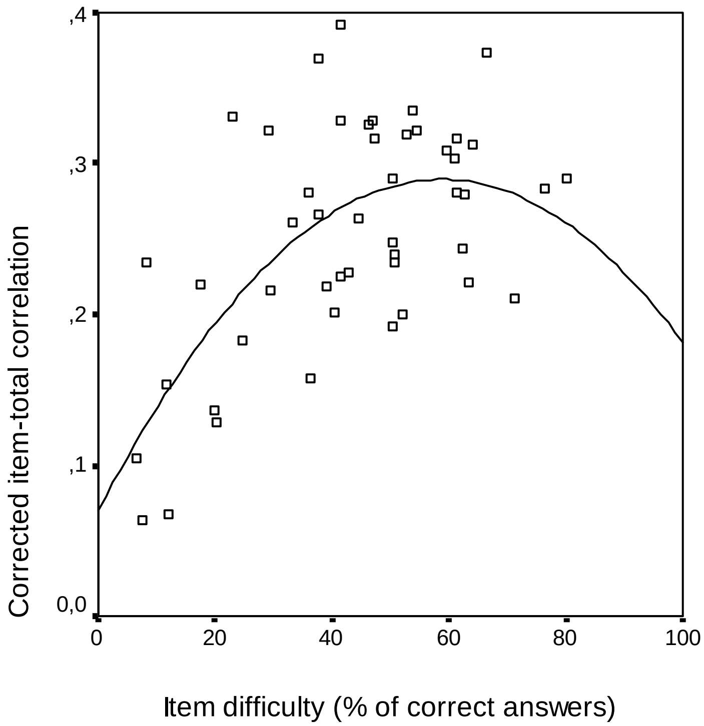

Figure1b:
Example of the parabolic relation between item difficulty and corrected item-total correlation (48 items of T-1, T-2, T-3, c.f. Lienert, 1969)

Statistical procedures

For each test and student, a total score was calculated as the number of correct answers. Due to the partial lack of normal distribution of our data, nonparametric methods were applied (Kolmogorov-Smirnov test with Lilliefors modification three-test group T-1, p = .036, T-2, p = .195, T-3, p = .019; one-test group T-2, p = .001). Consequently, we used boxplots as graphical charts. The statistical significance of changes of scores within all three test schedules was analysed using the Friedman-test, followed by pair-wise analyses from T-1 to T-2 and T-3 and from T-2 to T-3 using the Wilcoxon signed-rank test. The Mann-Whitney-U test was employed to test for pair-wise intergroup differences. An alpha level of .05 was used for all statistical tests.

Results

Our data analysis revealed an (unexpected) result with regard to the three-test control group, i.e. an increase of knowledge despite the lack of intervention, Friedman test $\chi^2(2, n = 62) = 8.673$, $p = .013$ (Figure 2).

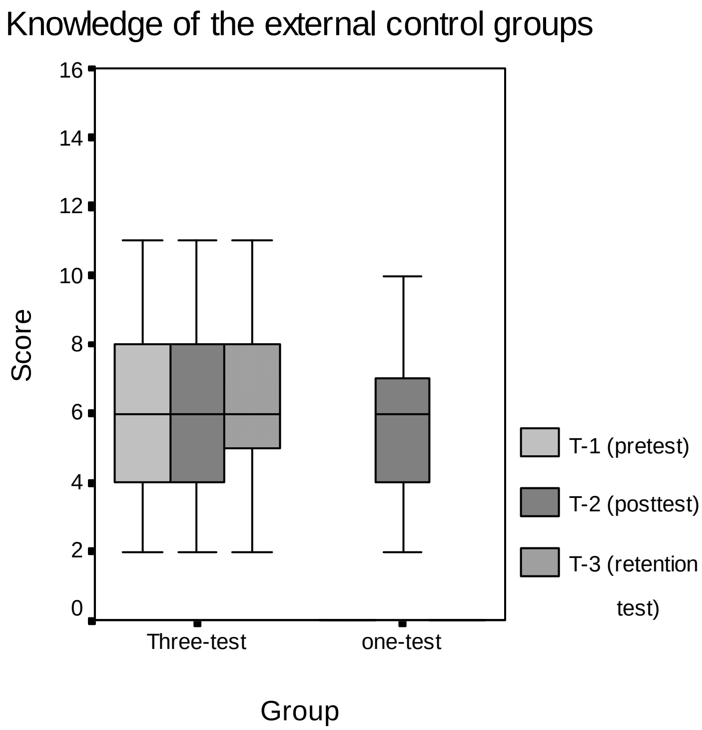

Figure 2:
Knowledge scores of the three-test and the one-test control groups without any intervention

Subsequent pair-wise analysis showed a statistically significant change from T-1 to T-3 (Wilcoxon signed-rank test Table 3).

Table 3:
Comparison of the one-test and the three-test group with regard to knowledge

	Control groups[a]			
Statistics	One-test	Three-test	Three-test	Three-test
Test dates	T-2	T-1 to T-2	T-2 to T-3	T-1 to T- 3
Mdn (grouped)	5.4	5.6 to 6.2	6.2 to 6.4	5.6 to 6.4
Z[a]	-	1.857	1.204	2.137
p	-	.063	.228	.033

Note. [a]*In total N = 117, one-test group n = 55, three-test group n = 62.*
[b]*Wilcoxon signed-rank test (based on negative ranks).*

Comparison of the one-test group scores with the scores of the three-test group revealed a statistically significant difference only at the testing schedule T-3 (Mann-Whitney-U-test T-3, $p = 0.029$, Table 4).

Table 4:
Comparison of the two external control groups with regard to knowledge

	Comparison of the one-test group with the three-test group				
Statistics	as a whole			subsample-1	subsample-2
Test schedule	T-1	T-2	T-3	T-3	T-3
Mann-Whitney-*U*	1630.000	1442.500	1308.500	944.000	272.500
Z	0.414	1.448	2.188	0.378	4.521
p	.679	.148	.029	.705	<.001

Note: *Mann-Whitney-U-test shows statistically significant differences at T-3 between the one-test group and the three-test group as a whole and its subsample-2, in contrast to subsample-1 (see Figure 3).*

Subsequent analysis of the three-test group on the level of individual courses indicated a distinction of two separate subsamples (Figure 3), one with no significant change over the three survey schedules, the other with substantial change, Friedman-test subsample-1 χ^2 (2, $n = 36$) = 1.350, $p = .509$; subsample-2 χ^2 (2, $n = 26$) = 27.482, p <.001.

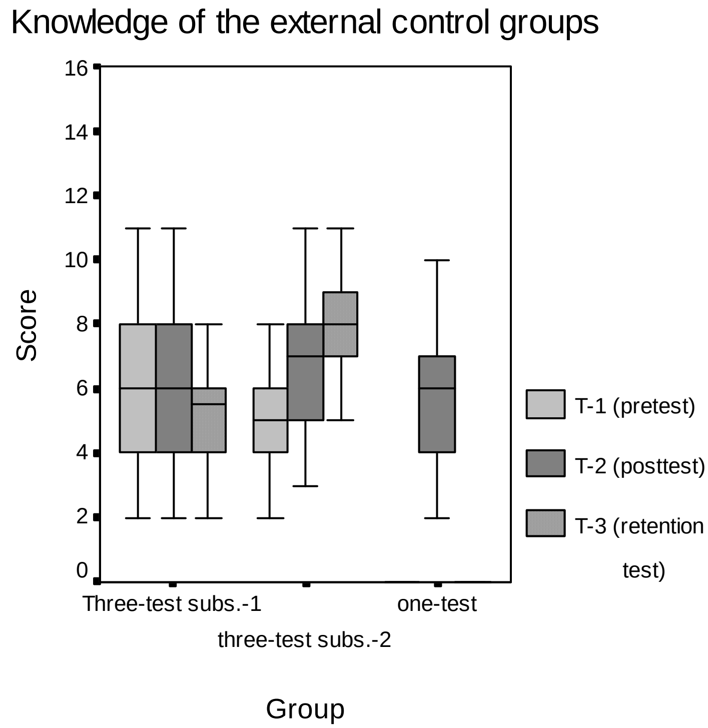

Figure 3:
Test scores of two subsamples (= subs.) extracted from the three-test group and compared to the score of the one-test group

Subsample-2 differed statistically significantly from the one-test group at T-3 (Mann-Whitney-U-test $p < .001$, Table 4). A comparison of both subsamples' scores on the three different tests also showed a significant difference at the testing schedule T-3 (Mann-Whitney-U-test $p < .001$, Table 5).

Table 5:
Comparison of both subsamples of the three-test group with regard to knowledge at the different test schedules

	Test schedule		
Statistics	T-1	T-2	T-3
Mann-Whitney-*U*	433.000	405.000	166.000
Z	0.504	0.907	4.353
p	.615	.365	<.001

Note. *Subsample-1 with no significant change over the three survey schedules (n = 36), subsample-2 with substantial change (n = 26).*

A subsequent pair-wise analysis of subsample-2 (Wilcoxon signed-rank tests, Table 6) showed significant changes across the three distinct survey dates. Despite the lack of intervention pupils of this specific subsample achieved a step by step increase in their level of knowledge, especially in the final T-3 (the retention test six weeks later). We conclude that some unknown factor must have affected the knowledge level of this subsample of the three-test control-group.

Table 6:
Cognitive achievement in subsample-2 of the three-test group

Statistics	Three-test group subsample-2[a]		
Test dates	T 1 to T 2	T 2 to T 3	T 1 to T 3
Mdn (grouped)	5.5 to 6.5	6.5 to 7.9	5.5 to 7.9
Z^a	3.063	3.689	3.811
p	.002	<.001	<.001

Note. ^{a} *n = 26*.
[b]*Wilcoxon signed-rank test (based on negative ranks).*

Discussion

The result of the three-test group in subsample-2 was quite unexpected, and we have no explanation for the surprising increase of achievement scores. However, we agree with Keeves (1998) that 'there is more information available in most well-designed evaluation studies', especially, when instruments used have the potential to be more generally introduced besides the original intent. Consequently, an external group scoring may enable us to gain additional insights into potential effects with regard to suitability of control-group.

Although many studies of the efficacies of laboratory activities rarely employ the Solomon four-group design or indeed any special retest control at all (e.g., Killermann 1998; Yager et al., 1969), we incorporated such a modified design to take account of potential pretest effects. Neither pupils nor teachers had any contact either with each other or with any course used as treatment group in our main study. Such contact was excluded due to the distances between the individual testing sites and the general survey schedule. Furthermore, the control courses did not take part nor planned to do so in any educational laboratory elsewhere.

The significant difference at T-3 between the one-test group and the three-test group as a whole firstly might hint at an effect of repeated testing. There may be a learning effect of participation in a pretest (T-1) that carries over to a following posttest (T-2). However, we found no difference between the one-test group and subsample-1 of the three-test group, suggesting that there is no bias due to repeated testing.

The unexpected gain in achievement level in subsample-2 of the three-test group may be attributed either to the students alone or to the teachers as well. With regard to the students, Cook and Campbell (1979) suggest that maturational factors may cause the increase in scores. This argument seems improbable since maturation may be assumed common to the sample as a whole and not specifically affect just this subgroup. Although students were unaware of the repeated testing schedules, the scores increased over time despite a gap of about six weeks between the posttest and the retention test. The second major explanation is the introduction of possible bias by the teacher: We have no evidence for this, but the hypothesis cannot be excluded as source of the disruptive influence.

Neither can we exclude the possibility of external influences via regional or supra-regional media. Schweiger and Brosius (1999), for instance, described the potential influence in their external control-group of news concerning the possible cloning of humans on pupils' specific attitudes towards gene technology. This too we find unlikely because all groups would be subject to such influences. Consequently, an intervention by the teacher seems more likely whether it occurred unconsciously or deliberately. Thus, the students may have been prepared for a routine assessment test perhaps by repetition of the selected knowledge necessary for the survey. A step-wise achievement effect is also feasible, resulting from a teacher either specifically preparing his/her pupils for the surveys or simply announcing that a second survey was to be conducted, thus motivating his/her pupils to recall previous test details in order to achieve better results.

A further explanation may be social desirability. Some indication of this is the fact that the investigator and the test administrator differed in the control-group surveys (in contrast to the treatment-groups). In all external control-groups the teachers acted as mediators. Some teachers in the three-test group may not have intervened with their samples. Mediators might have a specific interest in conflict with the investigators' interests in adhering to the standards of an empirical study. A teacher as mediator may intercede in order to get better results on his/her pupils as a desirable social objective, particularly if there is repeated testing. Two potential reasons may explain this. Firstly, a mediator may fear that investigators may get a poor impression of his/her capability. Secondly, he/she may have doubts about the anonymity of a survey, and hence fear bringing shame on his/her school.

Conclusions

With regard to our specific results of this study, we encourage the use of one-test groups in quasi-experimental research designs following a modified Solomon four-group plan. This might facilitate the identification of the above mentioned pretest effects in this design (Wilson & Putnam, 1982), especially such a form of retest effect as the one we observed in our subsample-2 of the three-test group. Furthermore, the selection of external control-groups needs careful action to exclude or at least reduce the influence of mediators' own interests. For instance, investigators could explicitly refer to the survey's anonymity and/or point out that the analysis of particular courses would be irrelevant to the study. We therefore emphasize to the complex issue of control-groups' selection in general, and to the necessity of ensuring that those control-groups function as intended. Taken this in account, a quasi-experimental design might provide more convincing empirical results with regard to the treatment groups as it is the case in our main study: Compared to the conventional learning location at school, hands-on activities with authentic experiments in out-of-school laboratories supported a substantial increase in knowledge (Scharfenberg et al., in press). As a consequence for science teaching, we suggest to offer teachers such out-of-school laboratories, especially, when authenticity is available (which is impossible to achieve at school). Nevertheless, any out-of-school experiment should be integrated within a teaching framework in a laboratory situation, thus, enabling students to actualize existing prior knowledge (as a precondition for the attainment of new knowledge). In our case, a specific teaching and learning unit assisted our students to develop individual hypotheses before participating in any hands-on activity and to verify/falsify his/her hypotheses. We appreciate, of course, the same frame in the context of demonstration experiments, either in a lab or in a school situation.

Acknowledgement

The study was funded by the Bavarian State Ministry of Regional Development and Environmental Affairs (Bayerisches Staatsministerium für Landesentwicklung und Umweltfragen), the Bavarian State Ministry of Education (Bayerisches Staatsministerium für Unterricht und Kultus) and the German Science Foundation (Deutsche Forschungsgemeinschaft; grant KL 664/5-1).

We welcome the cooperation of teachers and pupils involved in this study. Additionally, we appreciate the helpful and valuable discussion of the manuscript with C. Lehner, M. Wiseman and S. Tomkins.

References

Bavarian Ministry of Education, Science, and Art (Bayerisches Staatsministerium für Unterricht, Kultus, Wissenschaft und Kunst, Eds., 1991). Lehrplan für das bayerische Gymnasium, Fachlehrplan Biologie [Syllabus of the Bavarian secondary school at highest stratification level, Specific Syllabus of Biology]. *Amtsblatt des Bayerischen Staatsministeriums für Unterricht, Kultus, Wissenschaft und Kunst, So.-Nr. 7*, 1125-1172.

Blanchard, C., & Spence, J. C. (1999). The effect of pretesting on felling states and self-efficacy in acute exercise. *Research Update Alberta Centre for Active Living*, *7* (1), 1-2.

Bortz, J., & Döring, N. (2001). *Forschungsmethoden und Evaluation* (3rd ed.) [Methods of research and evaluation]. Berlin, Germany: Springer.

Campbell, D. T. (1963). From description to experimentation: interpreting trends as quasi-experiments. In: C. W. Harris (Ed.), *Problems in measuring change* (pp. 212-244). Madison, WI: The University of Wisconsin Press.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Dallas, TX: Houghton Mifflin Company Boston.

Cohen, L., & Manion, L. (1994). *Research methods in education* (4th ed.). London: Routledge.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation. Design & analysis issues for field settings*. Chicago: Rand McNally College Publishing Company.

Davies, N., & Gould, D. (2000). Updating cardiopulmonary resuscitation skills: a study to examine the efficacy of self-instruction on nurses' competence. *Journal of Clinical Nursing*, *9*, 400-410.

Deutscher Bildungsrat (Eds.) (1970). *Empfehlungen der Bildungskommission: Strukturplan für das Bildungswesen* [Recommendations of the Commision for Education: Curriculum for he educational system]. Stuttgart, Germany: Klett.

Diehl, J. M., & Kohr, H. U. (1999). *Deskriptive Statistik* (12th ed.) [Descriptive statistics]. Eschborn, Germany: Verlag Dietmar Klotz.

Hofstein, A., & Lunetta, V. N. (1982). The role of the laboratory in science teaching: neglected aspects of research. *Review of educational research*, *52,* 201-217.

Huck, S. W., & Chuang, I. C. (1977). A quasi-experimental design for the assessment of posttest sensitization. *Educational and Psychological Measurement*, *37*, 409-416.

Keeves, J. P. (1998). Methods and processes in research in science. In: B. J. Fraser & K. G. Tobin (Eds.), *International Handbook of Science Education, Part Two* (pp. 1127-1153). Dordrecht, Netherlands: Kluwer Academish Publishers.

Killermann, W. (1998). Research into biology teaching methods. *Journal of Biological Education*, *33* (1), 4-9.

Lienert, G. A. (1969). *Testaufbau und Testanalyse* (3rd ed.) [Development and analysis of tests]. Weinheim, Germany: Verlag Julius Beltz.

Marlatt, A. G., Demming, B., & Reid, J. B. (1973). Loss of control drinking in alcoholoics: a experimental analogue. *Journal of Abnormal Psychology*, *81* (3), 233-241.

McNemar, Q. (1963). *Psychological statistics* (3rd ed.). New York: John Wiley and Sons.

Michel, Y., & Haight, B. K. (1996). Using the Solomon four-group design. *Nursing research*, *45* (6), 367-369.

Morgan, W. (1997). *Physical activity and mental health*. Washington, DC: Taylor & Francis.

Reigeluth, C.M., & Moore, J. (1999). Cognitive Education and the Cognitive Domain. In: Reigeluth, C. M. (Ed.), *Instructional-Design Theories and Models Volume II* (pp. 51-68). Mahwah, NJ: Lawrence Erlbaum Associates.

Rosenthal, R., & Rosnow, R. L. (1997). *Essentials of behavioural research. Methods and data analysis* (2nd ed.). New York: McGraw-Hill.

Saunders, W. L., & Dickinson, D. H. (1976). A comparison of community college students' achievement and attitude. Changes in a lecture-only and lecture-laboratory approach to general education (biological science courses). *Journal of Research in Science Teaching 16* (5), 459-464.

Scharfenberg, F.-J., 2005 *Experimenteller Biologieunterricht zu Aspekten der Gentechnik im Lernort Labor: empirische Untersuchung zu Akzeptanz, Wissenserwerb und Interesse* [Hands-on teaching in biology in an out-of-school laboratory with regard to gene technology: empirical study with respect to acceptance, knowledge achievement and interest]. Doctoral dissertation, University of Bayreuth, Germany. Retrieved December 4 from http://opus.ub.uni-bayreuth.de/volltexte/2005/176/ .

Scharfenberg, F.-J., Bogner, F., & Klautke, S. (in press). Learning in a gene technology lab with educational focus: Results of a teaching unit with authentic experiments. *Biochemistry and Molecular Biology Education*.

Schweiger, W. & Brosius, H.-B. (1999). *Von der 'Gentomate' zur Gentechnikakzeptanz. Eine Panelstudie zu Einstellungseffekten eines rollenden Genlabors an Gymnasien* [From the 'gene tomato' to acceptance of gene technology. A panel study to effects of a mobile gene technology lab at secondary schools with regard to attitudes]. Neuherberg, Germany: Gesellschaft für Strahlenschutzforschung.

Solomon, R. L. (1949). An extension of control group design. *Psychological Bulletin*, *46*, 137-150.

R. Y. Tsien (1998) The green fluorescent protein, *Annual Review of. Biochemistry*. *67*, 509-544.

Walton Braver, M. C., & Braver, S. (1988). Statistical treatment of the Solomon four-group design: a meta-analytic approach. *Psychological Bulletin*, *104* (1), 150-154.

Willson, V. L., & Putnam, R. R. (1982). A meta-analysis of pretest sensitization effects in experimental design. *American Educational Research Journal*, *19* (2), 249-258.

Yager, R. E., Engen, H. B., & Snider, B. C. (1969). Effects of the laboratory and demonstration methods upon the outcomes of instruction in secondary biology. *Journal of Research in Science Teaching*, *6* (5), 76-86.

The Origin and Extent of Student's Understandings: The Effect of Various Kinds of Factors in Conceptual Understanding in Volcanism

Burckin Dal
Istanbul Technical University

Abstract

This study investigated and compared 130 students' perceptions of volcanoes and volcanic activity from an inner city elementary school (Year 6), middle school (Year 9) and student teachers in the science education department. A qualitative and quantitative methodology was used for this investigation. The data collection was based on three research stages: collection of information by the association of ideas, a Q-Sort test and a questionnaire with open-ended questions. The findings indicated that the sample possessed an incomplete picture of volcanoes and volcanic activity including many alternative conceptions about it. Both the students and the student teachers had surprisingly similar alternative conceptions despite the fact that the latter received more instruction on this topic. Moreover, over the course of the curriculum, a closer relationship between alternative conceptions and accepted scientific knowledge was evident. Hence, it was possible to map out the categories of alternative conceptions of volcanism and to measure the influence of the curriculum by looking at the evolution of these alternative conceptions. Based on the results, some suggestions to help teachers and students avoid critical barriers to learning that may be difficult to overcome later in their education are presented.

Correspondence should be addressed to Burckin Dal, Istanbul Technical University Eurasia Institute of Earth Sciences 34469, Maslak, Istanbul / Turkey
Email: dalbu@itu.edu.tr

Introduction

At Key Stages 2 and 3 (7-14 years), the French National Science Education Standards require children to develop a scientific understanding of the Earth's materials and processes and, logically therefore, of the Earth's structure that provides the context for such an understanding (B.O.E.N, 2000a-b). In seeking to support children's understanding in this domain, constructivist theories of learning and teaching provide a model which highlights the importance of children's existing ideas as the focus for conceptual change.

What are the students' ideas in Earth Sciences? Is it possible to classify their ideas into categories, e.g. physics, biology or chemistry? Is it possible to analyze and explain their ideas which persist, the difficulties and barriers to learning?

During the last three decades, there has been an increasing awareness of the importance of the variation amongst the ways students conceptualize and think about the phenomena they encounter in either science classes or in daily life (Driver, Guesne & Tiberghien, 1985; Giordan & De Vecchi, 1987). Various science conceptions constructed in students' minds are called "misconceptions", "preconceptions" or "alternative frameworks" (Schnotz, Vosniadou & Carretero, 1999). If these terms are examined for similarities, they have almost the same meaning. In this article, the term *alternative conception* is used to describe any conceptual difficulties, which is different from or inconsistent with the accepted scientific definition.

In most cases, alternative conceptions are not necessarily spontaneous ideas. In fact, they arise from *knowledge at different levels* (Chevallard, 1991). These alternative conceptions may be a product of instruction or come from the analogies used by teachers or from textbooks. At the same time, they might arise from the discrepancy between daily language and scientific language as well as the students' social environments (Happs, 1985; Ford, 2005). Moreover, when the teachers do not fully understand the scientific content of and believe their conceptions are correct, they may cause the students to have alternative conceptions (Osborne & Freyberg, 1985).

If teachers have alternative conceptions, they will have difficulties identifying the students' alternative conceptions and correcting them. For this reason, an investigation of alternative conceptions amongst science student teachers who have had instruction on fundamental concepts of science and a comparison of the student teachers' alternative conceptions to those of students would both be worthwhile.

Researchers in many countries have tried to answer a number of important questions related to children's conceptions such as which alternative conceptions occur, what their sources may be, how extensive they are, why they are so resistant to instruction, and what teachers can do to facilitate conceptual change.

Most research in Earth Science alternative conceptions has focused on concepts in;

- The Earth's materials: minerals and rocks (Happs, 1982, 1985; Dove, 1998; Ford, 2005);
- The Earth's Structure (Lillo, 1994; Sharp, *et al.* 1995; Blake, 2005);
- The Earth's processes: mountains, volcanoes, earthquakes, weathering and erosion (Russell, *et al.* 1993; Bezzi & Happs 1994; Sharp, *et al.* 1995; Blake, 2005; Dal, 2005)
- Geological time (Ault, 1982; Trend, 1998).

Whilst there have been a number of informative studies on children's ideas about the Earth's materials such as rocks and minerals, there is very limited research on children's thinking about the Earth's processes and particularly about volcanoes. Findings take into consideration language students and teachers employed to describe and define volcanoes, the criteria used by novices to classify and identify the products from volcanic eruption, the nature of the materials constituting the volcanic structure, the make-up of this structure and its link to the Earth's make-up.

According to the criteria of researchers in the previous literature, the student's ideas appeared to represent three different alternative conceptions of descriptive and causal understanding for volcanoes and volcanic activity. Bezzi & Happs (1994), Sharp, *et al.* (1995) and Dal (2005) found that the student who failed to recognise volcanoes and volcanic activity as natural events (Level 1). The students who did recognise volcanoes and volcanic activity as natural events, but said that they did not occur in the past (Level 2) (Ault, 1982; Lillo, 1994; Trend, 1998; Blake, 2005), while the students who believed they had occurred in the past as well as the present (Level 3) (Ault, 1982; Sharp, *et al.*

1995; Trend, 1998). With regard to causation, Level 1 was assigned to children who did not know what caused volcanic eruptions ("it happened by itself") or offered some human cause, e.g., loud noises or people walking on the ground causing vibrations (Bezzi & Happs, 1994; Sharp, *et al.* 1995; Dal, 2005). Level 2 causes were naturalistic, but either eccentric (for example, the volcano gets too full of lava), or approached some degree of scientific understanding (for instance, internal heat from the core or magma being forced up by gas pressure beneath the Earth) (Ault, 1982; Bezzi & Happs, 1994; Lillo, 1994; Blake, 2005). Level 3 responses related volcanic activity to the action of crustal plates (Sharp, *et al.* 1995; Trend, 1998; Dal, 2005).

These studies confirm that children develop their own, mostly non-scientific, understanding of volcanoes concepts before instruction, and describe and interpret these in everyday terms that are familiar to them. In addition, they have shown not only how generally limited children's understanding of the volcanoes, the products from the volcanic eruption and the nature of the materials constituting the volcanic structure is, but how different their conceptions are from those of Earth scientists. Importantly, key "critical barriers" to children developing a scientific understanding in this domain have been identified (Ault, 1982; Bezzi & Happs, 1994; Trend, 1998):

- large scale patterns in the environment and the physical changes they represent;
- geological time.

It is important to note that these early studies have tended to focus more on the ideas about volcanoes and volcanic activity held by samples selected only from one course of the curriculum. Different from previous literature, this paper reports on a study that investigated the ideas about volcanoes and volcanic activity of samples selected from different types of curriculum. In addition, investigation and comparisons will be made of the understanding of the alternative conceptions of the student teachers and Year 6 and Year 9 students to determine if the student teachers' alternative conceptions can be the source of the students' alternative conceptions.

This paper begins with a discussion regarding the nature and the complexity of knowledge at different levels in the conceptually confined domain of volcanism. There then follows an investigation and a comparison of the levels of understanding of children from Year 6 (10-11 years), Year 9 (13-14 years) and student teachers in the first two years of their professional training about concepts of volcanism. Such discussions will reveal how, over the course of the curriculum, although some alternative conceptions persist, others appear at given times or evolve. Equally, the role classroom materials play in this evolution will be evaluated. This study is not centred on the description of barriers to learning, but it may constitute a preliminary study to aid their identification. Identifying barriers to learning, not only helps to develop teaching strategies but also provides a first step for future research, which should address the effects of developing guide materials and teaching strategies as well as organizing workshops for inservice and preservice teacher training. Finally, such implications of the findings of this study for further research and for classroom practice will be addressed.

Research Design: the knowledge concerned

The Importance of the Structure of Knowledge

The study on knowledge at different levels provided the epistemological framework of research: it has allowed an *ex ante* analysis to take place and to sketch out the methodology for both the collection of the data and for interpretation.

Let me to clarify what the definition of knowledge at different levels is and what kind of relationship amongst these different definitions of knowledge exists.

Students are presented with different ideas about the Earth's processes during their schooling at all levels (Happs, 1982, 1985). According to Giordan & De Vecchi (1987), these ideas originate from knowledge that has been gradually acquired by the scientific community during the development of science, and which shall be named *reference knowledge*. According to Chevallard (1991), the *reference knowledge* chosen is decontextualised knowledge. This *reference knowledge* is therefore the object of a *transposition* (recontextualisation, or even a redefinition) to be taught at a given level. This first transposition from *reference knowledge* to *knowledge to be taught*, is thus, in fact, followed by a second transposition, which, carried out by teachers (but also by Official Instructions and textbooks etc.) establishes a *school knowledge* with its own particularities. So, the *school knowledge* is a knowledge rebuilt specifically for teaching. Finally, the last stage of the student's work is to interpret the knowledge "the way he or she can" during various steps which will lead him/her to transform it into *acquired or assimilated knowledge* in a particular context.

In fact, it appears obvious that the teaching offered plays a role in the construction of the students' knowledge (Driver, *et al.* 1994), but what exactly is this role? How is the received information integrated? Is it the logic of knowledge which organises the student's mental models?

The research was carried out into the acquisition of this logic of knowledge by analysing its' structure at different levels: from *reference knowledge* to *school knowledge*. Thus it was possible to establish *ex ante* hypotheses about the way in which the pupils are most likely to approach volcanicity and the notional fields concerned. These two stages served as milestones in constructing the finally employed methods.

In addition, the history of Geology is informative about the various ways in which the problem of volcanoes has been tackled over time: these other approaches correspond to different contexts and times when the current techniques and theories did not exist: they were helpful in the interpretation of answers of pupils who did not yet have scientific theories at their disposition either.

Thus the two following questions addressed were: How does the knowledge concerning the case of volcanism come about**?** How has the geological phenomenon been interpreted since Antiquity?

How Does the Knowledge Concerning the Case of Volcanism come about?

The analysis of knowledge was based upon: with regards to *reference knowledge*, on possible sources of documentation of middle school Geology teachers, thus on works illustrating a first sphere of knowledge transfer from the field of research to that of university education (e.g. Francis, 1993; Scarth, 1999) or to that of a popularisation considered as satisfactory in the scientific world (e.g. Sigurdsson, *et al.* 2000; Frankel, 2005); with regards to the *knowledge to be taught*, on the curriculum and on Official Instructions (B.O.E.N, 1978a-b, 1979a-b and 1985a-b) for the levels considered as well as textbooks. According to Stern & Roseman (2004*)* and Ford (2006)*,* the textbook "proposes an order for learning", as much with regard to the general organisation of contents as to the organisation of teaching. Hence, it is related to the knowledge to be taught. However, it is also "a teaching aid for the students". It accompanies the teacher's action in the classroom and prolongs it outside the classroom. It represents an interpretation of Official texts. Thus, it already belongs to *school knowledge*.

These analyses mainly use the overall structure of the documents (in particular chapter division), indexes, glossaries and keywords. They make it possible to highlight the following points:

- For *reference knowledge*, as stated in the curriculum and textbooks (which are the means of knowledge transfer), it is possible to use a general definition of the object "volcano". To use the example offered by Sigurdsson, *et al.* (2000, p.42): "a volcano is an exit point where magma can get to the surface in an intermittent or continuous manner". However the *typical* volcano does not exist. Nevertheless, there are various shapes of volcanic structure which depend on the types of volcanic activities. These types of activity can be explained within the theoretical framework of plate tectonics, which allows us to form a model of Earth's internal dynamics (Vosniadou & Brewer, 1992).
- For the *knowledge to be taught*, similar to findings by Bezzi & Happs (1994), Lillo (1994), Sharp, *et al.* (1995) and Dal (2005), volcanism can be studied in various manners: three angles of approach were defined, corresponding to each of the particular notional fields; which are:
    - The human angle where all that corresponds to the relationship between man and volcanoes can be found: positive or negative consequences, with their catastrophic aspects, emotional or sensory; methods of study of the phenomenon and the vulcanologists' work; applied geology (Bezzi & Happs, 1994; Dal, 2005);
    - The descriptive angle in which all the geological objects related to volcanism are studied, for example, the products of the eruption (Sharp, *et al.* 1995);
    - The explanatory angle where the phenomenon is linked to the structure of the globe, the related mechanisms and matter transformations (Lillo, 1994).

In the textbooks used in French elementary and middle schools, these angles are weighted differently, according to the curriculum level and according to the publishers: for example, the Magnard Year 9 textbook seems to give more weight to the descriptive aspects than the Nathan textbook.

If we take a closer look at the contents of the curriculum, the study of the evolution of the contents of the curriculum was based on Year 9 and 10 levels in France (B.O.E.N, 1978a-b, 1979a-b and 1985a-b). Indeed, geology, as a taught discipline, was introduced into elementary school only with the Official Instructions of 1985 (B.O.E.N, 1985a) and the curriculum has not yet undergone any modifications at this level. This study highlights an evolution of the structure of the *knowledge to be taught*.

Before 1985, the starting points of Year 9 studies were regional; the teacher based initial activity on the observation and the analysis of a local landscape (B.O.E.N, 1978b, 1979b). The whole of the program was based on the study of objects or external phenomena. It was only in Year 10 that studies reached a planetary scale, where a *structural* model *of the Globe* was proposed. Global tectonics was approached with care, simply to specify and synthesise extra-curricular information**:** only “the existence of slow movements of the continents” was confirmed and that it was possible to see “the hypothetical cause of internal dynamics activities in these movements” (B.O.E.N, 1978b, 1979b).

After 1985, on the other hand, the study of landscape and local observations became little more than a teaching aid which “allows the study of global tectonics” and puts the explanatory value into perspective and which must be “invested in the explanation” of studied geological phenomena. The theory is not disputed: “the discussion of global tectonics does not concern the Year 9 class” and a certain bigotry appears here: use of affirmative forms without qualifications: it “shows that sedimentary accumulations have drifted... lithosphere functioning should be represented as a relevant system able to explain the genesis and setting of rocks” (B.O.E.N, 1978b, 1979b). The theory of global tectonics became a central point; it structures the unit of the *knowledge to be taught*. The explanatory aspect, (or interpretation), took the lead over the more descriptive aspects, which had been prominent before 1985.

How Has the Geological Phenomenon Been Interpreted Since Antiquity?

To analyse the way in which volcanism has been represented over the course of previous centuries, it is necessary to turn to works on the history of Geology (e.g. Gohau, *et al.* 1991; Simkin & Lee, 1994; Sigurdsson, 1999; Stanleys, 2005). A review of these brings to light descriptions of volcanism which are, on the whole, precise and correct. Through a study of the explanations attributed to the phenomena, conceptions emerge which are presumed to be present in different eras, revealing contemporary, social and religious influences, as well as scientific techniques of the time (e.g. volcanism was compared to the effects of gun powder until the 17th century). Similar to Bezzi & Happs, (1994), Sharp, *et al.* (1995), Blake (2005) and Dal’s (2005) analysis in their articles, amongst the most common interpretations and those seen in the students’ answers are the following.

- The existence of a central fire, fuelling the volcanic spouts (from Plato, fourth century B.C. to Hutton, 18^{th} century, passing through Descartes’ Principles of Philosophy, 17^{th} century; “the Globe is fuelled by a central fire”).
- Formation of volcanoes by the lifting up and bursting of the Earth’s surface (from Albert the Great in the Middle Ages to the theory of Leopold von Buch in the 19th century)
- The likening of the volcanic structure to a “hollow mountain” (from Lucretius; “Etna is an entirely hollow mountain inside which an extremely violent wind circulates”; to Needham, 18^{th} century)
- Volcanic emission of fire which originates from combustion, reasonably deep down, fuelled by mineral or organic materials and stirred up,

depending on the individual case, by winds, the Sun's rays or reactions with water…(from Antiquity until the 18th century; "volcanoes are caused by coal fires which melt neighbouring rocks into lava flows", Werner).

The Importance of the Epistemological Analysis For an Ex Ante Analysis

The analysis of knowledge helped to establish an *ex ante* hypothesis, clarifying which notional fields and which angles of approach the questioning had to deal with, as well as putting several research hypotheses forward. Indeed, it is possible to think that the influence of *school knowledge* on the student's ideas can be seen throughout the curriculum in the following ways:

- the progressive acquisition of scientific knowledge (Driver, *et al.* 1994);
- the importance of explanation compared with that of description (Vosniadou & Brewer, 1992);
- the development of objectivity in relation to that of subjectivity, with possible persistence of alternative conceptions distinct from *reference knowledge* (Schnotz, Vosniadou & Carretero, 1999).

In methodological terms, this can be interpreted as the need to observe the way the students approach the subject: for instance, do they think firstly about human consequences of volcanism? What vocabulary do they use? What are they referring to when explaining the formation of volcanoes, or their activity? These are the main objectives of this study.

Methods

Context

Teaching Earth sciences in France begins with a brief introduction of the Earth's processes (Volcanism, Mountain Building, Weathering and Erosion, and Digenesis) in Year 6 (10 to 11 years of age). Then, the introductory concepts such as the Earth's materials (rocks and minerals), the Earth's structure (the Land Surface, underground and inside the Earth) and Earth's processes are taught in middle school in Year 9 (13 to 14 years old). Formal earth sciences lessons start in Year 10 (14 to 15 years old).

Teacher training is a 2 year course offered to high school graduates who are selected from a nationwide examination. Student teachers take modules in Science and Social Science during their course. Students do their practice teaching in mainstream schools in their final year.

The Sample

All the schools in which the research was done are situated in the North of France. The study, which included a total of 130 people, took place in: two Year 6 classes from an inner city elementary school, containing 22 to 23 students (between the ages of 10 and 11); two Year 9 groups, from an inner city middle school, containing a total of 32 students (between the ages of 13 and 14) and 29 first year student teachers and 24 second year student teachers from an inner city university education department. These four groups are referred to in the study as ST1 (student teachers in their first year), ST2 (student teachers in their second year), YR6 (Year 6 students) and YR9 (Year 9 students).

Each student is represented in three sets of data, four even, for Year 9, in the form of questionnaires handed out, before the subject of volcanism had been tackled in class. It is considered that the alternative conceptions which appeared in Year 6 mainly refer to extra-curricular learning; in Year 9, these conceptions should refer back to what was learnt in Year 6; for student teachers, conceptions should consist of an inventory of fixed knowledge on volcanism at an adult level.

Instruments and the Data-Collection Procedure

The methodology of the study comprises three research stages: collection of information by the association of ideas, a Q-Sort test and a questionnaire with open-ended questions.

The association of ideas made it possible to determine the angles in which students approach the subject of volcanism and to compare it with the angles defined during the analysis of knowledge. It also revealed a first alternative conception of volcanism, which is global and consistently represented. The technique, partly based on *the autonomous method* of *construction of knowledge* of Di Lorenzo (1991), consists of making the subject draw up a list of link-words in relation to the word *volcano*, given as an inductive term. It should be pointed out that this technique was used twice for Year 9: the first time in the case of a geology lesson and the second, roughly six weeks later, in a French language lesson concerning a written poetry activity.

The open-ended questionnaire and the Q-Sort test made it possible to pinpoint the alternative conceptions in relation to the geological concepts concerned (see Appendix).

The questions were written based on assumptions about the notional fields of Earth Science (Francis, 1993; Sigurdsson, *et al.* 2000; Frankel, 2005), and indications provided by the History of Geology (Gohau, *et al.* 1991; Simkin & Lee, 1994; Sigurdsson, 1999; Stanley, 2005), and finally on the results of the previous data collections (quoted in the related literature e.g. Happs, 1982, 1985; Bezzi & Happs 1994; Lillo 1994; Sharp, *et al.* 1995; Blake, 2005; Dal, 2005). The corresponding notional fields relate to:

- the products from the volcanic eruption;
- the location, formation and transformation of the lava;
- the nature of the materials constituting the volcanic structure;
- the make-up of this structure and its link with the Earth's make-up;
- its formation.

Children enjoy drawing and are able to use drawings to communicate the identifiable features of objects they have been asked to draw (Haynes, *et al.* 1994), although caution is needed when using drawings to represent children's understanding as what is not drawn does not necessarily imply the absence of these mental structures (Newton & Newton, 1992). Data from the drawings were cross-referenced with information from other sources to provide a more complete "picture" of what a child understands.

To examine how students visualise these concepts, questions relating to the last two points asked for the answer in the form of a drawing. This kind of technique has been used by many researchers in the related literature (e.g., Lillo, 1994; Blake, 2005; Dal, 2005).

The cross-referencing of the data obtained using these two types of questionnaires made it possible to refine interpretations already made. In effect, they are two kinds of methods, one where the subjects are in a situation of production, the other more directive, the Q-Sort test, where they have to use operations of selection. For the Q-Sort test, it is a question of choosing the correct solution/s, or the best amongst a range of proposed solutions: the case could be that the correct solution appears in the cognitive field of the student whereas it was not there at the beginning; in addition the answer needs no structuring on the student's behalf. In the questions of production, the process is different: the student needs to recognise or to determine the correct response from the elements of his own cognitive field and to structure it correctly.

The open-ended questionnaire was therefore given first, then the Q-Sort questionnaire. The second questionnaire builds on the answers to the open-ended questionnaire in order to help understanding the sense of the first answers given by checking their coherence and thus allowing the validity of our interpretations. It therefore enabled pinpointing of the alternative conceptions concerned.

Data Analysis and Results

Association of Ideas (lists of words)

The elements taken into account are firstly, the average number of words suggested by the students at each level, and secondly, the vocabulary used. Similar to findings by Bezzi & Happs (1994), Lillo (1994), Sharp, *et al.* (1995), Blake (2005) and Dal (2005), the latter was classed into five categories.

- the "geological objects", i.e. elements characterising, for example the shape of the volcanic structures, the products of the eruption…(e.g. lava, fire, mountain…) and corresponding to the descriptive angle (Bezzi & Happs, 1994; Sharp, *et al.*1995);
- the "adjectives" often used at the same time as one of the preceding "objects" or in direct connection with this category (e.g. *hollow* mountain, *liquid* fire…) (Blake, 2005);
- the "phenomena" (e.g. eruption, fusion) corresponding to the explanatory angle (Lillo, 1994);
- words related to the "emotional and aesthetic fields", also comprising some isolated adjectives but, evoking feelings or impressions (e.g. danger, to destroy, fantastic) and connected to the human angle (Bezzi & Happs, 1994; Dal, 2005);
- some "variants" (e.g. dinosaur).

The percentage of answers in each category was quantified and translated into a graphical representation to facilitate comparisons. It should be noted that qualifying adjectives, which were not directly associated to an "object", were classified in the "emotional-aesthetics" category.

It can be seen (Table 1) that certain terms are found at all levels of the curriculum and can be interpreted as alternative conceptions of volcanism, like a *destructive eruption of red hot lava*; as for "fire" and "mountain" (the mountain that spits fire), they are consistently found in YR6, YR9 and ST1.

***Table 1.* Elements of analysis in a list of words: The words most often seen are highlighted in bold.**

	YR6	YR9	ST1	ST2
Affective aesthetics	**Dead** **Killer** **Devastator** **Destructor** Red Hell Fear Noise	**Devastator** **Destruction** **Fear** **Death** **Red** Uncontrollable Flakes	**Destruction** **Catastrophe** **Red** Power/Strength Force Magic	(Very few responses) destruction Red
Objects	**Lava** **Fire** **Mountain** Ice cones	**Lava** **Mountain** **Fire** Rocks Flows Vent	**Lava** **Fire** **Rocks** **Stones** **Craters** Water Sparks Flames Clouds/Smoke/Steam	**Lava** **Rocks** **Craters** **Flows** **Magma** Lapillis Phenocrystal s Caldeira Rift Hotspot
Adjectives	**Hot Burning** Liquid	**Hot Burning** Viscous	(Very few) Hot	(Rare) Hot
Phenomena	**Eruption** **Earthquake**	**Eruption** **Earthquake** Melting Fissuring	**Eruption** Melting	**Eruption** **Melting** Rising terrestrial movements

It is possible to see on one hand, an increase of scientific knowledge from YR6 to ST2, in line with the level of the curriculum, and on the other hand, a correlative reduction of expression and subjectivity. The scientific angle of approach (objects, phenomena) is increasingly represented, especially the descriptive aspects (objects), whereas the explanatory angle is only well represented at the ST2 level (see Figure 1). Scientific teaching seems to have played its part in the acquisition of knowledge and a greater objectivity. The following results will make it possible to clarify this role.

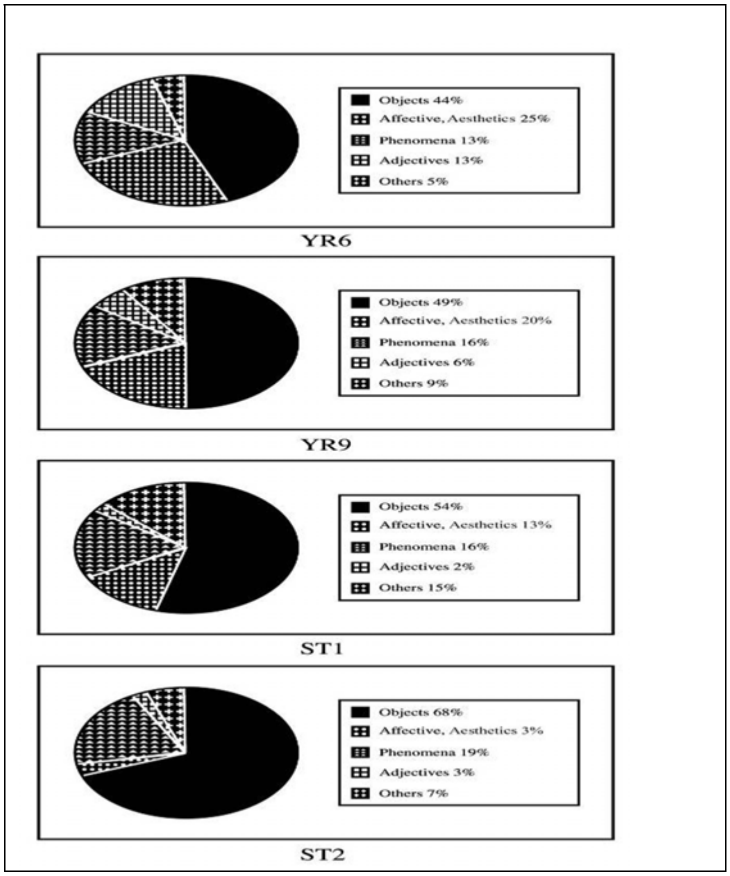

Figure 1. The angles of approach on different levels-The answers to the first questionnaire.

Open-Ended Questionnaire and the Q-Sort Test

Only the global result of the answers to the two types of questionnaires will be discussed here, taking into account the fact that the corresponding data was treated successively, the Q-Sort test having been conceived after the study of the answers to the first questionnaire. As for the data processing, the notional fields defined earlier constituted the reference grid. The analysis of the written answers was based on the description and the counting of "keywords" found in the students' conceptual levels (Bezzi & Happs, 1994; Sharp, *et al.* 1995; Blake, 2005; Dal, 2005).

The results obtained underlined the persistence of certain mental models: a volcano perceived like generally a cone-shaped relief, often hollow (Bezzi & Happs, 1994); a double origin of the volcanic structure: the up thrust mechanism given as important a role as (or even a more important role than) the accumulation of eruption products (Blake, 2005); the origin of the lava at the centre of the Earth (Lillo, 1994; Dal, 2005); short time scale in describing the products from volcanic eruption and the nature of materials constituting the volcanic structures (Ault, 1982).

The findings indicated that both student teachers and Year 6 and Year 9 students had some similar alternative conceptions (see Table 2). However, over the course of the curriculum, a closer relationship between alternative conceptions and accepted scientific knowledge was evident. Indeed, similar to Bezzi & Happs (1994), Lillo (1994), Sharp, *et* *al.* (1995), Blake's (2005) and Dal's (2005) analysis in their articles, seven categories (or types) of alternative conceptions were identified from idea generated answers (drawings and written responses). These categories were numbered from 1 to 7. Types 1 to 3 correspond to the more elementary conceptual levels, with a "leap" of knowledge in the third since it is at this level that the students began to allot a deep point of origin to lava; types 4 to 7 represent those with an increasingly elaborate knowledge or more scientific terminology. Examples of the students' drawings related to the various types are given in Figure 2.

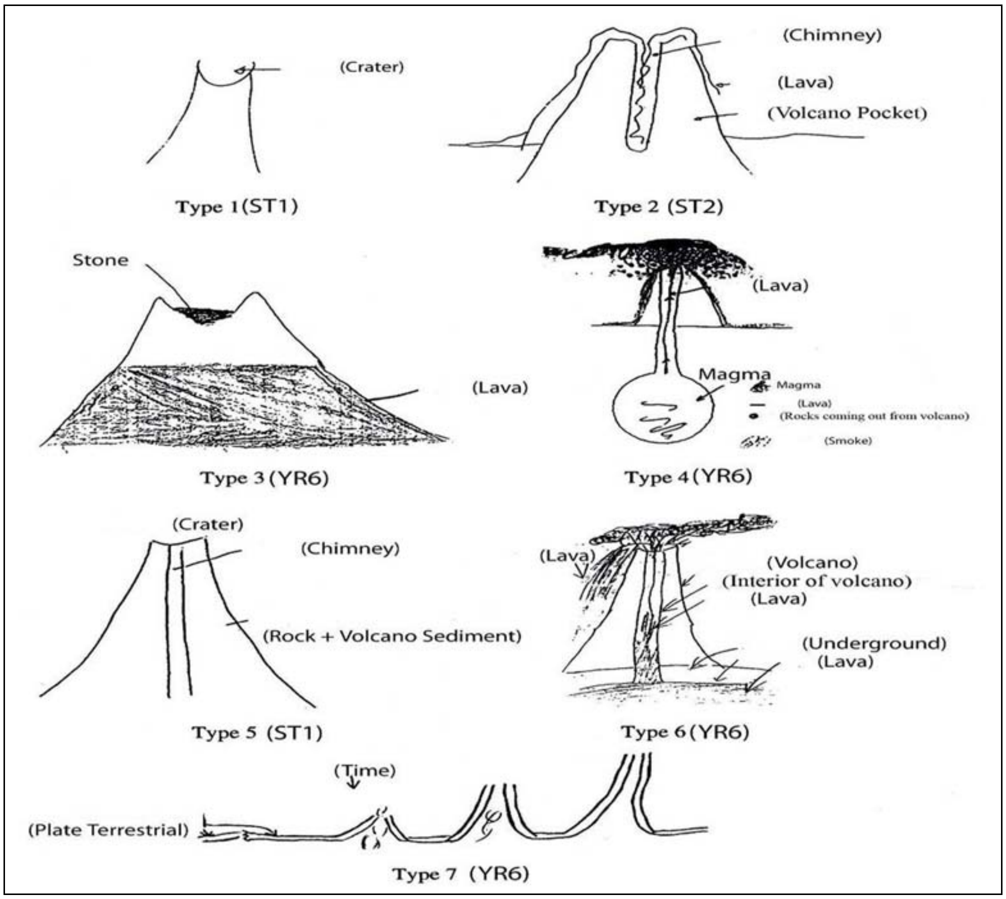

Figure 2. Some examples of the students' drawings.

The defined categories are the following:
Type 1: a volcanic cone without structural detail, including diverse and unclear answers concerning the origin of lava, even the absence of answer; no mention of time scales (16 %) (Blake, 2005).
Type 2: a well demarcated cone, with an internal structure; the students producing this type of drawing give the lava a superficial origin related to the volcano ("interior of the volcano", "at the top of the volcano"); no mention of time scales and specific mention of short time scales (years or less) (36 %) (Sharp, *et al.* 1995; Dal, 2005).

Type 3: a cone filled with lava, associated to a major (but vague) origin including "the lava is under the ground and goes up to fill the volcano"; specific mention of short time scales (years or less) (Sharp, *et al.* 1995)
Type 4: the presence of a shaft linking the volcanic structure with a reservoir, deep point of origin of lava; specific mention of short time scales (years or less) and general terms of time used (36 %) (many years).
Type 5: a structure with an unfinished shaft; deep point of origin of lava; specific mention of short time scales and general terms of time used (many years) (Lillo, 1994).
Type 6: a shaft linking the volcano with a layer or lava cap; deep origin but often evoking "the Earth's crust", the "basement", and the "the Earth's internal layers"; general terms of time used (many years), specific mention of medium time scales (thousand) (9 %) and specific mention of long time scales (millions) (3 %) (Bezzi & Happs, 1994; Dal, 2005).
Type 7: volcanic structure comprised of two "terrestrial plates" which clash; deep origin of lava; specific mention of long time scales (millions) (Blake, 2005).
These alternatives conceptions are distributed as follows, according to curriculum level (see table 2).

***Table 2.* Distribution of alternative conceptions.**

	YR9	YR6	ST1	ST2
Type 1	7	1	2	0
Type 2	7	3	1	1
Type 3	11	4	0	0
Type 4	8	10	7	9
Type 5	5	9	6	2
Type 6	3	6	4	16
Type 7	1	0	0	2
Total	42	33	20	30

The simplest types are more numerous at Year 6 level and decrease, and even disappear over the course of the curriculum, whereas, the most elaborate types characterize students at the most scientific level.

Thus it was possible to map out the categories of alternative conceptions of volcanism, in particular concerning the formation of the volcanic structure, the formation of lava and the understanding of geological time of the volcanic activities and to measure the influence of the curriculum through looking at the evolution of these alternative conceptions. It seems that, at the same time as the information they assimilate in their studies, similar to findings by Bezzi & Happs (1994), the students also acquire a specialised vocabulary (e.g. rock, magma) and that they all build a certain knowledge base, as the answers to the Q-Sort test confirm. They begin to link volcanism with the Earth's internal activity, understanding that: it is lava and not "fire" which comes out of a volcano during an eruption and this lava "transforms into rock" (see Figure 3). Descriptive approaches are, however, better represented than the explanatory approaches and appeared earlier in YR9, ST1 and ST2 as had been envisaged.

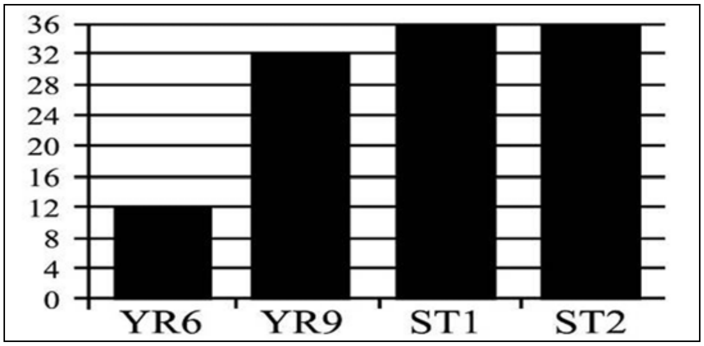

Figure 3. Example of the distribution of the answers to the Q-sort test: question 10: When lava comes out of a volcano, it cools down and transforms into rock?

The numbers express the “weight” of the given responses at each level for the same number of individuals, i.e. the score obtained by the addition of the points attributed to the questions and the total number of responses given; the possible choices went from -2 (completely disagree) to +2 (completely agree), (see questionnaire in the appendix).

Discussion: Do the Alternative Conceptions Listed Shed Light on Barriers to Learning?

Similar to findings by Driver, *et al.* (1994), it was noted that even though the alternative conceptions which are the most distinct from reference knowledge decrease as the amount of scientific teaching received, they do not disappear completely.

Therefore, taking only a few examples, the majority of students, including student teachers, consider that lava comes from the centre of the Earth, a place of mysterious and fantastic activity, an idea which dates back to Antiquity. For them, as at the time of Von Busch, a 19th century geologist, the volcanic structure is primarily formed by rising, under the influence of an internal push, “like a button” (Sigurdsson, 1999).

In addition, it is difficult to imagine the products from the volcanic eruption in another form (e.g. lava transforms into rock) (Sharp, *et al.* 1995; Blake, 2005). Furthermore, today we find them on the surface of the Earth, thus in thermodynamic conditions different from those within which they are formed. Moreover, the actual surface also creates phenomena that the students have not seen occur (Trend, 1998). Another reality that reinforces this complication is the fact that the transformations that they know happen over a short period of time, whereas they don’t see the products from the volcanic eruption and the nature of the materials constituting the volcanic structure are the consequences of a slow processes (Ault, 1982; Happs, 1982; Blake, 2005).

Even if some children were not taught the relevant information, the conceptions of some adults, whatever they were taught, are sometimes inspired by *imagination, personal perceptions or analogical thought process* (Blake, 2005).

Similar to the relevant literature, the latter may be explained by certain characteristic difficulties of geology: particularly the difficulty of reproducing geological phenomena in experiments and of obtaining direct observations; or even the problem of understanding time and space (Trend, 1998).

Concerning the first point, available testimonies are generally fragmentary and are spread both over different historical periods and geographically: it is necessary to put the "puzzle" back together (Ault, 1982). In addition, especially in the classroom, the use of significant experiments and modelling which could be used as a substitute for reality and/or as conceptual aids, is difficult and generally only approximate.

Concerning the second point, the duration of geological phenomena is usually measured in millions of years and the notion of geological time is difficult to grasp, even for an adolescent or an adult. Ault (1982) and Trend (1998) have shown that, for students between the ages of 13 and 14, the most common barriers to learning concern the incapacity to evaluate the chronological succession, the duration of events and the concept of terrestrial dynamic. Sharp, *et al.* (1995) and Blake's (2005) works, already mentioned, also develop the discussion of difficulties of assimilation of the concept of geological time.

Thus, much attention is devoted to activities which illustrate the immensity of deep time. Moreover, the objects of geological knowledge vary dramatically in scale, from molecule (e.g. crystal lattice) to universe (e.g. planetology, astronomy), but it is often on a continental or planetary scale that they have to be represented (e.g. orogenesis is treated on a planetary scale. i.e. the birth of a chain of mountains, within the framework of the theory of plate tectonics).

Volcanism was partly chosen as it was thought that its difficulties could be tackled, volcanoes are, at least partially, accessible to observation, directly or thanks to the media, and their most visible activities can be studied on a human scale in time and space. Indeed, as Scarth (1999) says

> an Earth that has two speeds does not exist! The short duration activities are only transitory signs of a phenomenon of very long duration. This is where the difficulty of the study of earthquakes and volcanoes. It is necessary to find methods to connect completely disproportionate time scales. (p.115)

Thus, it seems that there are specific epistemological barriers to learning in many disciplines of Earth sciences.

However, the alternative conception analysis has shown that this is not the only type of barrier to learning that must be taken into account. Indeed, certain alternative conceptions, distinct from scientific knowledge seem to develop parallel to the teaching received: thus it is Year 9 students, who had studied volcanism in Year 6 and the ST2 who are the most common groups to consider that "lava passes between the plates of the Earth's crust", as the answers to the Q-Sort test confirm. Similar to findings by Sharp, *et al.* (1995) and Vosniadou & Brewer (1992), these students know about the existence of the lithospheric plates, but the mental model that they build gives rise to false interpretations of the observed phenomena.

The epistemological barriers to learning thus increase the didactic barriers to learning. These barriers to learning seem to be linked to the *didactic transposition* process, a concept that is attributed to Chevallard (1991). Several points of reflection can thus be considered in relation to elements likely to play a part in the persistence or the appearance of alternative conceptions distinct from the *reference knowledge*.

The first of these elements is the operation of the education system. Indeed, time devoted to each point of the curriculum is limited; this leads to a pace of teaching which takes into account neither the student's ideas nor the time required for learning. As Driver, *et al.* (1994, p.7) points out: "the education system tends to […] want to identify teaching time and learning time, and to treat in terms of educational failure or delay, any difference between these two rhythms". In addition, the choices which are made concerning the concepts to be taught, and the relations which organise them, lead to an important degree of abstraction, prioritising the mechanisms and interpretations even though students cannot yet conceive the relevant object of observation. This may explain the alternative conceptions which describe a concertinaing relationship between the dimensions of the plates and the magma hot spots (see drawing of type 7 of Figure 2).

A second element is the nature of the explanations that accompany diagrams in textbooks, and the way in which they are used by the teacher in the classroom. To quickly illustrate this point, it is necessary to compare examples of persistent alternative conceptions with the explanations used in textbook diagrams.

The first example is that of the existence of a layer of magma under the Earth's crust whereas the mantle is in fact in a solid state. In Year 6, Bordas Publishing's Sciences and Technology textbook explicitly confirms this interpretation by proposing it like an "explanatory model" of the structure of the Globe. The explanatory model, concerning what is under the lithosphere, suggests that:

> Rigid shell = The Lithosphere, Melting rocks = The Magma, Rigid materials = The Central Core, Zone of compression: Formation of a chain of mountains, Zone of separation: Formation of an ocean (Tavernier, 1987, p.12).

In Nathan Publishing's Year 9 Geology textbook, to help the students understand the convection currents in the mantle, a model is proposed, in the form of diagrams with accompanying notes, of water (thus a liquid) heated by an electric resistance. It is explained that:

> Reaction of water in a kettle: Water heated by resistance rises towards the surface. In contact with colder water, the hotter water is cooled, falls back to bottom and reheats… Such a phenomenon is called convection. It evacuates heat and creates zones at the surface of the container, some hotter than others (Périlleux & Thomas, 1988 p.53).

The solid state of the mantle in fact only indicated by two words, right at the bottom of a whole page of illustrations. It is written that:

> Convection on a global scale: The differences in temperature between the interior of the globe and the surface create movements, at the solid state, of rocks in the mantle. These movements allow the Globe's internal heat to escape (Périlleux & Thomas, 1988 p.53).

The second example is that of the formation of a volcanic cone by rising, without the products of the volcanic eruption playing any part. In Bordas Publishing's Year 6 textbook, an illustration likely to confuse children can be found. At the bottom of a whole page of illustrations concerning the formation of a volcanic cone, it states the following:

> The volcano Paricutin is found in Mexico. It appeared in 1943. Its growth was very fast. It stayed active for 9 years and it ejected 3.6 billions tons of material (Tavernier,1987, p.46).

Similarly, the following is found in the key for the "Forecast of Volcanic Eruptions" diagram in Magnard Publishing's Year 9 textbook. The diagram is accompanied by the following descriptions:

> "Period 1 the swelling of the volcano starts, Period 2 peak of swelling – the horizontal and vertical distances lengthen - the slope increases", Period 3 Eruption- Deflation – "the slope decreases" (Salviat & Desbeaux, 1988, p.55).

However no scale is given to indicate the relative value of this swelling and to situate the phenomenon in relation to the mouth of the volcano.

These explications, which go together with the diagrams, aim to facilitate the comprehension of relatively inaccessible phenomena thanks to simplifications or analogies; but in both cases, if the teacher is not careful, the alternative conceptions can be consolidated or generated by teaching. They may create barriers to learning of a didactic origin.

Conclusion

The analysis of the responses showed that the alternative conceptions defined in biology, physics, chemistry and more generally in science education, are also found in Earth Sciences. Thus, it is possible to establish categories of alternative conceptions, to identify barriers to learning, and to examine the role played by the didactic transposition.

The findings indicated that both student teachers and Year 6 and Year 9 students had some similar alternative conceptions. Based on the findings, we concluded that since student teachers had alternative conceptions similar to those held by the Year 6 and Year 9 students, any instruction that student teachers had received from Year 9 onward had little effect on their alternative conceptions. Furthermore, since the teachers are a prime source of instruction in the study context, their alternative conceptions can be easily transferred to their students. Therefore, if we take the alternative conceptions into account in planning future activities in teacher education and science curriculum development, students may have a better chance to scientifically develop the fundamental concepts of science. Since these concepts are building blocks for latter learning, their development will help students to meaningfully grasp the advance concepts of science.

In the science education literature, it is stated that alternative conceptions are resistant to change. Teachers affect their students mostly in instructional environments. If the instructional environment is not well designed, conceptual learning and conceptual change will be difficult to achieve. In addition, the classroom environment and teaching strategies should create enough dissatisfaction with existing knowledge so that students' alternative conceptions may change. Therefore, we should not expect student teachers to develop high quality materials that provide conceptual learning or conceptual change or completely intervene in the students' alternative conceptions.

Furthermore, science-textbook writers and teachers do not adequately take into account the students' previous learning, particularly their alternative conceptions. In other words, they expect that students already understand the underlying concepts that are prerequisite for further or advanced learning. When students had inadequate concepts, the study teachers generally blamed the teachers of earlier stages of schooling, and they claimed that the earlier teachers did not teach the concepts at an appropriate level. In fact, France has a centralized educational system, and all schools implement the same curricula. Different authors write the textbooks by taking into consideration the curriculum for each subject area. The science teachers in each school are free to choose one of the science textbooks for their teaching. As is the case in many other countries, teachers can prepare supportive teaching materials for their students. Therefore, this study can help current science teachers rethink their way of teaching as well.

Hence, science teacher educators should develop appropriate activities and strategies to introduce students' alternative conceptions to student teachers and also address the student teachers' own alternative conceptions. Thus, as science teacher educators, we can eliminate one of the most important sources of alternative conceptions. This approach can provide experiences in an actual classroom environment about how alternative conceptions are dealt with and remediated from a professional. Moreover, student teachers will become aware of the importance of prior knowledge and will be able to develop activities and build new knowledge during their teaching. Furthermore, current science teachers in schools should be provided with inservice training so that they can relinquish their alternative conceptions.

Schnotz, *et al.* (1999) noted that student conceptual understanding becomes clear when multiple-source data are interpreted from multiple theoretical perspectives. At that point, it may be concluded that we should take into account individual differences. That is, we should identify students' perspectives and then devise guide materials, which should be designed to facilitate conceptual learning. Moreover, inservice and preservice programs should be reorganized for better teacher preparation.

Given the aforementioned conclusion, although certain aspects are prevalent only in specific disciplines, such as problems linked to time and space and the problems of modelling, results of research in didactic, and particularly barriers to learning and their origin, should improve understanding of learning phenomena in this field. They can also aid adapting Earth Science teaching to knowledge transfer methods so that Earth Science concepts are accessible, whilst still remaining accurate, at all levels of the curriculum.

In the science education literature, learning as conceptualised in a constructivist framework is dependent on authentic experiences that help learners make these critical connections between scientifically rigorous concepts and their own sense-making (Vosniadou & Brewer, 1992). This is a particularly challenging task when thinking about geoscience education, however, as student and teachers alike generally have had fewer experiences with formal geoscience education. To what extent do children make connections between everyday encounters with geology and their formal classroom learning? In the country in which the research was done, there aren't any volcano sites visible to students. It is important to note that this study didn't try to search an answer to this question. This study explored the students' ideas of the volcanoes and volcanic activities as the focus for conceptual change. A further research that could be done in the Clermont Ferrand region in central France, where there are important volcanoes sites (*Puy-de-Dôme*) visible would report on the extent to which students make connections between their ideas about rocks learned in school and their experiences with volcanoes in their out-of-school lives.

References

Ault, C.R.Jr. (1982). Time in geological explanation as perceived by elementary school students. *Journal of Geological Education,* 30, 304-309.

Bezzi, A., & Happs, J.C. (1994). Belief systems as barriers to learning in geological education. *Journal of Geological Education,* 42, 134-140.

Blake, A. (2005). Do you children's ideas about the Earth's structure and processes reveal underlying patterns of descriptive and causal understanding in earth science. *Research in Science Technological Education,* 23(1), 59-74.

Chevallard, Y. (1991*). La transposition didactique.* Grenoble : La Pensée Sauvage.

Dal, B. (2005). The initial concept of fifth graduate Turkish's students related to earthquakes, *European Journal of Geography*, 326.

Di Lorenzo, G. (1991). *Questions de savoir.* Paris: ESF, Coll. Pédagogies.

Dove, J.E. (1998). Students' alternative conceptions in Earth science: A review of research and implications for teaching and learning. *Research Papers in Education,* 13, 183-201.

Driver, R., Guesne, E., & Tiberghien, A. (eds). (1985). *Children's Ideas in Science.* Buckingham: Open University Press.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher*, 23, 5-12.

Ford, D.J. (2005). The challenges of observing geologically: third graders' descriptions of rock and mineral properties. *Science Education,* 89, 276-295.

Ford, D.J. (2006). Representations of science within children's trade books. *Journal of Research in Science Teaching*, 43 (2), 214-235.

Francis, P.W. (1993). *Volcanoes, a Planetary Perspective.* Oxford: Clarendon Press.

Frankel, C. (2005). *Worlds on Fire.* Cambridge: Cambridge University Press.

Giordan, A., & De Vecchi, G. (1987). *Les origins du savoir.* Paris: Delachaux & Niestlé.

Gohau, G., Carozzi, C., & Carozzi, A. (1991). *A History of Geology.* London: Rutgers University Press.

Happs, J.C. (1982). Rocks and minerals. Science Education Research Unit Working Paper No. 204. New Zealand: University of Waikato.

Happs, J.C. (1985). Regression on learning outcomes: some examples from the Earth Sciences. *European Journal of Science Education,* 7(4), 431-443.

Haynes, D., Symington, D., & Martin, M. (1994). Drawing during science activity in the primary school. *International Journal of Science Education,* 16 (3), 265-277.

Lillo, J. (1994). An analysis of the annotated drawings of the internal structure of the Earth made by students aged 10-15 from primary and secondary schools in Spain. *Teaching Earth Sciences,* 19(3), 83-87.

Newton, D.P., & Newton, L.D. (1992). Young children's perception of science and the scientist. *International Journal of Science Education,* 14(3), 331-348.

Osborne, R., & Freyberg, P. (1985). *Learning in Science: The Implications of Children's Science.* London: Heinemann.

Périlleux, E., & Thomas, P. (1988). *Biologie Géologie 4.* Paris: Nathan.

Russell, T., Bell, D., Longden, K., & McGigan, L. (1993). *Primary SPACE Research Report: Rocks, Soil and Weather.* Liverpool: Liverpool University Press.

Salviat, N., & Desbeaux, B. (1988). *Sciences et techniques biologiques et géologiques 4.* Paris: Magnard.

Scarth, A. (1999). *Vulcan's Fury: Man against the Volcano.* New Haven, Conn: Yale University Press.

Schnotz, W., Vosniadou, S., & Carretero, M. (eds). (1999). *New perspectives on conceptual change.* London: Pergamon.

Sharp, J.G., Mackintosh, M.A.P., & Seedhouse, P. (1995). Some comments on children's ideas about Earth Structure, volcanoes, earthquakes and plates. *Teaching Earth Sciences,* 20(1), 28-30.

Sigurdsson, H., Bruce, F. H, Stephen, R.M., John, S., & Haxel, R. (2000). *Encyclopedia of Volcanoes.* New York: Academic Press.

Sigurdsson, H. (1999). *Melting the Earth: The history of ideas on Volcanics Eruptions.* Oxford: Oxford University Press.

Simkin, T., & Lee, S. (1994). *Volcanoes of the World: A regional Directory, Gazetteer and Chronology of Volcanism during the last 10,000 years. 2d (ed).* Tucson, Ariz: Geoscience Press.

Stanleys, S.M. (2005). *Earth System History (2nd edn).* New York: W.H.Freeman & Co.

Stern, L., & Roseman, J.E. (2004). Can middle-school science textbooks help students learn important ideas? Findings from project 2061's curriculum evaluation study: Life science. *Journal of Research in Science Teaching,* 41 (6), 538-568.

Tavernier, R. (1987). *Sciences et Technologies, CM.* Paris: Bordas.

Trend, R. (1998). An investigation into understanding of geological time among 10-and 11 years old children. *International Journal of Science Education,* 20(8), 973-988.

Vosniadou, S., & Brewer, W.F. (1992). Mental models of the Earth: a study of conceptual change in childhood. *Cognitive Psychology*, 24, 535-585.

Official Instructions :

B.O.E.N - 1978a, Programmes du cours moyen.
B.O.E.N - 1979a, Programmes du cours moyen.
B.O.E.N - 1985a, Programmes des du cours moyen.
B.O.E.N - 1978b, Programmes des 4ème des collèges.
B.O.E.N - 1979b, Programmes des 4ème des collèges.
B.O.E.N - 1985b, Programmes des 4ème des collèges.
B.O.E.N – 2000a, Programmes du cours moyen.
B.O.E.N - 2000b, Programmes des 4ème des collèges.

Appendix

The Open-ended questionnaire and the Q-Sort Test were presented in the following manner:

The Open-ended questionnaire:
Dear student,
1. What comes out of a volcanic eruption?
2. In your opinion, where does lava come from?
3. How is lava formed?
4. What does lava become after a volcanic eruption?
5. In your opinion, are volcanoes still active?
6. What materials are volcanoes made of?
7. Please draw a diagram, with notes, of the longitudinal section of a volcano on the other side of this sheet of paper.
8. Are they several types of volcanoes?
9. Could you explain how you imagine different phases of the formation of a volcano (you may draw them)?

The Q-Sort Test:
Dear Student,
Please check the box that corresponds to your opinion (-2 completely disagree; -1: mostly disagree; +1 mostly agree; +2 (completely agree)

	- 2	-1	+1	+2	comments
1. Volcanoes have a cone shape?					--------------

2. A volcano is a mountain which opens to release lava?

3. Lava forms a continuous layer under the Earth's crust?
4. Volcanoes are very tall?
5. The visible part of a volcano is only formed with volcanic rocks?
6. It is fire that comes out of volcanoes?
7. Volcanoes are very tall because they were formed through uplift in the Earth's crust?
8. Lava comes from the fusion of rocks constituting a volcano?
9. Volcanoes are hollow?
10. When lava comes out of a volcano, it cools down and transforms into rock?
11. The visible part of a volcano is always formed with the same rocks as the rocks that form the Earth's crust?
12. Lava moves between plates in the Earth's crust?
13. There are volcanoes in France?
14. Lava comes from the Earth's centre?
15. It is solar energy that causes volcanic eruptions?
16. Volcanoes form a relief because they are formed through the accumulation of volcanic rocks.
17. The birth of a volcano is linked to (an earthquake)?
18. There are volcanoes under the sea?

Investigating the Effects of an Aquatic Ecology Graduate Course for Teachers: Linking Teaching to the Environment and Community

Yvonne Meichtry
Northern Kentucky University

Jeff Smith
Northern Kentucky University

Abstract

The purposes of this study were to determine the impact of a graduate teacher education course on the confidence levels and classroom practices of teachers**.** The three-credit hour, field-based course was taught during the summer using a two-week workshop and one follow-up day format. Place-based teaching approaches were utilized during the course. These approaches were designed to immerse teachers in studies of their local aquatic environment and community-based resources that are associated with the aquatic environment. Pre, post, and delayed post-survey data were analyzed using MANOVA and ANOVA measures to determine changes in the teachers' confidence levels and classroom practices. Positive changes were found in the teachers' confidence and classroom teaching in the use of various instructional technology, standards-based teaching strategies, community resources, field investigations, and in the teaching of water quality topics, real life topics, societal issues, and career education. An analysis of responses to open-ended questions on the delayed post-survey revealed the strengths of the course in regard to the learning of science content, instructional pedagogy and applications to classroom teaching, the potential impact on K-12 student learning, and barriers to implementing desired classroom practices. Implications and recommendations are presented that can be generalized across a variety of educational programs.

Correspondence should be addressed to Yvonne Meichtry, Northern Kentucky University, Email: meichtryy@nku.edu or Jeff Smith, Northern Kentucky University, Email: smithj@nku.edu

Introduction

Place-based educational approaches are designed to develop a sense of connectedness to where one lives by grounding learning in the local natural and community-based environment. The characteristics of place-based learning that make it a distinctive approach to education, as summarized by Woodhouse and Knapp (2000), are the emergence of education from the particular geography, ecology, sociology, & politics of a local community, a focus of study that is inherently multidisciplinary and experiential, and the connection of place with self and community.

Place-based education is an approach to education that is aligned with the goal of improving K-12 educational outcomes, as evidenced by an increasing number of studies reported in the literature. Among the educational benefits of place-based education evidenced in K-12 schools are an improved performance on standardized tests in all academic subjects, a reduction of discipline problems and absenteeism, an increase in engagement and enthusiasm for learning, greater pride in accomplishments, and greater teacher job satisfaction (Athman & Monroe, 2004; Audubon Washington, 2004; Ernst & Monroe, 2004; Heimlich, 2002; Lieberman & Hoody, 1998; NEETF, 2000; Powers; 2004, SEER, 2000; Smith, 2002; Sobel, 2005).

In addition to the educational value of place-based education, this approach to learning is viewed as a potential means of sustaining the culture and natural environment (Orr, 1994; Smith & Williams, 1999). According to David Orr (1994), people must have knowledge of ecological patterns, systems of causation, and the long-term effects of human actions on those patterns if they are to work on behalf of sustaining the cultural and ecological integrity of the places they inhabit.

While the value of place-based education is gaining increased recognition, it is not yet a mainstream approach used to design K-12 curricula, nor is this approach yet a substantial part of teacher education. For schools to successfully implement place-based teaching approaches, it is critical to provide teachers with training in the use of teaching practices that many have not experienced themselves.

Recent studies conducted with K-12 teachers in Kentucky revealed specific environmental science education needs of teachers that are relevant in building a foundation upon which teachers can effectively utilize place-based educational approaches. In a statewide survey conducted by Doug Carr (2005), 67% of teachers reported incorporating environmental content in their teaching, but relatively few incorporated it extensively. Few teachers received training related to environmental content within the past 3 years, but those that did appeared more likely to teach environmental content. The most important reasons identified for teaching environmental content were the relevance of the environment to the everyday lives of students and to teach their students about current issues. One of the most frequent reasons given for not teaching about the environment in this study was the lack of teaching materials and lesson ideas. The results of an environmental education needs assessment of K-12 teachers conducted by Meichtry and Harrell (2002) indicated that the three greatest needs of teachers, in order of frequency, were training in the use of outdoor learning sites, training in the alignment of curriculum with state standards, and the availability and use of curricula.

The Graduate Course

The graduate course used as the focus for this study was a three-credit hour course offered in the summer. The course was taught using a two-week workshop format and one day follow-up session. An overview of the course schedule is presented in Figure 1.

Figure 1: Overview of Graduate Course Schedule

<u>Week 1: Day 1</u>

1. Pre-course survey
2. Project WET activity – Humpty Dumpty (Restoration of aquatic systems)
3. Course overview
4. Overview of Field Station and programs

5. Video (After the Storm, EPA)
6. KY Watershed Watch overview/KY Watershed basins
7. Enviroscape Demonstration
8. Small group activity – Identify state standards addressed by lesson & develop open-response assessment for K-12 students.

<u>Week 1: Day 2</u>
Discuss previous day state standards & assessment questions
Set up river productivity line

Microbiology and Water Chemistry
1. Introduction to river ecosystem and sampling methods
2. Sampling of river-productivity line, collect plankton (pontoon boat) and water samples for chemistry testing, demonstrate use of YSI SONDE to instantly collect & graph multiple river parameters, & collect coliform samples
3. Measure oxygen levels in productivity bottles & calculate river productivity
4. Use microscopes and keys to identify microscopic life
5. Discuss impact of water parameters on biodiversity of the river

Identify state standards & open-response assessment for lesson.

<u>Week 1: Day 3</u>
Discuss previous day state standards & assessment questions
Observe & discuss coliform results

Geology and Chemical Cycling and Geologic History of Ohio River
1. Geologic History of Ohio River
2. Review of the hydrologic cycle
3. River systems & flood plain development
4. How various constituents can enter a river system
5. Collect water samples (from the Ohio River at field station and upstream from a tributary)
6. Field station: Analyze samples and plot data in histograms using Excel
7. Discuss results
8. Hypothesize concentrations in downstream tributary

Identify state standards & open-response assessment for lesson.

<u>Week 1: Day 4</u>
Discuss previous day state standards & assessment questions
- Demonstrate and practice use of LabPro and Dana technology to collect water parameter data

Field Trip to Ohio River Tributary: Stream Survey
- Biological Index
- Habitat Assessment
- Chemical and physical water quality parameters
- Fish seine

- Plankton & coliform sample
- Stream velocity

Identify state standards & open-response assessment for lesson.

<u>Week 1: Day 5</u>
Morning:
Discuss previous day state standards & assessment questions

- Microscope study of algae, protista, & microscopic invertebrates and study of macroinvertebrates collected in the Ohio River on Wednesday and in the Ohio River Tributary on Wednesday. Compare samples.
- Mussels of the Ohio River

Afternoon:

- 1:00-1:30-Speaker, Sierra Club Water Sentinels Program
- 1:30-2:00-Speaker, Ohio River Sanitation Commission (ORSANCO)
- Classroom curriculum activities

Identify state standards & open-response assessment for lesson.

<u>Week 2: Day 6</u>
Discuss previous day state standards & assessment questions

Terrestrial, Wetlands, & Upland Ecosystem and Biodiversity Study

- Wetlands and floodplains orientations (St. Ann's)
- Vegetation monitoring/research methods
- <u>Data interpretation/forest evaluation</u>
- Upland: Exotic species and their effects on ecosystems
- Watersheds/storm water management
- Calculate coefficient of similarity
- Habitat restoration (theory & practice)

Identify state standards & open-response assessment for lesson.

<u>Week 2: Day 7</u>
Morning:
Discuss previous day state standards & assessment questions
Field Trip: Lafarge Gypsum Plant – role of industry in protecting aquatic systems and biodiversity; education efforts for schools and community members

Afternoon: Field trip: Sanitation District #1 – best management practices for storm water runoff and education program and facilities.

Identify state standards & open-response assessment for lesson.

<u>Week 2: Day 8 - Evening session 5:00-10:00 PM</u>
Discuss previous day state standards & assessment questions

- Electrofish and aquatic organism study, fishes of the Ohio River
- Water quality parameters

Develop core content for assessment & open-response assessment for Wednesday lesson

Week 2: Day 9
Field Trip: Licking River study:
Canoe trip and field study
- Stream monitoring – chemistry, habitat assessment, and macroinvertebrate
- Canoe safety and paddling techniques
- Drainage patterns in watersheds
- Flooding and water management issues
- Point & nonpoint source pollution issues & best management practices
- Enjoy the river!

Identify state standards & open-response assessment for lesson.

Week 2: Day 10
Morning:
Discuss previous 2 days state standards & assessment questions
10:00- Speaker, Conservation Districts
10:40-Speaker, KY Energy Education Development Project (NEED)
11:20-Speaker, Ohio River Foundation
Identify state standards & open-response assessment for lesson.

Afternoon:
Individual Work time on curriculum projects
- Discuss previous day core content & assessment questions
- Professional river-based and education organizations
- Kentucky Watershed Watch and Licking River Watershed Watch
- Curriculum Resources

Follow-Up Session: 2 Weeks After Completion of 2-Week Segment of Course
- Teacher presentations of course projects
- Work on KAEE presentation
- Projects and notebooks due
- Post-course survey

Student Enrollment and Course Instruction

Students enrolled in the course were K-12 teachers seeking a Masters degree or Rank 1 certification. The course is cross-listed in the departments of Education and Biology and can be applied as a science content course requirement, an elective course, and/or as one of four courses that apply to an Environmental Education Endorsement.

The course was co-taught by two faculty; one with expertise in biology and environmental science and the other with expertise in science education and environmental education. Both instructors were trained and experienced in the use of the place-based teaching strategies that were used throughout the course. These strategies included experiential learning, use of the environment and community as a focus to integrate disciplines, inquiry-based learning, relevancy of learning to real life and current societal issues, and student reflection. Guest instructors during the course were university professors who specialized in the content areas of geology, microbiology, and botany/ecology.

Course Description

This course was designed to incorporate place-based teaching approaches as a means to improve K-12 science education outcomes and to address the needs of practicing teachers in the field of environmental science education. During the two-week segment of the course, teachers were engaged in field-based studies of aquatic systems, field trips to community facilities, presentations made by community-based guest speakers, classroom discussions to reflect on what they were learning, and small group work to apply what they were learning to their K-12 classrooms. Topics of the field studies used to investigate the local aquatic environment were microbiology, water chemistry, geology and chemical cycling, geologic history of the Ohio River and watershed area, stream survey components (macroinvertebrate sampling, habitat assessment, chemical parameters, stream flow, and coliform and plankton sampling), terrestrial, wetlands and upland ecosystems, and fishes of the Ohio River.

All field-based studies were inquiry-based. Teachers were required to keep a notebook record of all investigations made during the course. The standard format used to conduct investigations was the development of a question and a hypothesis, conducting the procedure to test the hypothesis, recording and analyzing results, and drawing conclusions.

The course utilized the monitoring protocols and scientific equipment used by the Kentucky Watershed Watch (Kentucky Division of Water, 2000a-c), which consisted of Lamotte dissolved oxygen and pH test kits, an aquatic thermometer, and a conductivity meter (Lamotte, 2006). Using state-recognized protocols made it possible for the teachers to become certified as volunteer monitors for the Kentucky Watershed Watch program

Community site visits were made to an industry to learn about the role of industry in protecting aquatic systems, the local water treatment agency to learn about it's best management practices for storm water runoff, and a canoe and kayak business to learn about the impact of flooding on local businesses and about water management issues. Information presented at each of the community sites included the educational opportunities offered for schools and community members.

Representatives from the community also served as guest speakers, representing county government, an Ohio River regulatory agency, two non-profit organizations, and the National Energy Education Development project. These speakers described their role in protecting aquatic systems, presented information about programs that they offered for schools, such as classroom resources, teacher professional development, K-12 field trips and grant programs, and citizen volunteer opportunities.

Teachers were led in a 4-hour canoe trip as a means to experience the river. Half of the teachers in the course had not previously been in a canoe or kayak. It was therefore a unique experience for them to be on a river, and an experience which helped to connect them to place as the river they canoed or kayaked was the main stem of the river which formed the watershed basin in which they live and work.

The course was designed to increase the potential that teachers would transfer their course learning to their classroom teaching by explicitly addressing the state science standards. In the early stages of designing the course, the instructors identified the K-12 state science standards to be addressed in the course. The general areas of science standards taught in the course included properties and changes of properties in matter, and transfer of energy in the physical sciences; structure of the earth system – lithosphere, hydrosphere and atmosphere, geochemical cycles, and formation and ongoing changes of the earth system in the earth sciences; diversity and adaptations of organisms, behavior of organisms, populations and ecosystems, biological change, and interdependence of organisms in the life sciences; science and technology, science in personal and social perspectives, history and nature of science, and scientific inquiry.

All course topics and experiences were then aligned with these standards. Reflection and discussion about these standards were built into the daily course activities. Teachers met in grade level groups at the end of each class session to identify the standards that were addressed by the lessons taught that day. During the beginning of each day, a class discussion was facilitated by the instructors to discuss the standards taught during the previous day. In addition, teachers were assigned homework each night to develop an open-response assessment item related to the standards taught that day. These questions modeled the type of questions that K-12 students are required to answer as part of the state testing system. Teachers were required to record the science standards and their assessment question on a daily basis in their course notebooks. These notebooks were graded at the end of the course by the instructors.

Another major assignment that required teachers to utilize the state science standards was the final project. At the end of the two-week session of the course, teachers spent two weeks developing a unit of study based on their course learning that they would teach to their students. This unit was aligned to the state science standards.

A follow-up session, held two weeks after the two-week segment of the course, focused on classroom applications of the two-week training. Teachers presented their unit of study that they had designed to teach their students. All projects were posted on the course website as a means for teachers to share ideas and resources with each other and with other educators.

Ongoing Support for Teachers

Support for teachers to make and sustain changes in their classrooms is recognized as a critical component of training programs (AAAS, 1998; NRC, 1996; Powers, 2004; Rhoton, et al., 1999). The types of support made available in this course included: 1) establishing a network of university and community-based professionals; 2) providing classroom resources to teachers; 3) requiring that a unit of study, aligned with state standards, be developed and used in the classroom by teachers; 4) conducting a follow-up sessions for teachers to share their units; and 5) developing a course website, http://www.nku.edu/~enved/aet.htm and a group email list, which allowed teachers continued access to information from each other, the course instructors, and the community-based experts they encountered during the course.

Place-Based Course Components

A broad goal of the program was to use the local environment as an integrating context to teach about the interactions between environmental systems and human systems. A more specific program goal was to promote awareness and understanding of the human and environmental forces that impact the health of a watershed. Instruction to accomplish these goals addressed the three important ideas that shape the instructional vision as stated in the Guidelines for the Preparation and Professional Development of Environmental Educators (NAAEE, 2004c). These ideas emphasize a systems approach to education, the interdependence between human systems and ecological systems, and the importance of where one lives.

Salient aspects of placed-based teaching approaches used and modeled in the program were using the environment as an integrating context across disciplines, collaboration between program leaders, participants and members of the community, reflective learning, experiential learning, relevancy to real life and current societal issues, and citizenship education.

Structuring time during the course for teachers to reflect about what they were learning is an important practice within educational programs (Clark, 1994; Ginsbury & Clift, 1990; Henson, 1996; Johnson, Guice, Baker, Malone, & Michelson, 1995; Meichtry, (1998); Reynolds, 1992; Rhoton, Madrazo, Motz, & Walton, 1999; Shulman, 1986). Teachers in this course reflected on their experience and applications to teaching through journaling, class and small group discussions, the course assignments, and the course evaluation.

The practice of using outside experts is supported as a way to enhance learning and of increasing the potential for community change (e.g., Bouillion & Gomez; 2001; Ciffone, Morelock, Turner, Sivek, & Daudi, 2002; Jakowska, 1987; Niesenbaum & Gorka, 2001; O'Neill & Gomez, 1998; Rhoton, et al., 1999). To this end, seven community-based specialists and three university faculty were scheduled throughout the course to share their expertise about the environment and/or community-based efforts and resources.

Experiential learning is advocated as a teaching approach for accomplishing educational objectives in both the cognitive and affective domains (Chawla, 1998; Chawla, 1999; Heimlich & Daudi, 2002; Jarvis, 1987; Niesenbaum & Gorka, 2001; Reeder, 1998; Rome & Romero, 1998; Uno, 1990). The experiential study of an aquatic system, river monitoring, and interaction with community-based experts accomplished each of the five objective areas of the Tbilisi Declaration (1978). The teachers developed awareness, conceptual understandings, attitudes and values, citizen action skills, and citizen action experience. The course also addressed each of the four curriculum goals derived from the Tbilisi Declaration objectives by Hungerford, Peyton, and Wilke (1980). These goals are ecological foundations, conceptual awareness about issues and values, investigation of environmental issues and evaluation of alternative solutions, and training in skills and action for the purpose of achieving equilibrium between the quality of life and quality of the environment.

Purpose of Study

The objectives of the course were to increase the level of confidence and degree to which the teachers a) use technology in their teaching, b) use standards-based teaching strategies,
c) integrate the sciences, d) integrate science with other subject areas, e) use the local environment, f) conduct field-based investigations, g) use community-based resources, h) teach watershed topics, and i) teach real-world current issues. The purpose of the study was to evaluate the impact of the course on teachers' confidence levels and classroom practices which related to the program objectives.

Methods

Participants

There were 16 course participants. Four of the participants taught K-4th grades, seven taught 5-8th grades, and five taught high school. Their number of years teaching experience ranged from one year to thirteen years. The teachers represented eleven school districts; three taught in rural schools, nine in suburban schools, and four in urban schools. Thirteen of the teachers taught in public schools and three taught in private schools.

Participants in the course received three graduate credits which they could apply to a Masters degree in Education or Rank 1 certification, which is 30 credits beyond a Masters degree. Twelve of the teachers were enrolled in a Masters degree program, with some at the beginning, some at the mid-point, and some near the end of their program. Four of the teachers had completed their Masters degree.

The design and measures utilized in this study were developed as part of an earlier professional development program evaluation (Meichtry & Smith, in press). The authors' descriptions of the design and measures from this previous study are included in the following two sections.

Design

This design was a repeated measures pre-test, post-test, delayed-term post-test design. The independent variable was the time of testing: pre-program, post-program and long-term (9 months) post-program. The nine-month post measure was included in order to assess the long-term impact of the program. Dependent measures of confidence in the ability to teach program related topics were assessed at all three time periods. Dependent measures of actual teaching of program related topics were assessed only at preprogram and long-term post program time periods. The major advantage of this type of repeated design is that each participant acts as his/her own control, resulting in the need for fewer subjects and a higher level of statistical sensitivity (Martin, 1991; pp. 67-70).

Measures

Participants' confidence in the ability to teach course relevant topics were assessed just prior to the beginning of the course, immediately after the course, and nine months after the end of the course, using a 5-point Likert scale with the response options being very low (1), low (2), average (3), high (4) and very high (5) confidence.

Participants' use of course related instructional techniques were assessed just prior to the beginning of the course and nine months after the end of the course using 5 point Likert scales. The response options, depending on the phrasing of the question, were as follows: never (1), 1-2 times a year 2), 3-4 times a year (3), 5-6 times a year (4), over 6 times a year (5) or never (1), rarely (2), sometimes (3), often (4), and always (5).

Five areas of assessment were developed to be consistent with the five major course curriculum areas. These areas were, confidence in: 1) the ability to use workshop demonstrated teaching technologies (9-items), 2) the ability to use workshop demonstrated instructional strategies (5-items), 3) the use of community resources, (3-items), 4) the ability to conduct field-based investigations (7-items), and 5) the ability to teach water quality topics and the connections between science and real life, social issues and science related careers (4-items). See Survey Instrument presented in Appendix A for a listing of the specific items.

The actual teaching of course topics by participants was assessed just prior to the beginning of the course and again nine months after the course ended. The areas of assessment and number of items were the same as the confidence areas listed above, except that they addressed the actual use of classroom practices rather than level of confidence in using these practices.

In addition to the statistical measures used to determine the course impacts, an open-ended questionnaire was administered to the participants as part of the delayed post-survey assessment. The questionnaire asked participants to identify the strengths of the course, the single most beneficial aspect of the course related to content, pedagogy and classroom teaching, impact of the course on student learning, and barriers to implementing the course material in a K-12 setting.

Results

Confidence Ratings

Comparisons between pre, post and long-term post participant confidence measures were made. Confidence measures were grouped into five areas, which were confidence in 1) the ability to use workshop demonstrated teaching technologies, 2) the ability to use workshop demonstrated teaching and instructional strategies, 3) the use of community resources, 4) the ability to conduct field-based investigations and 5) the ability to teach watershed topics and teach the connections between science and real life, social issues and science related careers.

A MANOVA was performed to make a pre, post, and long-term post comparison using the nine-dependant variable assessing the use of technology. The MANOVA was conducted because it creates a combined dependent measure for interrelated items which reduces the probability of type 1 error when there are multiple interrelated dependent variables being analyzed (Pallant, 2005, p. 247). Due to insufficient degrees of freedom, the MANOVA could not be calculated. Separate ANOVA analyses were performed on the nine individual items making up the scale. All nine of the individual scale items were significant. Post hoc Bonferroni comparisons analyses found greater confidence on eight of the nine items on the long-term post measures as compared to the pre workshop assessment, with the exception being the use of Excel which was found to be significantly different from post to long-term post; see Table 1.

TABLE 1. ANOVA and mean pre, post and long-term post confidence ratings for the use of instructional technologies.

Items	Pre	Post	Long-term	F(2,26)	p
Water quality sampling kits	3.0a	4.3b	4.2b	20	.001
Water study data probes	2.5a	3.9b	3.7b	14	.001
Excel spread sheet program	3.1ab	2.9a	3.6b	5	.01
Internet	4.0a	4.4ab	4.6b	4.7	.02
Microscopes	3.8a	4.0a	4.6b	7	.004
Videoscopes and/or display monitor	2.9a	3.2a	3.9b	7.1	.003
Digital camera	3.8a	4.1a	4.8b	8	.002
Global positioning system	1.9a	3.7b	3.8b	28	.001
Two-way radio	3.1a	3.9ab	4.3b	9	.001

Note: Means not sharing a common letter are significantly different at the p<.05 level using the Bonferroni procedure.

Five items assessing confidence in the ability to use effective instructional strategies were compared across the time of testing using a MANOVA. A significant multiple F was obtained (Wilks' Lambda=.03; F(10,6)=17.5, *p*=.001). Separate ANOVAs were performed on all five dependent variables. All five variables were found to be statistically significant with greater confidence found in the post and long-term post measures as compared to the pre workshop assessment; see Table 2.

TABLE 2. ANOVA and mean pre, post and long-term post confidence ratings for the use of instructional strategies.

Items	Pre	Post	Long-term	F(2,26)	*p*
Use hands-on instructional strategies	4.1a	4.7b	4.8b	9	.001
Use inquiry-based teaching strategies	3.9a	4.6b	4.4b	5.3	.01
Address gender and minority equity	3.1a	4.3b	4.1b	17	.001
Integrate the sciences in teaching	3.8a	4.4b	4.6b	9	.001
Integrate science as a subject with other subject areas	3.7a	4.4b	4.4b	10	.001

Note: Means not sharing a common letter are significantly different at the p<.05 level using the Bonferroni procedure.

The three items assessing confidence in the use of community resources were compared across time of testing using a MANOVA. A significant multiple F was obtained (Wilks' Lambda=.18; F(6,8)=6, *p*=.012). Separate ANOVAs were performed on all three dependent variables. Significant differences were found for all three, with greater confidence found in the post and long-term post measures as compared to the pre workshop assessment; see Table 3.

TABLE 3. ANOVA and mean pre, post and long-term post confidence for the use of community resources.

Items	Pre	Post	Long-term	F(2,26)	*p*
Guest speakers	3.6a	4.1b	4.1b	3.2	.001
Natural environment field sites related to watershed studies	2.9a	4.3b	4.1b	19	.001
Field trips to watershed-related community resource sites	3.0a	4.4b	4.1b	13	.001

Note: Means not sharing a common letter are significantly different at the p<.05 level using the Bonferroni procedure.

Seven items assessing confidence in the ability to conduct field investigations were compared across time of testing using a MANOVA. A significant multiple F was obtained (Wilks' Lambda=.13; F(14,1)=5.5, *p*=.001). ANOVA tests were performed separately on all seven dependent measures. Significant differences were also obtained for all seven measures, with greater confidence found in the post and long-term post measures as compared to the pre workshop assessment; see Table 4.

TABLE 4. ANOVA and mean pre, post and long-term post confidence in the ability to conduct field investigations.

Items	Pre	Post	Long-term	F (2,26)	*p*
Geological study of water Systems	2.3a	3.6b	3.8b	26	.001
Microscopic study of aquatic Life	2.8a	4.1b	3.9b	17	.001
Macroinvertebrate Study	2.6a	4.3b	4.2b	31	.001
Habitat assessment	2.7a	4.1b	4.3b	31	.001
Fish Study	2.5a	3.9b	3.8b	38	.001
Terrestrial ecology	2.6a	3.9b	3.8b	24	.001
Water Chemistry	2.9a	4.4b	4.5b	37	.001

Note: Means not sharing a common letter are significantly different at the p<.05 level using the Bonferroni procedure.

Four items assessing confidence in the ability to teach watershed topics and the connections between science, real life, social issues and science careers were combined into a MANOVA. A significant multiple F was obtained (Wilks' Lambda=.15; F(8,54)=11, *p*=.001). Follow-up ANOVAs were performed on each of the separate dependent measures. Significant differences were obtained for all of the measures. Post hoc tests found greater confidence in the post and long-term post measures as compared to the pre workshop assessment for all of the items; see Table 5.

TABLE 5. ANOVA and mean pre, post and long-term post confidence in the ability to teach watershed and science linked topics.

Items	Pre	Post	Long-term	F (2,26)	*p*
Water quality topics	2.9a	4.4b	4.4b	61	.001
Connections between science and real life	3.8a	4.6a	4.6b	21	.001
Connections between science and societal issues	3.6a	4.6b	4.5b	22	.0001
Connections between science and science-related careers	3.8a	4.2b	4.2b	3.9	.03

Note: Means not sharing a common letter are significantly different at the p<.05 level using the Bonferroni procedure.

In summary, the positive impacts on teachers' confidence are evidenced by the overall significant gains in all five of the confidence level measures. The course had a strong impact on teachers' confidence to teach in all five of the major program curriculum areas. Compared to the pre-course assessment, greater confidence was reported in all but one of the 28 post-survey measures and in each of the 28 delayed post-survvey measures.

Classroom Practice Assessments

Pre workshop and long-term follow-up comparisons were made of the actual use of 1) workshop-demonstrated teaching technologies, 2) workshop-demonstrated instructional strategies, 3) use of community resources, 4) conducting field-based investigations and 5) the teaching of watershed topics, connections between science and real life, social issues and science related careers.

The nine items assessing the use of technologies were combined in a MANOVA. A non-significant multiple F was obtained (Wilks' Lambda=.17; F(9,4)=2.1, *p*=.23). Separate pre to long-post comparisons were preformed on each of the nine dependent variables using an ANOVA. Two of the seven, the use excel and Internet web sites were found to be significant, with greater use reported after the workshop. A third measure, the use of video scopes was found to approach statistical significance, see Table 6.

TABLE 6. ANOVA, Mean pre and long-term post use ratings of instructional technologies.*

Items	Pre	Long	F(2,26)	*p*
Water quality sampling kits	2.0	2.78	1.9	ns
Water study data probes	1.8	2.15	.55	ns
Excel	2.38	2.69	7	.02
Internet websites for research and support materials	4.0	4.9	11	.006
Microscopes	2.9	3.23	.34	ns
Videoscopes	2.08	3.31	3.7	.08
Digital camera	3.46	3.46	2.9	ns
Global positioning systems	1.07	1.38	2.2	ns
Two-way radio	2.08	1.46	.43	ns

*Scale values: 1 = never, 2 = 1-2 times a year, 3 = 3-4 times a year, 4 = 5-6 times a year, 5 = over 6 times a year. Note: p<.05 was determined to be significant.

The five items assessing the use of instructional strategies were combined in a MANOVA. A non-significant multiple F was obtained (Wilks' Lambda=.54; F(5,8)=1.4, *p*=.33). Separate pre to long-post comparisons were preformed on each of the five dependent variables using an ANOVA. The use of inquiry based teaching strategies was found to significantly differ pre to long-term post, with greater use reported after the workshop. A second item, integrate the sciences in teaching, was found to approach significance; see Table 7.

TABLE 7. ANOVA, mean pre and long-term post use ratings for the use of instructional strategies.*

Items	Pre	Long-term	F(2,26)	*p*
Use hands-on instructional strategies	4.69	5.0	1.7	ns
Use inquiry-based teaching strategies	4.31	4.69	7.5	.02
Address gender and minority equity	1.92	2.38	.51	ns
Integrate the sciences in teaching	4.23	4.85	3.5	.09
Integrate science as a subject with other subject areas	3.92	4.62	2.6	ns

*Scale values: 1 = never, 2 = 1-2 times a year, 3 = 3-4 times a year, 4 = 5-6 times a year, 5 = over 6 times a year. Note: p<.05 was determined to be significant.

The three items assessing the use of community resources were combined into a MANOVA. A non-significant multiple F was obtained (Wilks' Lambda=.86; F(3,10)= .57, *p*=.65). Each of the three items making up the scale were also compared pre to post using an ANOVA. However, no significant differences were found; see Table 8.

TABLE 8. ANOVA, mean pre and long-term post use ratings for the use of community resources.

Items	Pre	Long-term	F(2,26)	*p*
Guest speakers	2.15	2.46	.45	ns
Natural environment field sites related to watershed studies	1.54	2.54	1.9	ns
Field trips to watershed related community resources sites	1.46	1.77	.79	ns

*Scale values: 1 = never, 2 = 1-2 times a year, 3 = 3-4 times a year, 4 = 5-6 times a year, 5 = over 6 times a year. Note: p<.05 was determined to be significant.

Seven items assessing the use of field investigations were combined into a MANOVA. A non-significant multiple F was obtained (Wilks' Lambda F(7,3)=.88, *p*=.60). Individual ANOVA comparisons on each of the dependent measures found none of the individual items to be significant; see Table 9.

TABLE 9. ANOVA, mean pre and long-term post use ratings for the use of field-based investigations.

Items	Pre	Long-term	F (2,26)	*p*
Geological study of water systems	1.69	2.23	.87	ns
Microscopic study of aquatic life	2.08	2.0	.23	ns
Macroinvertebrate study	1.58	2.0	.07	ns
Habitat assessment	1.84	2.3	.07	ns
Fish Study	1.77	1.92	.0	ns
Terrestrial ecology	1.58	2.17	1.0	ns
Water chemistry	1.92	2.62	1.6	ns

*Scale values: 1 = never, 2 = 1-2 times a year, 3 = 3-4 times a year, 4 = 5-6 times a year, 5 = over 6 times a year. Note: p<.05 was determined to be significant.

Four items assessing the teaching of watershed topics and the connections to life were combined in a MANOVA. A significant multiple F was obtained (Wilks' Lambda F(8,8)=8.1, *p*=.006). ANOVA comparisons on each of the dependent measure found significant difference in two of the four measures, teaching about watershed topics and connections between science and real life, with greater teaching in the long-term post measures as compared to the pre course assessment: see Table10.

TABLE 10. ANOVA, mean pre and long term post means on the extent of teaching watershed and science linked topics.

Items	Pre	Long-term	F (2,26)	*p*
Waters quality topics	2.85	3.54	6.7	.03
Connections between science and real life	4.23	4.85	9.1	.012
Connections between science and societal issues	4.08	4.67	1.8	ns
Connections between science and science-related careers	3.75	3.83	0	ns

*Scale values: never (1), rarely (2), sometimes (3), often (4) and always (5). Note: $p<.05$ was determined to be significant.

In summary, the results of the MANOVA revealed significant gains in the teaching of watershed topics and connections to real life and social issues. While the MANOVA results showed a lack of overall significance in the use of instructional technologies, standards-based instructional strategies, field-investigation, and community-based resources, the program did have a significant impact, as evidenced by the ANOVA procedure, on the use of Excel, Internet-based resources and inquiry-based teaching strategies, and on the teaching of water quality topics and the connections between science and real life.

Delayed Post-Survey Open-Ended Questionnaire Assessments

The two most frequently identified strengths of the course, reported by eight of the 16 teachers, were the hands-on activities and field studies taught throughout the course. Another five teachers reported the course strengths as the practical uses of content, skills, lessons, pedagogy, equipment, and the resources provided to them to use in their teaching.

Teachers were asked to identify the single most beneficial aspect of the course in regard to content, pedagogy, and classroom teaching. The most frequent response for the course strength related to content, reported by eight teachers, was the depth of content learned. Real life examples and the link between content, social issues, and environmental issues was the second most frequent response, reported by three teachers. The most frequent response for the course strength related to pedagogy, reported by seven teachers, was the use of hands-on activities in the course. The use of community resources (5 teachers), teaching materials (4 teachers), and the connection of course content to real life and current issues (3 teachers) were the most frequent responses to the strength of the course in regard to classroom teaching.

In response to the question of how their participation in the course will help improve their students' learning, ten of the teachers reported that their students would learn more content, five reported that their students would become more involved in activity-based learning and thus learn more, four reported that their students would experience more real life learning, three reported that their students would utilize an increased number of resources in their learning, and two reported that their students would learn more about human impact and responsibility and would become more excited and positive because their teacher was.

While six of the teachers reported that there were no barriers to teaching the course content in their classrooms, the other teachers reported barriers related to taking their students on field trips and issues related to curriculum. The barriers identified by teachers to utilizing field trips in their classroom teaching included the lack of usable sites that are safe and near the school (3 teachers), the difficulty of obtaining permission slips for all students to participate in field trips (1 teacher), limited ability to take field trips based on scheduling constraints (1 teacher), and the difficulty presented by large class sizes (1 teacher).

Barriers to implementing the course curriculum in the classroom were identified by five of the 16 teachers. A barrier reported by two of the teachers was the lack of a match between the curriculum taught during the course and the subject matter that they teach in the classroom. Other barriers reported were a limited amount of time devoted to the 5th grade science curriculum due to a focus on the social studies and math state assessment (1 teacher), the difficulty of adapting the course content to the level of elementary students (1 teacher), and the requirement to teach from a science kit that leaves little time to teach other activities (1 teacher).

Discussion

The results of this study support findings and recommendations of previous studies that incorporating several tenets of place-based education in training programs has the potential to improve educational outcomes. Educational benefits have been reported in the literature for the use of experiential education practices (Chawla, 1998; Chawla, 1999; Heimlich & Daudi, 2002; Reeder, 1998); utilizing different expertise in program leadership, building relevancy into the program by using a local setting and involving community-based experts (Bouillion & Gomez; 2001; Ciffone, et. al., 2002; Jakowska; Niesenbaum & Gorka, 2001; O'Neill & Gomez, 1998; Rhoton, et al., 1999); allowing time for participants to reflect about their learning (Clark, 1994; Ginsbury & Clift, 1990; Henson, 1996; Meichtry, (1998); Reynolds, 1992; Rhoton, et al., 1999; Shulman, 1986); and establishing means of ongoing support for the participants (AAAS, 1998; NRC, 1996; Powers, 2004; Rhoton, et al., 1999). In addition to these studies, which focus on the use of a single tenet of place-based education, Lieberman and Hoody (1998) found that using a comprehensive set of place-based teaching strategies when using the environment as an integrating context yielded positive educational outcomes for K-12 students.

For K-12 students to realize the benefits of place-based education, it is critical that classroom teachers be effectively trained in the use of place-based teaching strategies. Studies such as this are thus needed to determine the impact of educational programs that focus on teacher education and that utilize the comprehensive set of teaching strategies which constitute place-based education. These studies are necessary to help guide the design, implementation, and evaluation of teacher education courses and professional development programs that utilize a comprehensive set of place-based teaching strategies.

The comprehensive set of place-based education strategies used in this study as the basis for the teacher education course design and implementation resulted in a course that positively impacted the confidence levels of teachers to use place-based classroom practices. The areas impacted were the use of instructional technologies, the use of standards-based in teaching, the use of community resources and the natural environment in teaching, the use of field-based investigation in teaching, and the teaching of water quality, science, and societal topics.

There was a statistically significant gain evidenced in the areas of using Excel and the Internet as instructional technologies, the use of inquiry-based teaching, and the teaching of water quality topics and the connection between science and real life. While the statistical significance of classroom use measures was not evidenced in the majority of item measures, there were other results that revealed positive impacts of the course on the teachers' use of classroom practices. These results included an increase in 25 of 28 item means from pre-survey to delayedpost-survey and the responses of teachers to the delayed post-survey open-ended questions. Given the small sample size of 16 participants, the amount of time and support needed to enact change, and obstacles that existed within the school setting as reported by teachers, it was encouraging that the direction of the change in means indicated a positive change in classroom teaching practices.

Research design recommendations based on this study relate to the results of the delayedpost-measure, conducted nine months after the summer course ended. The fact that the measures were so specifically related to course activities lessened the likelihood that other interventions would have impacted the long-term positive outcomes of this study. The results of a repeated measures design, with the delayed post-measure analysis, revealed the extent to which positive impacts on teachers' confidence and use of classroom practices were sustained over time.

An analysis of the results of the open-ended questionnaire indicated that the teachers learned content, revealed areas of pedagogy that were learned and could be applied in their classroom teaching, and revealed areas of student learning that would be and were impacted due to the participation of teachers in the course. The questionnaire results also revealed barriers that made it difficult for some of the teachers to implement aspects of the course in classroom teaching. Knowledge of these barriers, which related to taking K-12 students on field trips and curriculum issues in this study, are important to ascertain so that instructors may plan ways to address barriers faced by teachers in future course sessions.

Instructors addressed the lack of opportunity faced by teachers for taking their students on field trips by conducting lessons during the course that demonstrated alternative ways to teach the same or similar field trip content in the classroom. These lessons demonstrated ideas for classroom experiments, simulations, outdoor education at the school site, role playing activities, the use of models for demonstration purposes, powerpoint presentations with digital photos of field trip sites, Internet websites, and community guest speakers. Resources that were necessary for these activities were made available by the instructors on a loan basis to the teachers.

It is recognized that providing support for teachers to make and sustain changes in their teaching is an important component of training programs (AAAS, 1998; NRC, 1996; Powers, 2004; Rhoton, et al., 1999). The types of support for teachers provided in this course were a network of university faculty, community-based experts and teachers, print and electronic media resources for classroom use, resources for loan, a course website, and a group email list. Other types of support to help teachers make classroom changes included the requirement that teachers develop a unit of study aligned with state science standards and share these units with one another. The degree to which the teachers used the types of support made available during the course over time was not measured as part of this study. However, a review of the units of study developed by the teachers to be taught in upcoming and future years indicated that teachers were using a variety of types of these support tools. Four of the 16 teachers used community speakers who spoke during the course as guest speakers in their classrooms, two teachers conducted field trips to community sites they were connected to during the course, seven teachers checked out resources for loan, and virtually all of the teachers made use of print and electronic resources provided during the class. The group email list continues to be used by the instructors to update the teachers on opportunities and resources related to the course topics.

While the results of this study indicated that the course had a positive impact on teachers' confidence levels, classroom practices, and potential impact on student learning, these results also revealed that the statistically significant gains in confidence levels of teachers did not translate into statistically significant gains in classroom practice for the majority of teaching areas measured. Follow-up qualitative studies are thus recommended to: 1) determine whether the reasons the teachers are not using particular classroom practices to a significantly greater degree are factors that can be addressed in future training programs; and 2) whether increases in the confidence level of teachers nine months after the program were based on aspects of the summer program, the experience of applying the summer learning to classroom teaching, or factors unrelated to the program. It is also recommended that follow-up studies be conducted to determine whether the reasons the teachers are using classroom practices to a greater degree than indicated on the pre-survey are related to the course.

Recommendations for the design, teaching, and evaluation of the course, based on the results of this study, are to increase the likelihood that classroom practices will be implemented by teachers through the following means:

Alignment of Course Content with State Science Standards and School Curriculum

- develop clearly stated objectives that are linked to the state and/or school standards that teachers are required to address.;
- explicitly connect course learning to K-12 classroom teaching, state standards, and school curriculum;
- design instruction to be explicitly connected to classroom practice by requiring units of study to be developed by teachers for use in their classroom teaching;
- help elementary teachers to adapt the course learning to their grade levels.

Course Teaching Strategies

- model effective teaching strategies for use with K-12 students in the teaching of the course;
- use assessment methods that model how teachers should asses their own students;
- build relevancy into the program by using the local environment and community as a context for integrating course topics, and by using investigation protocols and equipment used by the state or local government and citizen groups.

Support for Teachers

- provide classroom resources to teachers and/or develop a lending library;
- provide information on a website that teachers can utilize over time;
- provide a formal means for participants to establish networks with each other and with community experts;
- provide ideas for ways to overcome field trip barriers in K-12 schools.

Program Evaluation

- conduct program evaluation that is directly aligned to the program objectives and use the results to improve the course over time;
- conduct interviews and classroom observations of teachers as a means to improve the validity of the self-reported survey and open-response questionnaire data;
- collect demographic data such as prior teaching experience, advanced degrees held, type of school setting (rural, suburban, urban, private, public), and analyze these data to determine potential impacts to different classroom settings and across teaching contexts;
- measure the degree to which participants use the ongoing support components that were established throughout the course.

References

American Association for the Advancement of Science (1998). *Blueprints for reform in science. mathematics, and technology education" Project 2061*. New York, NY: Oxford University Press.

Atham, J. & Monroe, M. (2004). The effects of environment-based education on students' achievement motivation. *Journal of Interpretative Research*, 9(1), 9-25.

Audubon Washington. (2004). *2004 Report Card on the Status of Environmental Education in WashingtonState.* http://wa.audubon.org/new/audubon/userdocuments/EEReportCard.pdf.

Bouillion, L. & Gomez, L. (2001). Connecting school and community with science learning: Real world problems and school-community partnerships as contextual scaffolds. *Journal of Research in Science Teaching*, 38(8), 878-898.

Carr, D. (2005). *Incorporating environmental content: An online survey of K-12 teachers in Kentcuky concerning current educational practices and preferences for education resources.* Frankfort, KY: Kentucky Environmental Education Council.

Chawla, L. (1998). Significant life experiences revisited: A review of research on sources of environmental sensitivity. *The Journal of Environmental Education*, 29(3), 11-21.

Chawla, L. (1999). Life paths to effective environmental action. *The Journal of Environmental Education*, 31(1), 15-26.

Ciffone, E., Morelock, J., Turner, K., Sivek, D., & Daudi, S. (2002). Using community resources in environmental education. In J. Heimlich (Ed*.) Environmental education: A resource handbook*. Bloomington, IN: Phi Delta Kappa.

Clark, A. (1994). Student –teacher reflection: Developing and defining a practice that is uniquely one's own. *International Journal of Science Education*, 16(5), 497-509.

Ernst, J. & Monroe, M. (2004). The effect of environment-based education on student' critical thinking skills and disposition toward critical thinking. *Environmental Education Research*,10(4): 507-522.

Ginsburg, M. & Clift, R. (1990). The hidden curriculum of preservice teacher education. In R. Huston (ed.), *Handbook of research on teacher education*. New York: MacMillan.

Heimlich, J. (2002). Integrating environmental education and informal education. In J. Heimlich (Ed*.) Environmental education: A resource handbook*. Bloomington, IN: Phi Delta Kappa.

Heimlich, J. & Daudi, S. (2002). Making the experience of environmental education experiential. In J. Heimlich (Ed*.) Environmental education: A resource handbook*. Bloomington, IN: Phi Delta Kappa.

Henson, K. (1996). *Methods and strategies for teaching in secondary and middle schools* (3rd ed.). White Plains, NY: Longman.

Hungerford, H., Peyton, R., & Wilke, R. (1980). Goals for curriculum development in environmental education. *Journal of Environmental Education*, 11(3), 42-47.

Jackowska, S. (1987). Community-based education: Highlights of the prevailing trends and recommendations. In A. Baez, G. Knamiller, & J. Smyth (Eds.), *The environment and science technology education.* Oxford: Pergamon Press.

Jarvis, P. (1987). *Adult learning in the social context*. London: Croom Hilm.

Johnson, P., Guice, S., Baker, K., Malone, J., & Michelson, N. (1995). Assessment of teaching and learning in literature-based classrooms. *Teaching and Teacher Education*, 11, 359-371.

Kentucky Division of Water (2000):

a. *Watershed watch biological stream assessment*. Frankfort, KY: Author.
b. *Watershed habitat assessment protocols.* Frankfort, KY: Author.
c. *Watershed watch water chemistry sampling methods*. Frankfort, KY: Author.

Lamotte Water Monitoring Equipment, (n.d.). Retrieved June 11, 2006, from http://www.lamotte.com/pages/edu/mon-list.html.

Lieberman, G. and Hoody, L. (1998). *Closing the achievement gap: Using the environment as an integrating context for learning*. San Diego, CA: State Environment and Education Roundtable.

Martin, D. (1991). Doing Psychology Experiments. Brooke/Cole Publishing Company, PacificGrove: CA.

Meichtry, Y. & Smith, J. (In press, 2007). The Impact of a Place-Based Professional Development Program on Teachers' Confidence, Attitudes, and Classroom Practices. *Journal of Environmental Education.*

Meichtry, Y. (1998). Elementary science methods: Developing and measuring student views about the nature of science. In W.F. McComas (ed.), *The nature of science in science education: Rationales and strategies.* Dordrecht, the Netherlands: Kluwer Publications.

Meichtry, Y. & Harrell, L. (August, 2002). An environmental education needs assessment of K-12 teachers. *Journal of Environmental Education* 33(3), 21-26.

National Environmental Education Training Foundation (NEETF). (2000). *Environment-based Education: Creating High Performance Schools and Students.* Washington, DC: NEETF.

Niesenbaum, R. & Gorka, B. (2001). Community-based eco-education: Sound ecology and effective education. *Journal of Environmental Education*, 33(1), 17-20.

North American Association for Environmental Education (2004).

a. *Excellence in environmental education: Guidelines for learning pre K-12*. Washington, DC: NAAEE Publications.

b. *Environmental education materials: Guidelines for excellence*. Washington, DC: NAAEE Publications.

c. *Guidelines for the Preparation and Professional Development of Environmental Educators.* Washington, DC: NAAEE Publications.

National Research Council. (1996). *The national science education standards*. Washington DC: National Academy Press.

O'Neill, D. & Gomez, L. (1998). Proceedings of the ACM 1998 Conference on Computer Supported Cooperative Work, *Sustaining mentoring relationships online*. New York, NY: ACM Press.

Orr, D. (1994). *Earth in mind: On education, environment, and the human prospect*. Washington, DC: Island Press.

Pallant, J. (2005). *SPSS, 2nd Edition Survival Manual*. Berkshire, UK: Open University Press.

Powers, A. (2004). An evaluation four place-based education programs. *The Journal of Environmental Education*, 35(4), 17-32.

Reeder, B. (1998). Using water quality monitoring as a university-level teaching tool. *Journal of the Kentucky Academy of Science,* 59, 6-11.

Reynolds, A. (1992). What is competent beginning teaching? A review of the literature. *Review of Educational Research*, 62(1), 1-35.

Rhoton, C., Madrazo, G., Motz, L., & Walton, E. (1999). Professional development: A major component in science teaching and learning. *Science Educator*, 8(1), 1-9).

Rome, A. & Romero, B. (1998). Enhancing conservation education opportunities in tropical countries: A case study in Belize. *The Journal of Environmental Education*, 30(1), 34-39.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.

Smith, G. (2002). Place-based education: Learning to be where we are. *Phi Delta Kappan*, 83(3), 584.

Smith, G. & Williams, D. (Eds.) (1999). *Ecological education in action: On weaving education, culture, and the environment*. Albany, NY: State University of New York Press.

Sobel, D. (2005). *Place-based education: Connecting classrooms and communities*. Barrington, MA: The Orion Society.

State Education and Environment Roundtable (SEER). (2000). *California Student Assessment Project: The Effects of Environment-based Education on Student Achievement*. San Diego: SEER.

Tbilisi Intergovernmental Conference on Environmental Education. (1978). *Toward an action plan: A report on the Tbilisi Conference on environmental education*. A report developed by the FICE Subcommittee on Environmental Education. Washington, DC: U.S. Government Printing.

Uno, G. (1990). Inquiry in the classroom. *Bioscience*, 40(3), 841-843.

Woodhouse, J. & Knapp, C. (2000). *Place-Based Curriculum and Instruction: Outdoor and Environmental Education Approaches.*ERIC Clearinghouse on Rural Educationand Small Schools. (ERIC Document Reproduction Service No. ED 448012).

Appendix A: Confidence Measures for Pre, Post, and Long-Term Post

Please rate your <u>confidence</u> in your own ability to use the following technologies:	Very Low	Low	Average	High	Very High
1. Water quality sampling kits.					
2. Labware, probes, CBLs, and graphing calculators.					
3. Internet websites for research and support materials.					
4. Microscopes.					
5. Videoscopes and/or display monitor.					
6. Presentation technologies (slides, power point, etc.)					
7. Digital camera.					
8. Geographic Positioning System (GPS)					
Please rate your <u>confidence</u> in your own ability to:	XXXX	XXXX	XXXXX	XXXX	XXXXX
9. Use hands-on instructional strategies.					
10. Use inquiry-based teaching strategies.					
11. Address gender and minority equity through instruction.					
12. Integrate the sciences (physical, life, earth) in teaching.					
13. Integrate science as a subject with other subject areas.					
Please rate your <u>confidence</u> in your own ability to use the following community resources:	XXXX	XXXX	XXXXX	XXXX	XXXXX
14. Guest speakers (local, university, county, and/or state).					
15. Natural environment <u>**field**</u> sites related to watershed studies.					
16. Field trips to watershed related community resource sites (Museum, historical society, fish hatchery, farm site, etc.)					
Please rate your <u>confidence</u> in your own ability to conduct the following field based investigations:	XXXX	XXXX	XXXXX	XXXX	XXXXX
17. Water chemistry					
18. Macroinvertebrate study					
19. Habitat assessment					
20. Fish study					
21. Plankton collection					
22. Geology study with topo maps					
Please rate your <u>confidence</u> in your own ability to teach:	XXXX	XXXX	XXXXX	XXXX	XXXXX
23. about watershed topics.					
24. about connections between science & real life.					
25. connections between science & societal issues.					
26. connections between science & science-related careers.					
Please rate the general enthusiasm of the following groups of students for science:	XXXX	XXXX	XXXXX	XXXX	XXXXX
27. All student in my classes					
28. Male students					
29. Female students					
30. Minority students					
	0-10%	11-25%	26-50%	51-75%	76-100%
31. What percentage of your curriculum do you believe is aligned with the core content for assessment?					

Appendix B: Classroom Practice Measures for Pre and Long-Term Post

To what extent have you used the following types of technology in and/or for classroom instruction?	Never	1-2 Times a Year	3-4 Times a Year	5-6 Times a Year	Over 6 Times a Year
1. Water quality sampling kits.					
2. Labware, probes, CBLs, and graphing calculators.					
3. Internet websites for research and support materials.					
4. Microscopes.					
5. Videoscopes and/or display monitor.					
6. Presentation technologies (slides, power point, etc.)					
7. Digital camera.					
8. Geographic Positioning System (GPS)					
To what extent do you:	XXXX	XXXX	XXXXX	XXX	XXXXX
9. Use hands-on instructional strategies.					
10. Use inquiry-based teaching strategies.					
11. Address gender and minority equity through instruction.					
12. Integrate the sciences (physical, life, earth) in teaching.					
13. Integrate science as a subject with other subject areas.					
To what extent do you use the following community resources in your teaching:	XXXX	XXXX	XXXXX	XXX	XXXXX
14. Guest speakers (local, university, county, and/or state).					
15. Natural environment **field** sites related to watershed studies.					
16. Field trips to watershed related community resource sites (Museum, historical society, fish hatchery, farm site, etc.)					
To what extent do you incorporate the following types of field-based investigations in your teaching:	XXXX	XXXX	XXXXX	XXX	XXXXX
17. Water chemistry					
18. Macroinvertebrate study					
19. Habitat assessment					
20. Fish study					
21. Plankton collection					
To what extent do you teach:	Never	Rarely	Sometimes	Often	Always
22. about watershed topics.					
23. about connections between science & real life.					
24. connections between science & societal issues.					
25. connections between science & science-related careers.					

Learning Junior Secondary Science through Multi-Modal Representations

Bruce Waldrip
University of Southern Queensland

Vaughan Prain
La Trobe University

Jim Carolan
La Trobe University

Abstract

There is growing recognition that learning science in school entails understanding and linking verbal, visual and mathematical modes to develop knowledge of scientific concepts and processes. However, students face considerable challenges in engaging effectively with these literacies of science as they interpret and construct scientific texts. Our paper reports on two case studies on the topics of the particle theory of matter in Year 7, and force in Year 8. We aimed to identify (a) students' understandings of, and capacity to link, different representational modes to develop conceptual knowledge, and (b) teachers' perceptions of, and strategies to support, learning through this interlocking modal focus. Analyzed qualitative data included work samples, and focus-group interviews, as well as observations and interviews with participant teachers. The findings indicated that this multi-modal focus posed significant demands on learners, but had the potential to enable effective learning.

Correspondence should be addressed to Bruce Waldrip, University of Southern Queensland, Email: waldrip@usq.edu.au

Introduction

There is growing agreement in science education research that learning science entails learning the representational practices of this subject, including the reasoning processes, habits of mind, and rationale that underpin these practices. Science literacy is now understood as knowing how to interpret and construct these literacies of science (Norris & Phillips, 2003). From this perspective, learning scientific concepts and methods entails understanding and conceptually linking the purpose-built multiple and multi-modal representations of this domain (Ainsworth, 1999; 2006; Australian Academy of Science, 2005; Lemke, 2004; Gee, 2004; Russell & McGuigan, 2001). 'Multiple' refers to the practice of re-representing the same concept through different forms, including verbal, graphic and numerical modes, as well as repeated student exposures to the same concept. 'Multi-modal' refers to the linked use in science discourse of different modes to represent scientific reasoning and findings.

Given the increased use of new technologies to conduct and represent scientific activity in the science community and beyond, students' acquisition of this complex representational knowledge now poses very large challenges for effective classroom teaching and learning strategies. A key issue is to develop students' multi-modal thinking and reasoning in learning contexts that are consistent with current general principles of effective pedagogy for science learning. These principles emphasize the importance of catering for students' individual learning needs, preferences, and interests, and drawing effectively on students' current visual, verbal and numerical representational resources in acquiring the new literacies of science. At the same time, students need to be engaged actively with explanatory ideas and evidence that they can connect to real purposes and practices in their everyday worlds (Tytler 2003). This implies that student engagement with the key issue of how to represent science ideas and processes requires many iterations that are meaningfully contextualized and draw upon and expand their current repertoire of ways of showing what they know.

Much recent research on learning with representations generally, and in science in particular, has focused on identifying key design features of effective representations that promote successful student interpretation and learning (Ainsworth,1999, 2006; Schnotz & Bannert, 2003). The governing logic of this approach is that felicitous design features in representations can optimize student learning capacities. However, the highly complex nature of multiple representation environments poses many intractable questions for effective design. As noted by Ainsworth (2006) and others, design researchers are beginning to wrestle with many issues including the following typical questions. What number, type, style, and sequence of representations will maximize learning outcomes for different students? To what extent does brevity or redundancy of information in and across representations enhance learning, and under what conditions? To what extent do dynamic representations, such as spoken voice, animation, and dynamic graphs, enhance or impede interpretation of represented information when contrasted with static representations, and under what conditions? Are particular concepts better matched to particular representational modes, and how does the age and background knowledge of students affect learning outcomes? To what extent does interpretive constraint in a representation, such as graphic simplicity, help or hinder student understanding, and under what conditions? To what extent should science learning be focused only on domain-specific representations such as time graphs or cross-sections, or can learning be enhanced by including more domain-general approaches, such as visual displays and posters, and under what conditions and with what age or cultural groups might this mix be effective?

While various empirical studies have attempted to isolate and assess these different design options and sequences, and with mixed results, other research, including our own, has focused more on factors affecting students' own construction of scientific representations within mainstream classroom programs (diSessa, 2004; Prain & Waldrip, 2006; Russell and McGuigan, 2001; Tytler, Peterson & Prain, 2006). This research acknowledges that students must learn how to interpret science texts to achieve science literacy, but emphasizes a strong reciprocity between interpreting and constructing these representations. In constructing a science representation students are also involved in interpreting their own construction, its coherence and adequacy in representing their intentions and ideas, and the extent to which it will make sense to others, as well as its fit with the appropriate conventions for this kind of representation in science. We would also assert that students often need considerable practice in negotiating the construction of representational options, in order to understand in any depth the function and design of representational practices in science discourse. Students need to understand modal diversity in representations of science concepts and processes, be able to translate different modes into one another, as well as understand their co-ordinated use in representing scientific knowledge. While different classifications of these modes have been proposed, there is broad general agreement that these forms include such categories as descriptive (verbal, graphic, tabular), experimental, mathematical, figurative (pictorial, analogous and metaphoric) and kinaesthetic or embodied gestural understandings or representations of the same concept or process. There is increasing recognition that developing students' capacities to interpret and construct these complex science texts poses significant cognitive and pedagogical challenges.

In this paper we focus on case studies of teacher and student understandings and practices in engaging with multiple and multi-modal representations of the topics of force in Year 8, and changes to matter in Year 7 in mainstream classroom settings. As suggested above, our approach focussed mainly on lesson sequences where students were expected to construct representations of their own science investigations, drawing on their current representational resources supplemented by teacher guidance. We first review the theoretical framework and literature that guided our study's orientation.

Theoretical Framework of Study

The study was framed by current theoretical accounts of the nature of science discourse, learning as re-representation, and effective pedagogical conditions to promote student learning. These perspectives are viewed as compatible in that they link theories of science as a subject to how science can be learnt effectively, what should count as this learning, and broad socio-cultural factors affecting learning outcomes.

There is growing recognition that the discipline of science should be understood historically as the development and integration of multi-modal discourses (Lemke, 2004; Halliday & Martin, 1993; Kress, Jewitt, Ogborn, & Tsatsarelis, 2001; Norris and Phillips, 2003), where different modes serve different needs in relation to recording and integrating various kinds of scientific inquiry and reasoning. In this way, mathematical, verbal and graphic modes have been used individually and in coordinated ways to represent the knowledge claims of science discourse, with more recent technology-mediated representations of science consistent with, rather than a deviation from, this evolution of science as a discipline. By implication, students in the middle years of schooling need to learn about the multi-modal nature of the representation of scientific inquiry, and the different modes in which the same concepts in science can be represented as part of students' general development of science literacy.

Complementing this epistemic viewpoint, Ainsworth (1999) asserted that to learn from engaging with multiple representations of science concepts, students needed to be able to (a) understand the codes and signifiers in a representation, (b) understand the links between the representation and the target concept or process, (c) translate key features of the concept across representations, and (d) know which features to emphasise in designing their own representations. In this context, 'translation' means being able to recognise conceptual links between representations or invariant conceptual features across representations. Ainsworth (1999) posited that learner engagement with representations could support learning in three ways. These were (a) when the new representation complemented past understanding by confirming past knowledge, (b) when the new representation constrained interpretation by limiting the learner focus on key conceptual features, or (c) where the different representations enabled learners to identify an underlying concept or abstraction across modes or within the same mode of representation. This perspective is consistent with another account in cognitive science of the nature of learning as 're-representation' (Karmiloff-Smith, 1992). From this perspective, as noted by Russell and McGuigan (2001, p. 600), the developmental processes of student understanding involve the 're-coding of representations', implying that conceptual growth entails a process of re-representation, where learners generate and transform 'representations which are stored in different modalities, with meta-cognitive "explication" mediated by linguistic processes' (p. 600). From this perspective, learners use talk and other forms of representation to re-represent three-and two-dimensional experiences and records of understanding to clarify science concepts and procedures. This activity is also consistent with Paivio's (1986) theoretical account of the function and value of multiple coding in learning.

Our approach was also guided by current accounts of effective classroom pedagogy. A focus on representational diversity is consistent with recent calls for more student-responsive approaches to learning in the middle years of schooling (Gough, Beeson, Tytler, Waldrip, & Sharpley, 2002). Such an approach is viewed as likely to engage learners more than a traditional focus on restricted forms of representing scientific ideas evident in text books or usual classroom practices. This orientation is also consistent with recent research findings by Tytler and Waldrip (2002) that students learn most effectively in science, and engage more with the subject, where they are challenged to develop meaningful understandings, where individual learning needs and preferences are catered for, where a range of assessment tasks are used, where the nature of science is represented in its social, personal and technological dimensions, and where links are made between the classroom programme and the local and broader community that emphasise the broad relevance and social and cultural implications of science.

We considered that these broad theoretical orientations, in combination, provided an analytical framework for assessing factors affecting student learning in relation to representational choices and understandings.

Recent Research on multi-modal representations of concepts in learning science

Various studies have been conducted on student learning through interpreting and constructing different representational modes, including in primary classrooms (Russell and McGuigan, 2001) and in senior secondary physics (Dolin, 2001), with the use of some forms of representation researched in depth, (Glynn & Takahashi, 1998), such as the use of analogies for learning science (Coll & Treagust, 2001) and the role of scientific models in this process (Treagust, Chittleborough, & Mamiala, 2002). In asking primary school students to represent the same concept in different modes, Russell and McGuigan (2001) argued that the re-coding activity enabled learners to refine and make more explicit their understandings. In their classroom programme both student and teachers generated various representations of target concepts, and knowledge construction was viewed as the process of making and transforming these different representational modes, as they scaffolded their understandings in relation to their perceptions of the real world. Dolin (2001) noted that senior secondary physics students achieved enhanced understanding of concepts in physics when they attempted to translate different representational modes into one another in that subject.

Other researchers, such as Gobert and Clement (1999, p. 49-50), and van Meter (2001) have claimed that some modes may be more supportive of student learning than others, noting that students can 'draw to learn' effectively, where the visual media affords 'specific advantages over the textual media'. More recently, research in this area has focused variously on students' construction of self-explanation diagrams (Ainsworth and Iacovides, 2005), understanding concepts across multiple representations in different topics (Parnafes, 2005; Tytler, Peterson & Prain, 2006), and the role of visualization in textual interpretation (Florax & Ploetzner, 2005). Researchers in this field have also acknowledged the challenges learners face in constructing successful representations of science concepts. Ainsworth (2006, p. 186) noted that students needed to know how science representations encode information, including interpretive procedures, or 'operators'. They also needed to know how to construct an appropriate representation, in terms of its fit with the conventions of science discourse, including brevity, compactness, absence of ambiguity, and structural coherence, or systematicity. According to diSessa (2004, p. 298), "students start with a rich pool of representational competence" based on their past experiences with interpreting visual texts, and are 'strikingly good at … designing representations". He considered therefore that "rich and engaging classroom activities are relatively easy to foster " (p. 298) that are highly motivating for learners. However, like Gee (2004), Unsworth (2001) and others, he acknowledges that students also need to learn about the "sanctioned representations" (p. 294) of science, and justifiable strategies for their interpretation.

In a study of teacher perceptions in using multi-modal representations to support student learning in science, Prain & Waldrip (2005) noted that the teachers considered this focus could promote deeper learning, but was not easily accommodated within a tightly structured sequential learning process. Rather, teachers needed to respond flexibly to emerging learning opportunities and diverse student needs and capabilities. To succeed with this approach, students also needed to be familiar with the nature of the representational conventions in different modes in order to represent and translate concepts across modes. The teachers were aware that representations differed in their degree of abstractedness from, or visual similarity to the target concept, and that these differences posed further challenges for learners. While the teachers did not focus explicitly on these differences within individual representational modes with students, as recommended by Jewitt and Kress (2003), the teachers saw these differences as indicative of further complexities in choosing appropriate modes to enhance learning for students with different capabilities.

In summary, past research into an explicit focus on student engagement with specific representational modes and tasks has suggested the value and potential of this approach for promoting learning and for engaging a broad range of learners. Our study sought to investigate current teacher and student practices in relation to this negotiation of representational meanings. We also considered that this focus on opportunities for students to use multiple and multi-modal representations in the transitional years from primary to secondary education could meet the conditions for effective science learning, as outlined above by Tytler and Waldrip (2002).

Aims of Study, Research Methods and Context of Study

In this study we aimed to identify:

1. students' understandings of, and capacity to link, different representational modes to develop conceptual knowledge, and
2. teachers' perceptions of, and strategies to support, learning through this interlocking modal focus.

The study followed a mixed methods approach entailing collection and analysis of qualitative data (Cresswell, Tashakkori, Jensen & Shapley, 2003), including triangulation of different data sources to achieve convergence of results (Denzin & Lincoln, 1995). The research also included a case study approach (Merriam, 1998) that aimed to identify teacher and student perceptions in engaging with different representational modes. Initially, eight teachers were surveyed about their planning and usage of different representational modes. The survey preface indicated that "there is growing recognition that ideas can be represented in more than one way. These representational modes might include diagrams, cartoons, newspaper articles, photographs, written text, computer programs, images, models, analogies, drama, roleplay, acting out a process, data tabulations, numerical calculations, graphing, and posters". Teachers were asked about how they chose and used different representational modes to explain ideas, and what modes they might get students to make or use to engage with or show they understood ideas. The teachers were not told of modes the researchers might prefer. The surveys were analysed for patterns of common themes and differences in approach or emphasis.

From this initial survey, two teachers and their classes were selected for more intensive study of classroom practice. In this phase, lessons were observed, and interviews with teachers allowed insights into different pedagogic approaches from which it was possible to discern their views of learning and knowledge in relation to diverse modes of representation.

In this paper we report on two case studies, a unit on change of matter in year 7 taught by Meg, a teacher with over 20 years science teaching experience, and a unit on force in Year 8 taught by Barry, a secondary science teacher with over 30 years teaching experience. These units were taught in two regional Australian secondary state schools with students with predominantly low socio-economic profiles. Both units ran for eight weeks. Each teacher had participated in an in-service program with the researchers, that focused on the use of diverse representations, such as concept cartoons, and software programs to elicit and frame students' understanding of science topics. In devising the selection of resources and student tasks for each unit, the teachers collaborated with the researchers. Students' views and practices were also ascertained through classroom observations, surveys and transcription of group interviews of four students in Barry's class, four students twice in Meg's class. An earlier paper summarizes the teacher survey and some classroom observations (see Prain & Waldrip, 2005). Here we focus on both student and teacher perspectives and practices in the two classes studied.

Classroom Programs: The Unit on Change of Matter

Meg's main goal in this unit was for students to learn to recognize and utilize the application of a particle theory view of matter to a new situation/observation. She organized the unit into two phases, with the first five weeks including lesson content and strategies previously used by the teacher to develop student understanding of the application of 'Particle Theory' to the explanation of matter, its states, their properties, and transitions between states. In the last three weeks of the unit students were expected to use audiovisual hardware and software to develop a presentation linking different modal representations of particle theory to explain their observations of a laboratory demonstration. This was the major assessment task in conjunction with a standard test and several formative assessments during the teaching phase. Students had three 45 minute periods per week with two of these periods joined as a 'double'. Over the assessment phase, an extra two periods were utilized to allow sufficient access to ICT resources. The school Middle Years' structure encouraged teachers to teach one group for more than one subject area. The class was taught by the teacher for both science and mathematics. This allowed some flexibility in the provision of lessons and transfer of skills such as the use of spreadsheet software during the unit.

Meg considered the group "a strong class", with a balanced mix of girls and boys. In planning the assessment phase, Meg thought that sixty percent of the students were capable of attempting a conceptually demanding task. These were mostly girls who were subsequently observed to display very good communication, planning and organizational skills. Throughout the unit the particulate nature of matter was emphasized both verbally and by drawing student attention to diagrammatic representations on the board, and Meg also demonstrated the behaviour of particles using marbles on an overhead projector. Early in the unit students participated in a role-play enacting particle behaviour, such as the degree of attraction between particles, for one of the states of matter.

Meg began the lesson sequence with board notes for students to record in their books detailed notes on the scientific (textbook) understanding of matter, its states, their properties and transitions, and she introduced particle theory as an explanatory framework. This was accompanied by a 'brainstorm' activity where the students constructed a table of examples under the headings Solids, Liquids and Gases. The next lesson began with a practical investigation sourced from a Year 7 Science text book (Science Quest 1 Section 3.1). This involved students making observations of the properties of solids, liquids and gases and recording these in a table. The students also completed a 'silent card shuffle' activity where sets of representations (particle diagrams and written descriptions) matching each of the states were identified by students and pasted in their exercise books. In the next lesson students collected objects and recorded their physical properties. Meg believed that this activity was provided to give the students opportunity to begin using 'appropriate terms'. The students also watched a video titled 'What Matter is Made Of' and filled out a corresponding worksheet.

A single lesson was now dedicated to student observations of situations involving diffusion in different media. These were potassium permanganate crystals in water, bicarb soda in a Petri dish and eucalyptus oil scent in air. The teacher used this in the context of class discussion to emphasize verbally that particles were actually moving and so defined the new term and concept of ‘diffusion’. The learning sequence now turned to ‘change of state’. Students set up a situation where water was boiled in a beaker and condensed a ‘cloud’ below an ice-filled watch-glass set on top. Students recorded a labelled diagram of the apparatus and their observations. They then completed a picture of the “water cycle in nature” by labelling with terms such as ‘evaporation’ from a list provided.

Students then watched a video in the next lesson and completed a word ‘craze’ where they filled in the gaps in sentences with words provided, and answered questions regarding condensation and matched terms. In the next lesson students completed a comprehension activity where students read a description of the particle theory defining new terms and then answered questions requiring them to “apply particle theory to explain properties” (of matter). Students then used a lesson to complete activities from an interactive computer software program ‘Professor McClutcheon’ which the teacher noted “explains concepts well” including “visualizing particles”. This “reinforcement” activity required students to apply particle theory in order to work through the program. The students were provided with concept cartoons with the speech balloons blanked out. These portrayed two situations involving water boiling and students were required to complete the speech balloons to explain the situation to assess if they would employ the particle model without prompting. A practical skills activity followed where the particulate nature of the substances was not emphasized. The mass and volume of small objects were measured and entered into a software spreadsheet to determine density. The end of the teaching phase was punctuated by a ‘standard’ paper-based written test.

The learning/assessment task required students to apply a particle theory to a new situation using common laboratory apparatus, and providing an extended explanation via a multi-modal presentation. The choice of assessment aimed to facilitate a motivational function as well as provide adequate time for students to reflect on possible explanations in light of their recent learning. The students were introduced to different pieces of laboratory apparatus designed to provide examples of heat transfer, expansion, air pressure (vacuum) and diffusion. They were provided with a lesson to familiarize themselves with the correct operation of the apparatus and to make observations of its function and purpose.

They were then shown two examples of multi-modal audiovisual presentations made by older and younger students to explain concepts and processes. The former explained the approach to and solution of a mathematical problem, whilst the latter was a primary explanation of a thermometer. They were then encouraged to consider what combination of mode, content and explanation they would need to use to explain their observations of the apparatus previously explored. They were also encouraged to reflect on their recent learning to include a ‘theoretical’ explanation in their presentation. The students subsequently began collecting video and still images of their apparatus and developing verbal accounts. In the following lesson the student group was tutored in the use of software to facilitate the construction of a multi-modal presentation with a simple non-scientific example provided by the researcher. They were observed to have success with the use of the software, quickly demonstrating effective use of functions. The students then had approximately three lessons to develop their presentations with access to their apparatus, computer facilities, and video equipment.

Classroom Programs: The Unit on Force

This unit spanned approximately 8 weeks although with interruptions was described by the teacher as "4 weeks' work" because only approximately 14 periods of Science were taught. Students had three 45 minute periods per week with two periods of these joined as a 'double'. Barry's goal for the unit was for students to understand key concepts about force in relation to simple machines. The classroom contained computers which the students were able to use at their own initiation or the teacher's for learning/assessment activities. With the introduction of the new integrated curriculum in Victorian schools, the program structure focused on year level teams with paired form groups with two principal teachers teaching core subjects. The class was taught by the teacher for both Science (under the title of Trans-Disciplinary Studies) and Mathematics. This allowed flexibility in the provision of lessons and transfer of skills such as the use of spreadsheet graphing software during the unit. Barry considered the class as "better than average" without a "big bottom end" of struggling or disengaged students.

Barry introduced the topic over a few periods by watching the Honda Motor Car advertisement (http://www.youtube.com/watch?v=g2VCfOC69j) and looking at the work of U.S cartoonist Rube Goldberg. The students were then encouraged to produce their own "Goldberg style" cartoon. This was used by the teacher to facilitate "one to one discussion" with students and to encourage them "to look at movement and force ". Some students requested to produce their 'poster' by animating pictures using PowerPoint. The teacher later utilized this skill base in other learning and assessment tasks. He provided a number of worksheets to the students during the unit and instructed them on the recognition of simple machines, associated terminology (fulcrum, load, inclined plane etc), their taxonomy (class of levers) as well as the identification and labelling of direction and extent of forces and motions.

The students attempted one of a range of practical activities covering each of the major simple machines and focusing on the central concept of work. This was structured as a "jigsaw" activity designed so that individuals made measurements and observations of force and distance moved in a particular simple machine. To facilitate student re-representation of the work concept they were asked to record their learning on a worksheet in a table, written sentence, labelled diagram and mathematical equation. The teacher observed that this activity did not "work as well as I'd like it to", citing problems with students following some of the practical instructions and with the need for greater scaffolding prior to the task.

The class watched and responded to a number of video presentations (Stansfield & Boiteau, 1981) to introduce and reinforce simple machine, mechanical advantage and work concepts. Students had access to the internet via classroom computers which allowed access to simple machine websites to reinforce concepts previously covered. Sites included:

http://www.mos.org/sln/Leonardo/GadgetAnatomy.html Recognition of simple machines (in complex machines) activity –10 minutes;

http://www.mikids.com/Smachines.htm Naming simple machines e.g. propeller as an inclined plane. –15 minutes;

http://www.edheads.org/activities/odd_machine/index.htm Covers some forces revision as well as identifying and using simple machines. –15 minutes.

The whole class also had a session using the interactive website *Aspire – Simple and Complex Machines* (http://sunshine.chpc.utah.edu/javalabs/java12/machine/index.htm). This enabled students to conduct virtual experiments measuring force and distance for simple machines (wedges, lever, ramps, pulleys, inclined planes, screws, wheels and axles). Students then entered these measurements in a spreadsheet to calculate work done and note the conservation of work in each example.

Barry aimed to assess students' theoretical and applied (pertaining to actual "measurements") understanding of the simple machines concepts though two pieces of assessment. In pairs, students were provided with a word list which, after watching a general machine video (Beeston & Maude, 1997), they used to construct a sentence or sentences describing or explaining some aspect of simple machines. These were assessed for individual understanding of simple machine concepts such as the conservation of work. Following this, students were provided with an example of a compound machine such as a can opener, shifting spanner, or lever-action cork remover. They were instructed to observe closely the machine to identify each feature, how these worked, and measure the direction and extent of motions and forces to determine the mechanical advantage. They were then provided with an option of presentation formats to report and explain their observations, either, according to their preference, as a poster or using PowerPoint to animate diagrams.

Findings

As the summary accounts of each lesson sequence above make clear, students in each class participated in a diverse range of interpretations and constructions of multiple and multi-modal representations of science concepts and processes. These included group and whole class talk about different aspects of the topic, interpreting teacher notes and diagrams, re-representing three-dimensional practical activities in two-dimensional formats, making sense of video and other resources used to supplement classroom activities, participating in virtual web-based experiments using tables and graphs, interpreting in written language key concepts in a concept cartoon scenario, enacting understanding of concepts with physical actions and roleplay, and constructing their own multi-modal two-dimensional representations of practical investigations. The interview data, observations and examples of students' work, were analysed to identify major episodes of interactions; fine-grain analysis of interview transcripts within these interactions; and recomposing these smaller analyses into patterns to create assertions as to what are students' perceptions of the roles, forms and interplay between different multimodal representations in science classes. These assertions form the basis for identifying conditions and strategies that maximise the learning outcomes of this approach. Given also the diversity of student background knowledge and interests in science, it is very difficult to ascribe particular learning outcomes to specific representational work within these mainstream classroom programs. Clearly, too, the teachers in each lesson sequence understood effective learning opportunities as a re-iterative process whereby students re-visited negotiation of the meaning of concepts in different representational forms and across different contexts.

In the light of these complexities in the learning environment in relation to representation, and diverse contextual factors influencing learning outcomes, our reporting of findings focuses on indicative general teacher and student perceptions of learning effects rather than claims of tight causal links between an example of a representational construction or interpretation and a learning outcome. We draw on two specific examples of student work, as well as student and teacher reflective comments on this work, as indicative of general effects of a representational focus rather than as conclusive evidence of causal effects.

In the first student work example, three Year 7 students produced a Powerpoint to represent diffusion of particles of scent throughout the classroom. At this year level, students were expected to begin to use a simple particle model to relate the properties, behaviours and uses of substances to their basic material structure. In the Powerpoint, the students have clearly adopted a particulate view of matter as displayed by the diagrammatic representation both in the first and fourth slides. The latter also demonstrates that the students have a sound basic understanding of the properties of a gas, including particles filling a closed container, and the random spread of particles. The written component of the description reflects recognition of the importance and action of the role of forces between particles in determining behaviour of the substance, where weak force means low attraction so particles spread out as a gas. It also shows an understanding of diffusion as occurring ostensibly from areas of higher to lower concentration until equilibrium is reached. The written component could be seen to imply but not directly express the idea that the smell is the gas particles. The students may still consider that the non-particulate smell is somehow being carried by the gas particles and this may warrant further probing, perhaps verbally and/or by encouraging the students to refine further their explanation to achieve clarity of expression. The written account also mentions liquid evaporating but does not reflect this in the diagrams presented, and expansion here should further reveal the students' deeper understanding.

In the second work example, two Year 8 students produced a Powerpoint of a corkscrew opener to show understanding of force. In this unit students were expected to recognize and explain how mechanical systems can direct and modify force and motion. They were expected to identify simple machines such as pulleys, gears, levers and inclined planes, and describe their action in producing a mechanical advantage. The students in this work example have been able to break a compound machine down clearly into its component simple machines through investigation of its actions. They have been able to represent them both separately and in combination through clever use of animation showing their action in context. The students through their measurements have been able to recognize the source of the mechanical advantage of the machine as being gained at the expense of greater movement of the lever compared with the cork. Later they demonstrate an understanding of the mathematical relationship between effort and load and their distance from the pivot point for a lever. The students have not overtly expressed the benefit gained from the machines, such as the idea that less force is needed to remove the cork or top, nor mentioned mechanical advantage or the concept of work done, despite having covered these in previous learning activities. Ideally for a summative assessment this recognition or understanding should be expressed. However, Barry considered that the students understood this, based on his informal observations during class.

Student views on multiple representations

Meg and Barry had not focused on making explicit to their students why each unit entailed making multiple representations as a major part of assessment. As a consequence, when asked about this in subsequent interviews, the students did not have a strong sense of the purpose or rationale for this approach. However, in these interviews, students gave thoughtful accounts of how they constructed their representations, factors that affected their decision-making, and some sense of how this approach supported their learning. Students commented on the value of the role play in Meg's class, crediting this activity with giving them "a better idea of what they (the written descriptions) meant", and also noting that "doing it made you think about it - about what you are doing". They also commented on the 'silent card shuffle' activity where they identified sets of representations (particle diagrams and written descriptions) matching each of the states and pasting them in their exercise books. They saw this work as valuable in consolidating their understanding.

In the following interview segment, a Year 7 student in Meg's class explained the process her group used to construct the Powerpoint animation to represent the diffusion of scent as particles throughout the classroom:

Student: *With ours (the diffusion of scent) you couldn't take many pictures because you obviously can't take pictures of particles, so we just got the little circles and made shapes to show the jar thing and how they travel if they are let out*.

Researcher: *Did you start out with your pictures and then go to the words?*

S: *We started off with words and then we did the front heading (first slide) which had all the particles moving around, and then we did writing, and at the end Lauren thought of that picture which was really good and explained it more.*

R: *So with you, personally, what did you start thinking about, now that you have got this feeling about everything being particles. Were you seeing it as a picture or were you thinking of it as a spoken description?*

S: *Yeah, I was thinking of it, visualizing it, but with the computer it really helped with all the pictures.*

R: *So was it hard to go from that to the written part?*

S: *Yes, it was so hard because … all the other people got to do experiments where you could see actually what was happening, but with ours you could just sit it there (the jar of scent) and you had to wait for the smell to travel and it was really, really hard because you can't take pictures or videotape it.*

While these comments are mainly descriptive of the choices involved in the construction process, they point to some deeper understanding of the need to change representations to show understanding of the idea of particles in this particular context. In her responses the student recognized that the animation had to re-represent a process that could not use other obvious forms of evidence. In commenting further, she said "We thought we'd just have to write about it so we got the idea…of people putting up their hands when they could smell it (diffusing across the room) so we made a video of that".

This student also thought that the task would be easier if the teacher provided more explicit guidance on how to structure the representation. While the class was shown some unrelated examples of student-produced multi-modal texts, the students wanted additional guidance in "how do we do that", including both technical understanding for making the product, and advice on format and focus in linking modes, even though their products demonstrated a capacity to make modal linkages. In the case of the diffusion animation, visual and written texts were linked, and the students demonstrated knowledge of how to represent a process and passage of time through a three-step set of diagrams. The students who made the multi-modal representation of the animation of forces involved in the operation of a corkscrew opener linked visual text, labelled diagrams, measurement of force, arrows to indicate the direction of force, and animation of the stages of the process. They also used the convention of representing the corkscrew in a frontal perspective level with the viewer, thus emphasizing the objectivity of the representation. They also simplified the representation to highlight key aspects of the machine in a style typical of traditional labeled diagrams in science texts. In these ways both groups of student demonstrated knowledge of some of the conventions and aims of science texts relating to clarity and coherence of the representation.

In a follow-up conversation with a student from Barry's class, one of the two boys who worked on the Powerpoint of the corkscrew opener's action, the researcher asked what the student tackled first in constructing this representation:

Student: *Oh, the diagrams.*

Researcher: *And why was that?*

S: *Well I like to have a visual type thing to see it.*

R: *So did you have to talk much about your diagrams?*

S: *Just debating it, on how sort of accurate we were going to make it.*

R: *When you went from transferring it from your diagram to your writing was that hard?*

S: *When we first started it (examining the tool) I wrote down stuff in my book, just taking notes about what each thing did and we basically elaborated on that.*

R: *Did you have much discussion about what you wrote down because it seemed from before (when viewing the presentation) you picked it up and said that your partner might have written something down wrong? Why was it important for you that he might have changed that to make it a little bit different?*

S: *I'm kind of an individual kind of worker and then it's good to work with someone. It's all right, I mean it's both of our things (work).*

R: *So you had to negotiate what the end product was?*

S: *Yep.*

R: *Did you have to discuss whether it was right or wrong, the different things you put in there?*

S: *Oh yes. Just some of the things Chris pointed out that I hadn't put in that we needed to point out.*

These responses indicate the student is aware of the need for text and diagrams to cohere, and the need to check with a partner about the accuracy and clarity of the product. This student also commented that having to explain the text to a partner and other students helped him learn because "you have to learn more about it if you don't know enough about it".

A group of students in Meg's class were asked why she encouraged them to use a software program for their assessment task that enabled them to represent different modes such as video, diagrams, written text, and graphs on the same screen. Their responses indicated that they saw value for their learning in constructing this text:

S2: *Probably because we have to explain it ourselves and instead of using really big words to explain we have to use little ones that you can understand.*

S3: *Yeah, we have to think about it ourselves, like you can help us do it, but we have to try and do it ourselves and think about it.*

Researcher*: So why didn't we just give you a piece of blank paper and say write about it? How was what you did different from that …. and the test you were given which was pretty much just writing?*

S1: *You could definitely visualize with all the experiment stuff which was a lot easier, so you could write more about it, about what happened, instead of just writing about it.*

S2: *This doesn't feel like a test if you give us a piece of paper and say start now.*

S1*: All tense and everything.*

S3: *I think it helped us learn as well because like you had to think about it more rather than just writing it down. We went more into it and used it a bit better.*

Teacher views

Both Meg and Barry saw a range of positive effects in student engagement with multi-modal representations. Both teachers considered that asking students to design their own representation of what they had learnt about the topic was very motivating, but claimed students needed a template or finished product to guide their work. While the teachers thought that student exposure to a model product directly on the topic would only encourage mimicry rather than deeper learning, they claimed that students often need a framework or instructional model to guide their thinking. Barry considered that the lack of scaffolding and instructional support had undermined the effectiveness of the jigsaw activity of different representations of the same concept. On the other hand, with able students, he considered that the larger range of representational resources allowed strong students to produce outstanding work. In commenting on the work of the pair of student who produced the Powerpoint animation of a corkscrew opener's action he made the following observations:

> *It's the depth of learning that the kids are working at. It's not just engagement. That's what I reckon we really got out of this. With the good kids, particularly with those two boys who did the corkscrew, they are really bright kids, that's a pretty special bit of work. It demonstrates their understanding, and they've sat there and really enjoyed it. It's really quite deep understanding to put that together. They've had to talk quite deeply about it.*

Barry noted that these students were intently focused on making accurate observations for their representation, and were willing to tackle a multi-modal representational task beyond any task he had envisaged. He also considered they could demonstrate more knowledge through this format than through paper-based testing:

> *With the Powerpoint it led to that (the Powerpoint animation). I would never have thought of ever doing that. And they went with that. They can create the movement and show things. I found I wasn't limiting them to my limits, letting them go with what they could do.*

A key aspect of assessment for Barry was informal monitoring of student group discussion which he considered offered rich evidence of the level of students' conceptual understanding. He claimed that listening to students' discussion on how to represent their investigation strongly revealed both their practical and theoretical understanding of the topic.

Meg considered that students' work on constructing a multi-modal representation of an investigation was strongly motivating, and produced more effective learning than text-book reading, note-taking, or other strongly teacher-directed activity:

> *They will definitely have a better understanding of particles, because that was pretty much covered (in their Powerpoints), the part of the topic that deals with effect of heat on matter, and they will certainly have better knowledge, and I would suspect they would do better on their test. They would do better, because there was a focus on what the particles are actually doing.*

However, she also claimed that some students also required a model or a template to guide their work on their representation, and that she had to provide extensive support for some student groups. Reflecting on the broad effects of an explicit representational focus, Meg considered that it could enhance student learning:

> *If it became part of their learning, part of the way they did science, to come to expect that today we are going to try and demonstrate our understanding of a concept by use of a diagram, or next time by talking about it, then they would develop skills that would help, and the next topic there would be skills they could transfer to the new topic, assuming they are going to use a similar range of representations.*

Implications

These findings have various implications for developing students' understanding of the literacies of science. Teacher and student comments indicate that a major challenge entailed in student engagement with multi-modal representations is the question how much guidance and explicit teaching of representational conventions should be undertaken. The student work produced in each case study offers strong support for diSessa's (2004) contention that students are likely to have rich meta-representational competence in understanding aspects of the nature, purpose and preferable features of science texts, based on past experiences in science classes and previous experiences with many kinds of visual texts that aim to explain spatial and/or temporal relationships. This study suggests that the teacher needs to ascertain what students collectively and individually understand about these features, and then provide timely and relevant practice in representational tasks appropriate to the topic under investigation. For example, students need to have some knowledge about the purpose of graphs, their typical structural and functional features, and the operators that enable interpretation of graphs, before they are expected to represent their understanding of a topic in graphic form. As noted by diSessa (2004), there is some degree of developmental predictability in student acquisition of meta-representational competence, but teachers need to be responsive to the needs and knowledge of their students in framing representational challenges.

A further implication is the question of how much variation from convention or conceptual accuracy should teachers accept or tolerate in student-generated representations. Student products in both cases in this study, and illustrated in the two pieces of group work, indicated a tendency to experiment with the expressive possibilities of technological texts in ways that diverged from the typical conventions of science discourse. Students also might use a representational mode that is not suited to achieving precision or clarity in showing understanding, or they might construct a representation inaccurately. As evident in the responses of Meg and Barry to their students' work, there is a need for teacher to be keenly informed readers and viewers of their students' products and intentions. There is a need for a strong focus on the conceptual accuracy of student work, but also to guide the students towards standard practices. However, as Lemke (2004), diSessa (2004) and many others have noted, scientists are always designing new representations, especially as new technologies continue to change the conventions for how science is conducted and reported, particularly in net-based texts. As diSessa (2004) points out, the quality of a representation should always be judged by its purpose. While students are expected to learn that the science community highly values systematicity. brevity, completeness, and absence of ambiguity in scientific representations, students need to be given scope to learn through own experimental design work the aptness of these meta-representational values. This suggests that teachers need to exercise some degree of flexibility in the prescriptions they might offer students about appropriate conventions to draw upon as they interpret and construct science texts. It is probable that these teachers need to work with students developmentally along a continuum of skill competence with the conventions themselves.

The findings of this study complement current research on maximizing the effectiveness of designed representational environments by focusing on the need to take into account the diversity of learner background knowledge, expectations, preferences, and interpretive skills. The procedures that students use in constructing their own multi-modal representations, and the developmental pattern to these procedures (diSessa, 2004), provide insight into design features that could be explored in effective teaching representations for different age groups. There is a need for more classroom-based research on this interface between student- and researcher-designed representations, as undertaken by Russell and McGuigan (2001), and Dolin (2001).

Teacher and student perceptions of factors affecting learning in the two case studies reported in this paper are too diffused to offer any strong confirmation of Ainsworth's (1999) theory of how interpreting multi-modal representations enhances learning. As Ainsworth (2006) recently noted, the complexity of the cognitive tasks learners face in translating effectively across modes, and the effects of context, student knowledge and purposes on this translation work, are only starting to be appreciated fully in research in this field. However, the teacher and student responses suggest generally that designing multi-modal representations of science concepts enables learners to construct a deeper understanding. In the case of the Powerpoint on diffusion, the student's comments indicate that she understood the function of particles across different representations, implying an increasingly abstracted understanding of this conception of matter. However, further representational work with different applications would be required to confirm this point. Our study suggests that there is a need for more research that focuses in appropriate detail on students' perceptions and specific strategies in relation to this design and translation work across modes in particular science topics.

As noted at the outset of this paper, effective pedagogy in science must entail engaging students' interest and enhancing their perception of real-world applications in their learning. Our findings confirm there are potentially strong motivational gains in providing guided opportunities for students to design explanatory representations of their conceptual understandings, but this still leaves open the question of how these or other new representations might enable learners to connect what they have learnt with the world beyond the classroom.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*, 131-152.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16 (3), 183-198.

Ainsworth, S. & Iacovides, I. (2005, August). Learning by constructing self-explanation diagrams. Paper presented at the 11th Biennial Conference of European Association for Research on Learning and Instruction, Nicosia, Cypress.

Beeston, S., & Maude, R. (1997). *Scientific Eye: Physical Processes 3* (Machines and Moments) [Videorecording].London: Channel Four Learning.

Coll, R. & Treagust, D. F. (2001). Learners' mental models of chemical bonding. *Research in Science Education, 31*(3), 357-382.

Cresswell, J.W., Tashakkori, A., Jensen, K.D., & Shapley, K.L. (2003). Teaching mixed methods research: Practices, dilemmas and challenges. In A. Tashakkori and C. Teddlie (Eds.). *Handbook of mixed methods in social & behavioral research* (pp.619-638). Thousand Oaks, CA: Sage Publications.

Denzin, N., & Lincoln, Y. (1994). *The handbook of qualitative research in education.* Newbury Park, CA: Sage Publications.

diSessa, A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction 22*(3) 293-331.

Dolin, J. (2001). Representational forms in physics. In D. Psillos, P. Kariotoglou, V. Tselfes, G. Bisdikian, G. Fassoulopoulos, E. Hatzikraniotis, E. Kallery (Eds.). *Science Education Research in the Knowledge-Based Society. Proceedings of the Third International Conference of the ESERA* (pp. 359-361). Thessaloniki, Greece: Aristotle University of Thessaloniki.

Florax, M. & Ploetzner R. (2005, August). Effects of active integration of texts and visualization in learning. Paper presented at the 11th Biennial Conference of European Association for Research on Learning and Instruction, Nicosia, Cypress.

Gee, J.P. (2004). Language in the science classroom: Academic social languages as the heart of school-based literacy. In E.W. Saul (Ed.) *Crossing borders in literacy and science instruction: Perspectives on theory and practice.* Newark DE: International Reading Association and National Science Teachers Association.

Glynn, S. M., & Takahashi, T. (1998). Learning from Analogy-Enhanced Science Text. *Journal of Research in Science Teaching, 35*, 1129-1149.

Gobert, J. & Clement, J. (1999). The effects of student-generated diagrams versus student-generated summaries on conceptual understanding of spatial, causal and dynamic knowledge in plate tectonics. *Journal of Research in Science Education, 36*, 39-53.

Gough, A., Beeson, G., Tytler, R., Waldrip, B. & Sharpley, B. (2002). *Improving effective science teaching and learning within Australian Schools.* A paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans.

Halliday, M & Martin, J.(1993). *Writing science: Literacy and discursive power* (London: Falmer Press).

Jewitt, C., & Kress, G. (Eds.). (2003). *Multi-modal literacy*. New York: Peter Lang.

Karmiloff-Smith, A. (1992). *Beyond modularity. A developmental perspective on cognitive science.* Boston, MA: MIT Press.

Kress, G., Jewitt, C., Ogborn, J., & Tsatsarelis, C. (2001). *Multimodal teaching and learning: The rhetorics of the science classroom.* London, UK: Continuum.

Lemke, J. (2004). The literacies of science. In E.W. Saul (Ed.) *Crossing borders in literacy and science instruction: Perspectives on theory and practice* (pp. 33-47). Arlington, VA: International Reading Association/National Science Teachers Association.

Merriam, S, (1998). *Qualitative research and case study applications in education* (2nd Edition). San Francisco: Jossey-Bass.

Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education, 87*, 224-240.

Paivio, A. (1986). *Mental representations: A dual-coding approach.* New York: Oxford University Press.

Parnafes, O. (2005, August). Constructing coherent understanding of physical concepts through the interpretation of multiple representations. Paper presented at the 11th Biennial Conference of European Association for Research on Learning and Instruction, Nicosia, Cypress.

Prain, V, & Waldrip B. (2005). Teachers' perspectives about using multi-modal representations of concepts to enhance learning. Paper presented at the European Science Education Research Association Conference (ESERA), Barcelona, Spain, August 28-September 1.

Prain, V, & Waldrip B. (2006). An exploratory study of teachers' and students' use of multi-modal representations of concepts in primary science. *International Journal of Science Education, 28*(15), 1843–1866.

Russell, T. & McGuigan, L. (2001). Promoting understanding through representational redescription: an illustration referring to young pupils' ideas about gravity. In D. Psillos, P. Kariotoglou, V. Tselfes, G. Bisdikian, G. Fassoulopoulos, E. Hatzikraniotis, E. Kallery (Eds.) *Science Education Research in the Knowledge-Based Society. Proceedings of the Third International Conference of the ESERA* (pp. 600-602). Thessaloniki, Greece: Aristotle University of Thessaloniki.

Schnotz, W., Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning & Instruction, 13* (2), p141-156.

Stansfield, D., & Boiteau, D. (1981). *Eureka Series* (Gravity, Weight vs Mass, Work) [Videorecording]. Ontario, Canada: TV Ontario.

Treagust, D., Chittelborough, G., & Mamiala, T. (2002). Students' understanding of the role of scientific models in learning science. *International Journal of Science Education 24*(4), 357-368.

Tytler, R. (2003). A window for a purpose: developing a framework for describing effective science teaching and learning, *Research in Science Education, 33*, 273-298.

Tytler, R., & Waldrip, B.G. (2002). Improving primary science: schools experience of change *Investigating, 18,* 23-26.

Tytler, R., Peterson, S. & Prain, V. (2006). Picturing evaporation: Learning science literacy through a particle presentation. *Teaching Science, 52*(1), 12-17.

Unsworth, L. (2001) Evaluating the Language of Different Types of Explanations in Junior High School Science Texts. *International Journal of Science Education, 23* (6), 585-609.

Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology, 93* (1), 129-140.

Designing and Using Program-specific Evaluation Instruments

Leslie Garrison
San Diego State University

Olga Amaral
San Diego State University

Abstract

This paper describes the development and use of a program-specific evaluation instrument designed to measure the impact of professional development on classroom practice. The evaluation rubric describes a sequence of skills or proficiencies that guides teachers and professional developers toward improved practice. The design and implementation of the instrument: defines the program instructional targets, guides professional development, makes target instructional skills explicit to teachers, and aligns program evaluation with program instructional targets.

Introduction

The U.S. Department of Education, *No Child Left Behind* legislation, the National Science Foundation, and state Departments of Education have all called for the implementation of research based instructional practices in the programs they fund. In addition, funded programs are often required to do formal evaluations that expand research in the field. To this end, evaluations must include empirical evidence of the program's impact on classroom practice and student learning.

Few instruments are available that document changes in teacher practice from professional development activities. The generic measures of teacher practice which can be found, are seldom targeted on the specific goals of the professional development. For example, the Observation and Analytic Protocol, developed by Horizon Research (Weiss, Pasley, Smith, Banilower, & Heck, 2003) contains an impressive list of 29 aspects of classroom instruction. While all of these characteristics are important in lesson delivery, the list does not necessarily match with the specific training focus of an individual program. This paper provides an example of how an instrument, based on the key elements of the professional development, was designed and used for program formative evaluations.

The context for this project was a science reform initiative in the Imperial Valley of southern California. The science program there has been shown to be effective in providing professional development that leads to student achievement (Amaral, Garrison & Klentschy, 2002; Klentschy & Molina de la Torre, 2005). Researchers from San Diego State University's Imperial Valley Campus have worked with the science reform project to develop the Teacher Behavior Continuum, an instrument targeted at classroom practice (Table 1). This instrument is used to document teacher instructional practices when conducting inquiry-based science lessons and is aligned to the goals and objectives of the project.

Creating a Common Vision

Increasing the pedagogical content knowledge of teachers (Shulman, 1986), or in other words, the skills necessary to effectively teach science, comprises the focus of most professional development in the project. Pedagogical content knowledge in science has been described more broadly (Morine-Dershimer & Kent, 2002) as not only pedagogical knowledge and content knowledge, but also including knowledge of the learners, knowledge of the curriculum, and knowledge of the educational goals of the course.

Reform movements often have at least three groups that work together to improve student performance: program designers, staff developers, and teachers. Program designers envision the reform, write the grants or otherwise secure funding, as well as set the goals and activity framework. This vision is passed along to staff developers who, in turn, interpret the language of the reform, into professional development activities. They most often recruit participants for the professional development, as well as conduct training sessions and follow-up activities with teachers. The teachers and recipients of the professional development activities interpret what they have learned through individual lenses, those which embed their background experiences, classroom practices, and knowledge of their students. In short, they translate reform ideas between the professional development activities and their students. By the time the initial vision reaches the final stage of interpretation and implementation at the classroom level, it may have become almost unrecognizable from the original vision. The challenge for reform programs, therefore, is to establish and maintain a common vision among the three major groups to ensure unified goals and outcomes.

Establishing a common vision is best achieved by involving all three groups from the planning stages, including the development of the vision, goals, and activities. The vision can be formed by defining the goals of the project in terms of target teacher behaviors that can be observed in the classroom. Once the target behavior has been defined, levels of performance that build toward the target behavior can be established. The behavior descriptions for each level must be sequenced in ways that accommodate the novice teacher, that is, the instructional practice of the teacher for whom the target behavior is a new concept, and to the experienced teacher who is fully integrating the target behavior in daily practice. These successive levels building toward the target behaviors comprise the Teacher Behavior Continuum (Table 1).

Uses of the Behavior Continuum

Once established, the Teacher Behavior Continuum plays a central role in program design, implementation, and evaluation. First, it creates a common vision which all participants can support and strive to implement. This vision provides a focus and direction for the program leadership and drives orientation and activities for staff development. It can also be used to inform teachers about expected goals, serving as a self-assessment for teacher practice, as well as a vehicle to guide reflection and discourse of pedagogical content. It helps make concepts about the instructional process more concrete for practitioners. Finally, it provides a template for program evaluation, one that allows for ongoing assessment of implementation of strategies over time.

Case Study: Development and Use of the Teacher Behavior
Continuum in Science Reform Efforts

Science reform efforts started in Imperial Valley in 1996 with assistance of a National Science Foundation Local Systemic Change grant. During the ensuing five years, the Valle Imperial Project in Science (VIPS) trained teachers on the implementation of inquiry-based science instruction. A cadre of five science resource teachers supported the change initiative. The National Research Council talks about how the responsibility for inquiry science instruction is shifting from teacher to student resulting in shifts in inquiry lessons from guided to open. (National Research Council, 2000). The Teacher Behavior Continuum employs a five-point rubric that reflects this transition. The ability of teachers to carry out this transformation in their practice depends on how effective the professional development experiences are in advancing their pedagogical content knowledge. The Teacher Behavior Continuum has proven to be an effective tool for the VIPS leadership team as they continue to develop the target skills among teachers by focusing on Lesson Design training.

Table 1:
Teacher Behavior Continuum Video Analysis Instrument

STRAND	I	II	III	IV	V
Lesson Focus	Tasks, intent and purpose of the lesson are unclear	Tasks made clear but not the intent or purpose of the lesson	Lesson tasks and intent are clear but not set within a larger frame	Some linkages are made between the current activity and the Big Ideas of the unit	Lesson tasks and intent is clearly evident within the key concepts of the unit
Student Engagement	Many students not actively engaged in the lesson	Engages most of the students to participate	Engages nearly all students to participate at various point in the lesson	Most students engaged physically and intellectually in the lesson	Engages nearly all students physically and intellectually to contribute consistently throughout the lesson
Data, Claims & Evidence	Teacher doesn't require and/or provide direction for data collection	Teacher requires data collection but without sufficient student support	Teacher monitors and guides students to clearly and accurately record data from the lesson	Teacher ensures that students record data clearly and accurately and can interpret data	Teacher ensures that students record data clearly and accurately, can interpret data, and relate findings to the key concepts
Discourse/ Discussion	Teacher talks, students listen	Teacher engages students in procedural and management discussions	Teacher asks students fact based questions about what they did and found in the lesson	Teacher poses questions to develop student thinking that begin to link the lesson to the key concepts	Teacher poses questions that connect lesson to key concepts and requires students to explain their responses with clear lines of evidence
Closure/ Conclusion	Lesson ends without closing activity	Procedures reviewed to handle and put away materials	Lesson's activities and findings were reviewed. Teacher directed all students to the same conclusion	Lesson's activities and findings were reviewed with some reference to the key concepts	Lesson's activities and findings were reviewed and tied to lesson intent, purpose, and key concepts
English Language Development	Vocabulary assistance needed but not addressed	Vocabulary defined only after student request	Vocabulary presented verbally or in written form with no elaboration	Vocabulary introduced with context using active student participation	Vocabulary defined and integrated throughout the lesson and students required to use new words
Leslie Garrison and Olga Amaral					

Lesson Design

Lesson Design is a process developed by VIPS and instructors from California Institute of Technology to identify central ideas within science units that students need to know. The first step in this process is the identification of lessons that are essential to maximize student understanding of concepts in science. The second step is to determine the areas within each lesson that can lead teachers to maximize instruction and convey the "big idea" of science. In addition, it was important to develop strategies that were based on scientific processes. The habits of science are defined by Connolly (1989) as the collection of data, suspension of premature closure, and the ability to make claims based on data and evidence. Building on this premise, the team decided to include the following elements for each lesson: Lesson Focus, Student Engagement, Data, Claims, and Evidence, Discourse/Discussion, and Closure/Conclusion. Teachers receive guidance on the use of each element during professional development sessions. Trained support personnel are also available to provide ongoing classroom support. The challenge, then, is to determine the level of impact of each element in target classrooms.

The Elements

Each element is itself rooted in research. For example, the effectiveness of embedding classroom discourse in science instructional practices is well established in the literature. Teachers need to improve the intellectual level of tasks as well as the level of student communication and discourse (Ruiz-Primo, Li, & Shavelson, 2002). Rowell and Ebbers (2004) recognized three different categories of explanations developed in discourse: descriptive and relational explanations and explanatory models. Newton and Newton (2001) reported that there are certain types of oral discourse that can improve learning. Other elements have also been the focus of many studies: student engagement (Herrenkohl & Guerra, 1998; Brophy & Good, 1986); lesson focus (Garrison, 2004; Williams and Sternberg, 1993); and closure/conclusion (Baxter, Bass, and Glaser, 2000). An understanding of the research in the field provided the foundation upon which the program was built.

Guiding Question

The question that guided the work of the VIPS reform effects was: How does the professional development in Lesson Design impact the delivery of classroom instruction for grade 4-8 classrooms? The reform movement set out to help teachers implement lessons which were clear in focus, inquiry-based, enriched with discussion, inclusive of English learners, and engaged student thinking. The development of the Teacher Behavior Continuum helped the leadership team reach a common understanding and definition of the target strategies in that it required the group too not only operationalize each strategy, but to describe the sequence of skills that led toward its implementation. Videos of science lessons were made and evaluated for the presence of target instructional elements that promote student thinking and learning. The evaluation of the professional development centered on the impact of Lesson Design training on teacher practice.

Instrument

Teacher Behavior Continuum for Video Analysis

An instrument was designed for use in the evaluation of professional development relative to Lesson Design. The Teacher Behavior Continuum (Table 1) was developed collaboratively by a team of program designers, program evaluators, staff developers, and classroom teachers. The initial task before the group was to identify the central elements of lesson design and lesson delivery that the program aimed to establish. Specifically, what were the target instructional areas addressed by the program; and what specific teacher behaviors did the program seek to establish?

The areas for analysis on the rubric were agreed upon collaboratively among the program researchers, evaluators, and program personnel. These areas were: Lesson Focus, Student Engagement, Data, Claims, and Evidence, Discourse/Discussion, Closure/Conclusion, and English Language Development. Initially each area was discussed and described in a narrative form, which defined in general terms the behaviors that the group felt best described it. A summary of the groups' thoughts follows.

Lesson Focus: At the beginning of the lesson, the teacher needs to provide students with the opportunity to be clear about the tasks of the lesson as well as its intent. Students need to understand both what to do and why they were doing it.

Student Engagement: To learn, students need to be engaged in the lesson. Engagement has two parts, physical engagement and intellectual engagement. The former is indicated by the number of students who are 'busy' working on the task at hand. Intellectual engagement can be determined by the questions students asked and the questions that teachers posed. Tasks needed to be designed so that students had to analyze and to reflect as they worked.

Data, Claims and Evidence: Inquiry is the heart of the Lesson Design program. Therefore, student work needs to center around the collection and interpretation of data. Through this analytic process, students need to tie the information they gather to the central scientific ideas of the unit.

Discourse/Discussion: An important instructional principle of the Lesson Design professional development is the use of scientific discourse. The teacher should pose questions that deepen student understanding and connect the lesson generated data to the central ideas of the unit. Students are asked to explain their responses using clear lines of evidence.

Closure/Conclusion: Teachers are often very conscientious about having students put away the materials used in a lesson. Lesson Design training also encourages teachers to help students put the findings and ideas of the lesson in order as well. To do this, teachers review the lesson focus with students as well as the data collected, even if the collection process has not been completed. The day's activities are set within the larger context and next steps are reviewed.

English Language Development: Instructional strategies for English learners are assessed in both the formal introduction of new words and on how well the teachers integrate them throughout the lesson. The goal is to have the vocabulary and related concept development be addressed directly by the teacher and then observe the opportunities for the students to use the terms throughout the lesson.

Five point rubrics were developed for each of the areas, starting with Level 1, where the desired trait is least evident to a Level 5 where it is fully present. Level 1 presents a stage of development rather than a description of an inferior teacher. For example, Level I in Discourse/Discussion describes the condition where the teacher talks and the students listen. The approach to teaching described in Level I is not deemed inferior. It is merely the initial stage for many teachers as they start to incorporate discourse and discussion into their instruction.

Please note that the rubrics that are presented here reflect the target skills for one particular program. The rubrics can provide a model for other programs, but are not intended for direct application. For example, other programs may be focusing on students writing their own research questions, or may be targeting English learner strategies other than vocabulary development. When this is the case, the descriptors in the Data, Claims and Evidence and/or the English Learner sections would look very different.

Implementation

Participants

The 73 teachers participating in Lesson Design during the 2004-5 school year were self-selected from among the K-9 teachers in Imperial Valley. Demographics collected on the participants revealed that they had taught from 1.5 to 32 years with 9.45 representing the average number of years in the profession (Table 2). The teachers were a stable force having spent on average 80% of their career in the same school and nearly five years (4.84) in their current grade assignment. Twenty (20) teachers were very experienced with the science unit and had taught it as many as seven times while 30 others were attending the Lesson Design and content training prior to ever having taught the unit. One third of the teachers attended content training in an academic area that was not the same as the lesson they had videotaped. The classroom types also varied with 27 teaching in a regular classroom, and 46 were in classrooms that had Structured English Immersion (SEI) support. Four of the SEI classes were for newcomers, or students who were recent immigrants to the United States. Since in the Imperial Valley it is rare to have a class without any English learners, 64% of the teachers in this study had classes specially designed to address English Language Development. Class size ranged from seven (7) (private school) to 34 with an average class size of just over 21 students. Teachers reported teaching science, on average just over three times a week.

Table 2
Demographic Chart of Participating Teachers

Number of Teachers	Grade Range	Number Schools	Years Experience	Years at Site	Years at Grade	Experience with Unit	Lessons per wk	Class size	Reg/SEI Classes
73	K-9	27	9.45	6.96	4.84	1.3 times	3.14	21.37	27 / 46
Video Evaluation Group									
20	1-9	10	11.22	8.00	4.17	1.53	2.94	21.00	7 / 13

Data Collection

All teachers participated in the videotaped lessons and lesson evaluation. A random sample of 20 teachers was selected for videotape evaluation.

Data Analysis

The Teacher Behavior Continuum evaluation instrument was designed specifically for this program and matched to the program goals. After the first draft of the instrument was developed, the leadership team met and discussed the following: Did the rubrics describe what we intended? Did the descriptors describe what we wanted to see in the classrooms? Could the behaviors we did see in classrooms be rated according to the rubric categories? The instrument was further validated by through the iterative process of designing the rubrics, reviewing the design with the leadership team, using the rubrics to evaluate a range of video taped lesson, and then having the team suggest further modifications. This process clarified with the professional development team and the program evaluators the target skills of the project.

The reliability of the instruments was established through repeated calibrations. In these sessions, the leadership team and evaluators would watch the same video, rate it on the rubric and then share their scores. Discussions of scoring differences and the reasons that led to particular scores helped align the results. Program evaluators perform a recalibration when they have not used the instrument for several months.

Videos were viewed and individually analyzed by the researchers as per the Teacher Behavior Continuum. Findings were recorded using the rubric. Random sample lessons were chosen to be scored by both researchers in order to establish scoring consistency and reliability.

Teacher Behavior Continuum as an Instrument for Lesson Analysis and Evaluation

Twenty video tapes from randomly selected teachers were to be analyzed by the evaluation team, however only 19 were available. The teachers selected represented all grade levels from first grade through 9th grade. The lessons were analyzed for evidence of: Lesson Focus, Student Engagement, Data, Claims and Evidence, Discourse/Discussion, Closure/Conclusion, and English Language Development using the Teacher Behavior Continuum - Video Analysis Instrument. Results from the lesson analysis are summarized on Table 3.

Table 3:
Teacher Behavior Continuum Results

Teacher	**Grade**	**Lesson**	**Lesson Focus**	**Student Engagement**	**Data, Claims Evidence**	**Discourse Discussion**	**Closure Conclusion**	**ELD**
1	5	Water Cycle	2	2	1	1	1	**1**
2	9	Newton's 2ndLaw	1	3	3	3	1	1
3	1	Finding the Moon	3	2	1	1	2	**1**
4	3	Sound	4	3	3	2	2	5

5	5	Bones Skeletons	3	4	3	2	2	2
6	3	Brine Shrimp	3	3	3	3	3	5
7	3	Brine Shrimp	5	4	3	4	3	4
8	1	Soils	3	4	4	4	4	5
9	2	Sink & Float	1	3	2	3	4	3
10	5	Mixtures Solutions	4	4	4	4	4	5
11	6	Measuring Time	3	4	4	4	4	**3**
12	7	Micro-Macro	3	4	2	4	4	**3**
13	2	Butterflies	4	4	3	3	5	5
14	4	Electricity	4	5	5	5	5	**3**
15	8	Properties Of Matter	5	5	5	5	5	**3**
16	5	Mixtures Solutions	1	3	3	4	4	**4**
17	5	Mixtures Solutions	2	3	3	3	1	3
18	3	Brine Shrimp	5	5	3	5	2	4
19	2	Soils	5	5	5	4	4	4
Average Score			3.21	3.68	3.16	3.37	3.16	3.37

Lesson Focus: Scores in this area ranged from a low of one to a high of five. In three lessons the tasks and intent were unclear to students (level 1) and in four examples the tasks and intent were clearly evident and set within the key concepts of the unit (level 5). Most lessons (74%) contained clear tasks and intent (level 3 and higher). Teachers frequently began the lesson with a focus question and organized the lesson's activities around it. The average score in this area was 3.21 on a 5 point scale.

Student Engagement: This was the strongest area in the lessons with an average score of 3.68. Most lessons received a score of 4 or 5 which indicated that students were engaged both physically and intellectually in the lesson. Physical activity was gauged by the on-task behavior exhibited by students in the video. Intellectual engagement was shown by the level of student questions and discourse with the teacher and among their peers. The lessons shown were all inquiry-based and required active student participation. Students exhibited an enthusiasm for learning in this way by being readily involved in the lesson activities. Some teachers are reluctant to use hands-on methods with their students as they are concerned that students will be off-task and 'playing' during instructional time. This fear was not evident in the videotaped lessons evaluated for this project. Students were focused, on-task, and involved with the lesson activities.

Data, Claims, and Evidence: The Teacher Behavior Continuum emphasized that students collect data and use that data to make claims which they backed up with evidence. At numerous points during the videos that were evaluated, teachers explicitly probed students about the data they were collecting and what claims they could make based on that data. When students offered a claim, the teacher would request the evidence that led the student to the claim. Two teachers did not adhere to the lesson protocol and therefore students did not have the opportunity to collect data and make claims in the portion of the lesson that was videotaped. The other seventeen lessons required students to collect data, 16 of them provided the necessary student support to ensure success. These teachers monitored and guided students to clearly and accurately record data from the lesson. Teachers actively moved throughout the class, talking to students about their work, answering questions, and providing quality control for the information students were gathering and organizing. Five (26%)of the teachers also worked with students to help them correctly interpret the data they gathered and in several instances ensured that students were also able to link their findings to the key concepts of the lesson.

Discussion/Discourse: Again two (11%) of the teachers did not teach through inquiry during the video section and therefore had nothing upon which to base discourse. Two of the remaining 17 teachers engaged students in only procedural and management type discussions. Four (21%) of the teachers asked students fact based questions about what they did and found in the lesson activities. While these were important in checking student data collection procedures and findings, the discussion didn't link to the interpretation and connection to the major concepts of the lesson. Nearly one-third of the teachers posed questions that developed student thinking and helped them connect their findings to the lesson's concepts. This infusion of classroom discourse into the lesson is markedly different than what is seen in a traditionally taught, lecture-based lesson. There, students are seldom engaged in discussion and frequently asked single answer, fact-based questions.

Closure/Conclusion: This was a challenging element to evaluate in the video lessons as many lessons continued over a two day (or longer) period and the complete lesson was not captured on tape. That notwithstanding, closure of the day's activities was still seen as a critical part of the lesson. Just as the students are asked to organize and put away the materials they used during the lesson, they should also be given the opportunity to review and bring at least partial closure to the progress they have made during the day's lesson. This is an excellent opportunity to start sense-making with the students, have them start to see patterns, and prepare them for the next day's activities. With this in mind, each lesson was evaluated for closure. Five (26%) of the teachers either failed to have a closing activity or merely reviewed how to put away the materials. An equal number of the teachers worked with students to arrive at one conclusion which the whole class was to copy into their individual notebooks. While this might be a bit more understandable at the kindergarten and first grade level, especially at the beginning of the year when students may lack the writing skills to develop an original conclusion, the videos showed teacher directed conclusions at the third grade level. Students at this age should be able to formulate original responses. Primary grade students were able to write their own responses with teacher support.

English Language Development: Instructional strategies supportive of English Learners were present in 79% of the classrooms. The lesson vocabulary was presented in oral and written form in most of the regular classrooms, that is, classrooms not specifically designated for English learners. The ELD support strategies were even stronger in the Structured English Immersion classrooms. The average score on the rubric in the area of ELD for these classes was 3.85, a score that indicates that words are introduced in context and used active student participation. Students were also frequently expected to use the new words as the lesson progressed. The ELD strategies introduced in the Villas Institute such as working word walls and kit inventory lessons were evident in many classrooms.

The evaluation results made sense to the leadership team as they actively participated in the development of the instruments and spent many hours in teachers' classrooms so were aware of the range of teacher practices. In general, the evaluation results matched their observations, a condition that was created, at least in part by the earlier collaboration.

Teacher Behavior Continuum as an Instrument for Program Improvement

As can be seen from the descriptions above, the instrument was able to document teacher behavior in ways that led to both a numerical rating and an accurate depiction of classroom instruction. Since the Teacher Behavior Continuum was collaboratively developed, the areas evaluated were previously agreed upon and understood by the team members. The following conclusions were substantiated through the Teacher Behavior Continuum and shared with the leadership team. The program leadership team stated that the findings were extremely helpful in planning program changes and improvement for the ensuing year.

At the request of the project leaders, evaluations were used in the aggregate to give the professional developers feedback and to help guide up coming sessions. While this was the preference for this project, the evaluations could be used to guide teacher self evaluation and these results compared to the outside evaluation scores. This could provide a springboard for a rich discussion of classroom practice. The authors have used a Teacher Behavior Continuum developed for another site for this purpose and found it to be a valuable tool for teacher reflection.

Program Strengths

- Teachers were doing inquiry-based lessons. Almost without exception, students were actively involved in inquiry practices collecting data and using the data to draw conclusions. Teachers were actively involved with the students; moving around the room as students worked; asking questions; giving direction.
- Students were actively involved in the lesson's activities, working cooperatively in groups on the lesson's tasks. Students seemed comfortable and knowledgeable in data collection procedures and were able to record the lesson's findings.

Developing Practices

- Instructional strategies inclusive of English learners were seen in most of the classrooms. Vocabulary was introduced, discussed, and in some instances woven into the lesson. Working word walls with key words were included in some of the lessons and referred to in the written analysis provided by each teacher.
- There was a clear trend toward providing direction and focus for the lessons. Terms such as "Big Idea" and "Focus Question" were used by many teachers in an attempt to set a larger frame and context for the student activities. Focus questions were

frequently printed on the board and students were instructed to put them in their notebooks as part of the scientific process and as a guide to their thinking.

- A shift away from a lecture delivery model was evident. One indication of this shift was that teachers were posing questions to their students. However, questions frequently required only a factual explanation of what students had found. While factual knowledge can provide a critical first step in providing instruction based on student thinking, teachers need to practice posing questions that probe student reasoning. For example, "what happened" questions could be followed by "why" questions, that is, questions that ask students to look for patterns, trends, and relationships within the data that has been gathered.

Target for Improvement

- The area of greatest needed growth among the lessons presented is in lesson closure. Many of the lessons were not completed during the time allotted and teachers often left the day's accomplishments 'hanging.' Many teachers made no effort to set the activities in the context of the question, to look for trends or consistency in the data, to recap the day's work and project the next steps for the following day. There were some instances, however, where closure did take place, even when the entire lesson was not complete. Here, students had the opportunity to examine the results they had compiled so far and start to understand them in the context of the lesson focus.

Program Changes Based on Teacher Behavior Continuum Results

The VIPS team used the above observations and incorporated the findings into an action plan for the upcoming year. They stated:
"Science resource teachers will focus on lesson closure as a significant part of classroom coaching. This along with time management will strengthen the overall delivery of science instruction in our classrooms. Questioning strategies will also be modeled in the content of "how" and" what" types of questions versus questions that may be simply answered "yes" or "no." Videos from the digital video library demonstrating best practices from project teachers will also be available to model lesson closure and questioning strategies." (Klentschy & Molina de la Torre, 2005, p. 10).

This provides evidence of how data collected through the Teacher Behavior Continuum was used and incorporated in the decision making for program improvement for subsequent years. Information collected through the Continuum provided the type of detail that the program leadership team needed to make targeted and focused program changes.

Discussion

The Teacher Behavior Continuum holds potential for helping teams of educators interested in science reform to evaluate practices since they are developed locally and designed to meet specific program goals. The team designed instrument aligns to their specific needs when using the model provided here. The construction of the rubrics must take into account the range of possible teacher behaviors, from having a lack of knowledge about the topic to the highest level of knowledge, that of full integration in daily lessons. One of the powerful facets of using this approach is the interaction among the leadership team during in the development of the instrument. The team must reach consensus on the important program goals and what target teacher behaviors exemplify those goals. When goals are clearly established at the beginning of the program, they provide a roadmap and alignment for professional development and program evaluation. This article provides the instruments that were developed and used to evaluate an inquiry based science program. These tools can be used as models for other programs to design their own program-specific evaluation instruments.

References

Amaral, O. M., Garrison, L., & Klentschy, M. (2002). Helping English learners increase achievement through inquiry-based science instruction. *Bilingual Research Journal*, 26(2), 213-240.

Brophy, J. & Good, T. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Handbook of Research on Teaching*. New York: Macmillan.

Baxter, G.P., Bass, K.M., & Glaser, R. (2000). An analysis of notebook writing in elementary science classrooms. *CSE Technical Report 533*. Los Angeles: University of California, Los Angeles.

Connolly, P. (1989). Writing and the ecology of learning. In P. Connolly & T. Vilardi (Eds.), *Writing to learn mathematics and science* (pp.1-14). New York: Teachers College Press.

Garrison, W. (2004). Profiles of classroom practices in U. S. Public Schools. *School Effectiveness and School Improvement.* 15(3/4), 377-407.

Herrenkohl, L. & Guerra, M. (1998). Participant structures, scientific discourse, and student engagement in fourth grade. *Cognition and Instruction,* 16(4), 431-474.

Klentschy, M. & Molina de la Torre, E. (2004). Students' science notebooks and the inquiry process. In E. W. Saul (Ed.), *Crossing borders in literacy and science instruction.* Newark, DE: International Reading Association.

Klentschy, M. & Molina de la Torre, E. (2005). *California Mathematics and Science Partnership Mid-Course Program Report*. Sacramento, CA: Department of Education.

Morine-Dershimer, G. & Kent, T. (2002). The complex nature and sources of teacher pedagogical content knowledge. In J. Gess-Newsome, & N.G. Lederman, (Eds.). *Examining Pedagogical Content Knowledge: The Construct and Its Implications for Science Education*. Berlin, Germany: Kluwer Academic Publications:

National Research Council. (2000). *Inquiry and the national science standards*. Washington, DC: National Academy Press.

Newton, D. & Newton, L. (2001). Subject content knowledge and teacher talk in the primary science classroom. *European Journal of Teacher Education.* 24(3), 369-380.

Rowell, P. & Ebbers, M. (2004). Constructing explanations of flight: A study of instructional discourse in primary science. *Language and Education.* 18(3), 264-281.

Ruiz-Primo, M.A., Li, M., & Shavelson, R. J. (2002). Looking into students' science notebooks: What do teacher do with them? *CSE Technical Report 562*. Los Angeles: University of California, Los Angeles.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.

Weiss, I. R.; Pasley, J. D.; Smith, P. S.; Banilower, E. R, & Heck, D.J., (2003). Looking Inside the Classroom: A Study of K-12 Mathematics and Science Education in the United States. Retrieved on October 13, 2006. http://www.horizon-research.com/reports/2003/insidetheclassroom/looking.php.

Williams, W. & Sternberg, R. (1993). Seven lessons for helping children make the most of their abilities. *Educational Psychology.* 13(3/4), 317-332.