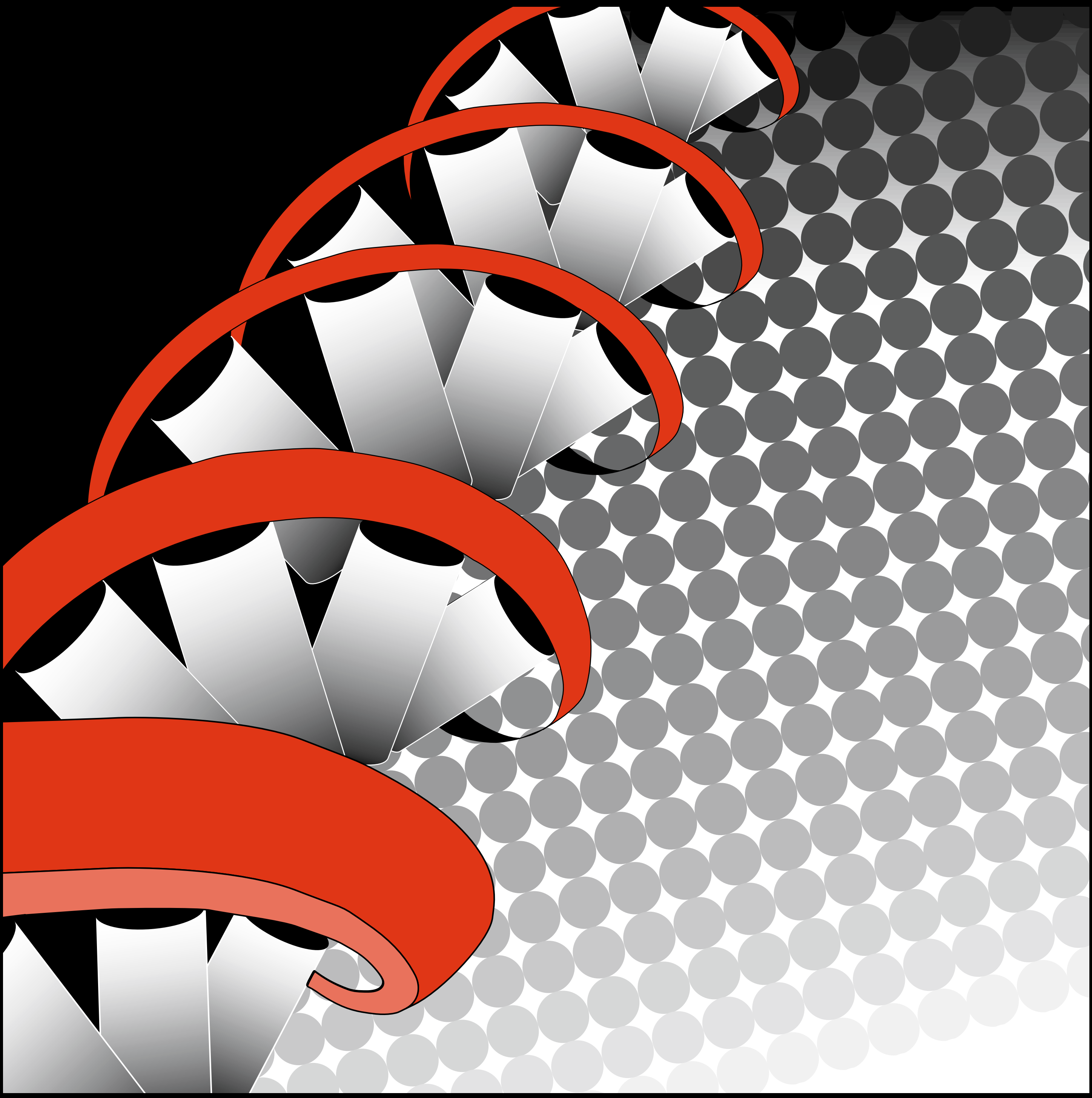
SCIENCE EDUCATOR

THE NATIONAL SCIENCE EDUCATION LEADERSHIP ASSOCIATION JOURNAL



Fall 2007 • Volume 16, Number 2

EDITOR: Jack Rhoton
East Tennessee State University

ASSOCIATE EDITOR: Bill Harwood
University of Northern Iowa

Fall 2007 Vol. 16 No. 2

Officers and Members of the Executive Committee
National Science Education Leadership Association

Table of Contents

Manuscripts: Submit manuscripts in Triplicate to Jack Rhoton, Editor, Science Educator, East Tennessee State University, PO Box 70684, Johnson City, TN 37614-1709 or electronically at rhotonj@etsu.edu. Refer to the information for authors elsewhere in this journal. The opinions and statements published are the responsibility of the authors, and such opinions and statements do not necessarily represent the policies of NSELA. Annual membership dues for NSELA are $45.00. Members may receive the Science Educator for $5.00 per issue, non-members will be assessed a charge of $10.00 per issue. For subscription orders and customer service, contact the editor.

The Science Educator: (ISSN 1094-3277). Copyright 2007 by the National Science Education Leadership Association (NSELA) is published bi-annually with provisions for quarterly publication in the future. Printed at the University Press, East Tennessee State University, Johnson City, TN 37614, the journal serves as a forum for presentation and discussion of issues pertaining to science supervision. It is indexed in the ERIC (Ebsco). General inquiries may be directed to: Susan Sprague, NSELA Executive Director, 2011 Holly Drive, Prescott, AZ 86305 (Ph: 928-771-1030; e-mail: susansprague@yahoo.com) or to Jack Rhoton, Editor.

Philip M. Sadler, Robert H. Tai

Advanced Placement Exam Scores as a Predictor of Performance in Introductory College Biology, Chemistry and Physics Courses

Survey data from 8,594 students in 55 randomly chosen colleges and universities finds that those having passed an AP science exam earn somewhat higher college science grades, but not enough to assume prior mastery. Moreover, half of this performance difference appears to be related to demographics and high school coursework and not to students' AP coursework.

Introduction

Students who take Advanced Placement (AP) examinations in high school seek the option of receiving credit for and bypassing introductory college level courses. These exams are constructed and graded by the College Entrance Examination Board (College Board) to *"allow participating colleges and universities to award high school students college credit, advanced placement, or both" (College Board, 2004, p. 6)*. While the granting of credit or placement remains at the discretion of each individual college or university, the College Board is unequivocal in its interpretation of AP exam scores:

> AP Exam grades of 5 are equivalent to the top A-level work in the corresponding college course. AP Exam grades of 4 are equivalent to a range of work representing mid-level A to mid-level B performance in college. Similarly, AP Exam grades of 3 are equivalent to a range of work representing mid-level B to mid-level C performance in college (College Board, 2006, p. 1).[1]

The purpose of this study is to appraise the College Board's claim of the equivalence of Advanced Placement courses as measured by AP exam scores with the comparable college course in three science disciplines: biology, chemistry, and physics. This is a complex issue because, to some degree, AP course-taking is a sorting process that draws primarily highly motivated, high achieving students. If this self-selection effect is ignored, the outcome of AP course-taking itself may be overestimated. Dougherty, Mellor and Jian (2006a) offered the following insight:

> Much of those [AP] students later success in college may be due not to the AP classes themselves, but to the personal characteristics that led them to participate in the classes in the first place – better academic preparation, stronger motivation, better family advantages, and so on. These selection effects will affect any comparison of AP and non-AP students (p. 3).

Colleges vary considerably in their policies and in their advice to their students concerning enrollment in introductory science courses.

Colleges vary considerably in their policies and in their advice to their students concerning enrollment in introductory science courses. As a result, there exists a sizable population of students taking introductory science courses who have done well in AP science courses while in high school,

1.In this research we refer to the numerical score earned on the Advanced Placement exam (1, 2, 3, 4, or 5) as the "AP exam score." We refer to the teacher's assessment of the AP students' coursework as the "AP course grade." In its literature, the College Board generally refers to the AP exam score as the "AP grade."

yet do not "place out" in college. Since the AP exam is taken prior to college entry, one would expect that these students, whom the College Board has awarded passing exam scores, would earn substantially higher grades, on the average, upon re-taking an introductory college course in the same subject. The apparent advantage is clear. Students exhibiting higher levels of AP performance would most certainly be expected to have a stronger grasp of the relevant college content. They should do better than students who have not had the advantage of what is typically a second year of study in the subject while in high school. The intent of this paper is to characterize the performance of these students as both an independent approach to the validation of Advanced Placement exams in biology, chemistry, and physics and as an estimate of the contribution of these advanced high school courses.

Background

The AP Program of the College Board provides "motivated high school students with the opportunity to take college-level courses in a high school setting."[2] High school AP courses corresponding to the 35 AP examinations administered by the College Board in both science and non-science disciplines are taught by 110,000 teachers in the US. These exams are central to the AP Program

High school AP courses corresponding to the 35 AP examinations administered by the College Board in both science and non-science disciplines are taught by 110,000 teachers in the US.

and offer a means of certifying each student's mastery of course content at the college level. In the sciences, students sat for a total of 311,595 exams in environmental science, biology, chemistry, or physics in 2005.[3] Among these science test-takers, 59% earned scores of 3 or higher (5-point scale). By comparison, 62% of mathematics test-takers earned scores of 3 or better. As a result, many of these students would have grounds to expect their prospective colleges to offer them course credit in these disciplines when they matriculate.[4] This outcome would be consistent with the original purpose of Advanced Placement envisioned by representatives of the faculties of the three private academies (Andover, Exeter, and Lawrenceville) and three private universities (Harvard, Princeton, and Yale) responsible for formulating the program.[5] To promote this policy, the College Board makes available research studies that support its contentions.[6]

With tests scores in hand, students are assured by the College Board that "Over 90 percent of the nation's colleges and universities have an AP policy granting incoming students credit, placement, or both, for qualifying AP Exam grades."[7] However, the fine print is more complex. Having an "AP policy" does not mean that colleges and universities follow the College Board's recommendation that an AP score of 3 should garner course credit or advanced placement. Many times, colleges only offer credit for scores of 4 or even 5. Some college policies do not grant either to incoming students.[8]

While the College Board puts much energy into the construction, monitoring, and scoring of AP exams in an attempt to attain a measure of college course equivalency, there is a considerable range of opinion concerning the value of AP courses in the sciences and mathematics. Many professors believe that all students can benefit from taking introductory science courses in college (National Research Council, 2002, p. 58). Others argue that college-level courses require the efforts of college-level faculty (Juillerat, Dubowsky, Ridenour, McIntosh & Caprio, 1997). Some colleges administer their own placement exams. Others allow credit only for students not majoring in that specific field or award credit only after a student has taken and passed a higher-

2. College Board AP Central website. http://apcentral.collegeboard.com.
3. http://apcentral.collegeboard.com/members/article/1,3046,152-171-0-39036,00.html
4. "ACE [American Council on Education] notes that the practice of granting provisional credit for a grade of 2 is becoming more frequent because comparability studies have shown that many of these students are potentially qualified for college-level work -- credit is made permanent if the student satisfies another criterion, such as successfully completing the next course in the discipline." http://apcentral.collegeboard.com/apc/public/colleges/setting_policy/index.html
5. "[For] the student of marked mathematical and scientific ability and interest, who is probably heading for a college major in science and a career in science or engineering … could take a one-and-a-half or two-year school course in chemistry, physics, or biology, designed to prepare him for admission to the sophomore college course in the field chosen for concentration in school [i.e. in college]. (Phillips Academy, 1952, p. 65.)
6. http://apcentral.collegeboard.com/article/0,3045,152-167-0-11592,00.html
   However, research that can be considered somewhat critical of the claims of the College Board concerning the AP program does not appear to be included at their website (e.g. Geiser & Santelices, 2004; Lichten, 2000).
7. http://apcentral.collegeboard.com/program/ 0,3060,150-0-0-0,00.html
8. Lichten's (2007) sampling of colleges and universities found that only 30% of colleges accept an AP exam score of 3 for advanced placement, while 89% accept a score of 4, and 95% accept a score of 5.

level course in the subject (National Research Council, 2002, p. 58). Some universities offer special transitional courses for those students who place out of introductory courses, but do not advance to the next level course (National Research Council, 2002, p. 59).[9] Lichen (2000) argues that the quality of AP courses has diminished to the point that the College Board's recommendations are no longer accepted by a large number of colleges, particularly those considered highly selective (e.g. Harvard, Yale, MIT, University of Pennsylvania).[10] An NRC committee recommends that colleges should be strongly discouraged "…from using college scores on AP an IB exams as the sole basis for granting automatic advanced placement out of specific courses for majors, or out of biology distribution requirements for non majors (Wood 2002, p. 125.)"

The population of students who have taken AP courses in high school and retake introductory courses has been has been largely neglected by researchers.[11] Surprisingly, researchers have rarely structured past studies to reveal the degree to which AP courses bestow an added benefit upon students who take them, "… the existing empirical evidence regarding the benefits of AP experience is questionable (Klopfenstein and Thomas, 2005, p. 5)." Klopfenstein and Thomas refer to the body of AP research, a majority of which lack or include insufficient controls. We should keep in mind that, in addition to attracting potentially more highly motivated and gifted students, there are other potential covariates: AP courses are more prevalent in wealthier communities, are typically taught by more experienced teachers (Burdman, 2000) and predominantly enroll students who have already studied the subject for a year. Much of the existing AP research gives no recognition to these factors. Studies lacking controls suffer the potential weakness of overestimating the effects of AP courses, because they conflate the influence of numerous factors and fail to account for the relationship of AP courses net of the covariates. While many assume that students who do well in AP courses in high school owe their success to their AP studies, other factors may well be involved, such as coming from more advantaged homes and attending schools that emphasize college preparation (Willingham & Morris 1986). Moreover, the fast pace of AP courses, due in large measure to the focus on test preparation, may not serve students as well as courses that progress at a more moderate pace, focusing on topics in greater depth and detail. This supposition led a recent national study to posit that, "It is possible that AP students were at a disadvantage in some classes or at some colleges" (National Research Council, 2002, p. 194). While most assume that such an outcome is unlikely, we are willing to subject such hypotheses to scrutiny. For example, students reporting the lowest score of 1 (out of 5) on the AP exam have not demonstrated content mastery. However, this outcome leads one to wonder whether this year for these students of study could have been better spent in a course with a different format. Also troubling are the low AP science exam passing rates among under-represented groups.[12]

Prior Research

Research on the value of AP courses in the sciences has primarily focused on accounting for differences in freshman grade point average or persistence to graduation, both rather indirect measures of the equivalence of AP courses and introductory college science courses. Others judge the performance of students who "place out of" one or two semesters of the introductory course in a scientific field and then enroll in a second or third semester science course equivalent to students who take the prerequisite course in college. Such studies have not controlled fully for the fact that AP students generally have stronger high school preparation in math and English, contributing to the ability to "skip ahead." Moreover, second and third semester science courses can cover content unrelated to the prerequisite course (e.g. second semester physics usually entails a study of electricity, magnetism, and light, topics distant from first semester kinematics and Newtonian mechanics). In most cases, studies examine the perfor-

9. None of the courses participating in this study have that profile.
10. USA Today, 3/20/2006, Advanced Placement: A detour for college fast track? Mary Beth Marklein
11. Studies comparing student performance in second-year college courses typically compare those who have "placed out" with high AP exam scores to those with students who take the introductory level college course with either a failing AP exam score or no AP experience (Dodd, Fitzpatrick, De Ayala, & Jennings, 2002). No mention is found in the literature of the experience of those students who chose to retake the introductory college course with a passing AP exam score.
12. Fewer than half of under-represented minorities are "passing" the AP exam with scores of 3 or higher, while scores of 1 (the lowest score possible) are awarded to more than half of African American and Chicano students. Put in another way, students not "passing" the exam constitute 70% of Chicano and African-American students, while more than 50% of Native Americans and Puerto Rican students score below 3. Based on this measure, AP courses do not appear to be an effective means of AP exam preparation for many students. For more information, please consult the AP National Report 2005, http://www.collegeboard.com/student/testing/ap/exgrd_sum/2005.html

The élan of having AP courses, especially in science and mathematics, is well known.

mance of students at a single college or university or within a single state, making generalization to the entire U.S. problematic. These studies also fall into two general categories: studies with adequate control variables and those without. The latter group typically compared grand means between samples of AP and non-AP students or correlated the performance of AP students with other performance-related variables. We found that these studies typically appear as reports that are not published in peer-reviewed journals. Often the source of funding for much of this work is unclear. On the other hand, the "with controls" studies more frequently appear in peer-reviewed journals, having been subjected to more scrutiny and review.

While we reviewed both forms of research, we summarized only the findings of the "with-controls" studies below, grouping the findings under four commonly analyzed outcome variables. The 1986 Willingham and Morris report was especially useful in constructing this list.

- **College Grade Point Average (GPA)**– The total number of AP courses taken in high school has little effect on predicting freshman college GPA, but AP exam scores are a strong predictor of sophomore grades (Geiser & Santelices, 2004). Using matched sets totaling 688 students at Indiana University, AP students were found to earn higher college GPA's (Chamberlain, Pugh, & Shellhammer, 1978). Willingham and Morris' (1986) found that about half of the grade differential (those with a college grade average of B or above) of AP students over non-AP students was accounted for by matching pairs of 1878 students on their academic background. Klopfenstein & Thomas (2005) found that of 28,167 students, those taking AP courses in high school did not earn higher freshman grades in college.
- **College Science Course Grades.** Ruch (1968) did not find a statistically significant difference between matched pairs of AP and non-AP students on college grade in the courses related to the AP course taken or overall freshman GPA. Dodd, Fitzpatrick, and Jennings (2002) studied the performance of three groups of 831 students in a second semester biology course at the University of Texas at Austin. They compare the mean grades of three student groupings:

1. Those who earned credit for the prerequisite course with a score of 3 or above on their AP exam (AP-CR).
2. Those who earned lower than a 3 in their AP course and took the prerequisite (AP-Class).
3. Those who did not take an AP class and enrolled in the perquisite class (Non-AP).

AP-CR and Non-AP groups were matched using high school class rank and SAT and ACT scores. Over a four-year period, the Non-AP group earned higher mean grades than either the AP-Class or the AP-CR groups in the second semester course. This difference was reported to be statistically significant for one year of the study. The AP-Class group earned slighted higher grades in the second semester course than the AP-CR group. Reported GPA's in other college biology courses were also higher for Non-AP students.

- **Persistence to college graduation** – The combined number of AP and honors courses taken in high school was not a significant predictor for persistence through freshman to sophomore years in college (Geiser & Santelices, 2004); Klopfenstein & Thomas (2005) found students taking AP courses in high school have no higher probability of persisting to a second year of college. Low-income students who take AP courses in high school graduate college at a greater rate (Dougherty, Mellor & Shuling, 2006a)
- **Choice of Further Study** – Ruch (1968) used 21 matched-pairs of students and found that that AP students were more likely to continue with the same subject in college. Chamberlain, Pugh, & Shellhammer (1978) found academic progress, college GPA, and course-taking measures were superior for AP students over 344 matched sets of students (Cahow, Christensen, Gregg, Nathans, Strobel, & Williams, 1979). Willingham and Morris' (1986) found that of 1878 students, AP students were no more likely to take more than one year's coursework in biological or physical science, nor to major in those fields.

All of the studies above controlled for students' high school performance (i.e. rank, HSGPA, or coursework completed) except for Chamberlain *et al*. (1978) and Dougherty *et al*. (2006a). Geiser & Santelices (2004), Klopfenstein & Thomas (2005), and Dougherty *et al*. (2006a) included measures of students' socio-economic status. Ruch (1968), Dougherty *et al*. (2006a) and Klopfenstein & Thomas (2005) included high school quality ratings, while Willingham & Morris (1986) matched students by college attended.

From this short list of controlled studies, it is apparent that the measure of AP coursework value varies with the outcome variable considered and the control variables included. We did not find any studies that measured performance in introductory college science courses while controlling for relevant covariates such as performance in high school courses considered prerequisites for AP science courses. Given that performance in prior coursework and is widely regarded as the strongest predictor of subsequent performance, it is essential that this factor should be included in any analysis seeking to isolate the relationship to AP coursework itself, rather than merely considering an aggregated analysis that lacks both detail and is unable to estimate the relative contributions of various factors. The inclusion of such controls provides an opportunity for refinement of models that more accurately estimate the relationship of AP coursework relative to other factors. Regression models must be constructed using variables representing viable alternative hypotheses to the main relationship being studied in order to be considered robust. For example, inclusion of SAT scores alone, entered as a control for high school-level learning, is insufficient to account for prior science knowledge learned in first-year high school science courses. It is incumbent on the scholar employing regression to explore alternative relationships that are potentially causal and to take great care in considering, measuring, and including factors that could account for additional variance. Put another way, it is the attention that scholars pay to controlling for alternative hypotheses that is often a measure of quality research.

But, why is research on the direct comparison of AP and non-AP students important? Conventional wisdom ascribes clear and overwhelming benefits to Advanced Placement classes in high schools. The élan of having AP courses, especially in science and mathematics, is well known. As a result, the College Board's AP program has grown dramatically over the last 20 years, with a sustained increase in the number of exams including all subject areas of 9.3% (with an 8.8% increase in number of students) annually. To offer some perspective, this growth is more than five times faster than the growth in the US high school population.[13] The ratio of students to exams tracked closely until recently; their separation is indicative of an increase in the number of students taking multiple AP exams (1.29 in 1970 vs. 1.72 in 2005). In 2005, 1.2 million students took 2.1 million AP exams. According to the latest available data, 2.1 million entering freshmen enrolled in degree granting institutions in Fall 2003. One might assume that on the average nearly half of all entering college freshmen have taken at least one AP course in high school.[14] In reality, many have taken more than one course. Our estimate is that in 2004, those taking AP science exams represented one of five college freshmen (Digest of Educational Statistics, 2004).[15]

The expansion of AP Programs is also currently propelled by economic and political support since several states currently extend financial incentives to high schools that offer AP courses[16], while others require all high schools to participate (Hershey, 1990; Willingham & Morris, 1986).[17] Many states have mandated through legislation the inclusion of AP enrollment in admissions decisions or the granting of college credit for high AP exam scores in public institutions of higher education (Lichten, 2000). The U.S. president has called for an addition of 70,000 Advanced Placement teachers in science and mathematics, a near doubling of current numbers.[18] In addition, the U.S. Department of Education subsidizes AP programs for low-income students (Klopfenstein, 2004).[19] This expenditure of public funds, in effect, subsidizes one particular educational program over others. Hence, it is desirable that Advanced Placement policies and programs be rigorously assessed so that scarce resources can be spent the most wisely.

A powerful motivating factor for students is the fact that one may enter college with course credits and in some

13. http://apcentral.collegeboard.com/members/article/1,3046,152-171-0-47040,00.html.
14. http://nces.ed.gov/programs/digest/d05/tables/dt05_180.asp.
15. http://apcentral.collegeboard.com/members/article/1,3046,152-171-0-39036,00.html.
16. Florida, Louisiana, and Utah.
17. South Carolina.
18. George W. Bush, State of the Union Address, 1/31/2006.
19. Higher Education Act Amendments of 1998, Title VIII, Part B, P. L. 105-244.

cases with enough credits to bypass the freshman year, saving a year's tuition (MacVicar, 1988; Pushkin, 1995). Troubling to many professors is that students may use AP credit in many colleges to satisfy their requirements for science courses and never take another science course when in college (National Research Council, 2002, p. 59). For example, Lichten (2000) found that only 22% of students who had scored "3" on their AP calculus exam enrolled in an advanced calculus course in college, while 17% took a remedial calculus course. Though high school AP teachers may be very talented, they generally do not have the breadth of knowledge nor the most recent and relevant content knowledge possessed by college faculty (Lichten 2000). By accepting AP exam-related course credits, students may be forego-ing the opportunity to learn about the most topical findings and issues from college instructors whose "other" job it is to do research.

The striking expansion of the AP program offers an opportunity to gauge the pedagogical influence of AP coursework. Merely presuming that AP courses are comparable with college courses ignores the opportunity we have to analyze the degree of this purported similarity as a means to improve college preparation through high school instruction.

Methods

The Sample

Factors Influencing College Science Success (Project FICSS) is a large-scale national study of introductory college students and their high school science experiences, which collected a total of nearly 18,000 surveys. The data includes a series of subsamples of college students in various college science courses. In this particular analysis, we analyzed survey data gathered from students in 124 different first semester introductory college biology, chemistry, and physics courses in 55 colleges and universities. These institutions were from a stratified random sample reflecting a nationally representative distribution of institutional size and selectivity, from small liberal arts colleges to large state universities. All participating introductory science courses met in lecture-style classes with weekly recitations sessions and laboratories and had enrollments of 10 students or more.

Of 9354 collected surveys in this subsample, 1029 respondents reported having taken an AP course in the subject in high school in which they were currently enrolled in college. Of the students who took AP courses, 316 earned scores of 3 or higher on the AP exam, roughly 3% of the entire sample. While this can be considered a small number compared to the 170,000 students who earned scores of 3 or higher on AP exams in these subjects in 2005, the data used in this analysis represents a national sample of students who have both taken part in Advanced Placement courses and exams in high school and also take introductory level college science courses. The theory and practice of inferential statistics has established that valid inferences (within appropriate confidence limits) can be drawn from relatively small number of participants, provided they are from a well-selected random sample.

... many students feel that while AP is a good preparation for college science, many also feel they benefited from taking the actual college science course.

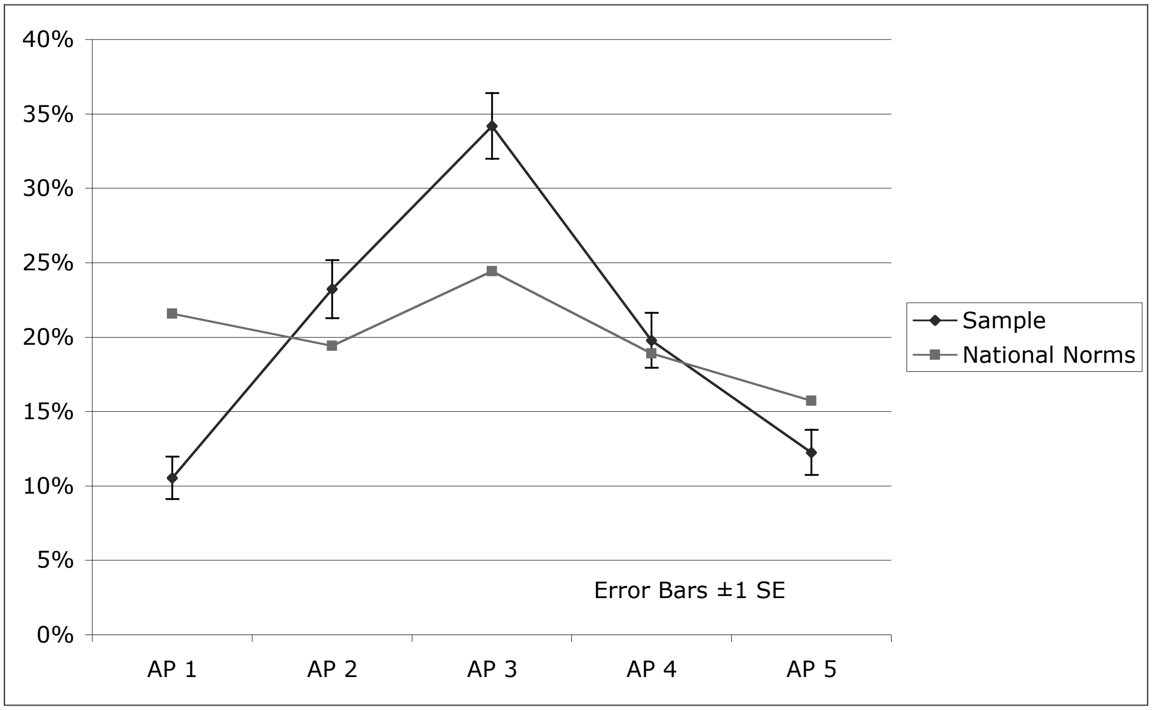
Choosing to focus only on undergraduates, we omitted cases of 103 graduates and 400 non-traditional students. We also omitted 397 students who spent their high school years outside of the United States. Allowing for some students in the intersection of these various groups, our final sample included 8594 undergraduates.

Of this total, 31% of the sample was enrolled in biology, 42% in chemistry, and 27% in physics. Females made up 60% of the biology students, 55% of the chemistry, and 42% of the physics students. Parental education was related to the highest level of mathematics taken by students, with representation of students taking calculus in high school ranging from 33% of student with parents having a high school diploma or less to 49% of students with a parent holding a graduate degree. The sample identified themselves as 77% White, 7% Asian/Pacific Islander, 6% black, 5% Hispanic, 3% multi-racial, and 2% other or unreported.

This study, by its very nature, may only examine students who take introductory-level college science courses. Students enrolled in colleges that offer the option of bypassing introductory-level science courses with sufficient AP exam scores and who also make the decision to take advantage of this option to place out are not included. Yet, since prior research studies have generally ignored this interesting population, studying these students can make a valuable contribution to understanding the value of Advanced Placement programs.

Students who did poorly on AP exams in a given science discipline may choose not to take the college course. By comparing the national statistics for AP exam scores with the

Figure 1: Distribution of AP Exam Scores in Sample vs. National Norms



distribution in our sample, one can see differences (Figure 1).[20] The five AP score categories each average about one-fifth of the students who take the exam.[21] It appears that students at the low end (AP exam scores of 1 and 2) and at the high end (AP exam scores of 4 and 5) are under-represented in the sample. Presumably a larger fraction of those who scored below 3 chose not to enroll in a college course. Many at the higher end arguably took advantage of placement to enter a higher-level course or received credit for the introductory course. The mean AP exam score for the national data is slightly lower (2.81) than the sample used in this analysis (2.99) reflecting a relative dearth of students earning low AP exam scores.

We were also careful to check for any anomalies in the makeup of students who earned scores of 3, 4, or 5 on their AP exams. Table 1 displays a comparison across seven measures of academic performance among nine high school science preparation categories ranging from Computed grades represent the mean grade for a particular group (on a 100-point scale where A=95, B=85, etc.). "Not Taken" for students reporting that they had not taken a high school course in the discipline of their introductory college science course to "AP 5" for students reporting they had taken the Advanced Placement exam in the discipline of their introductory college science course and received a score of 5.[22] In general, AP students outperformed honors students, who in turn outperformed regular students, who in turn outperformed students with no high school coursework in the corresponding college science discipline. Within the AP group, we see a clear pattern of higher performance corresponding to higher AP exam scores. High scoring AP 4 and AP 5 students consistently reported higher SAT scores, mathematics grades, in-discipline science grades for regular or honors courses, and grades in other science courses. Higher scoring students in our sample

20. The national statistics for AP biology, chemistry, and physics have been averaged.

21. One should note that the distribution of scores for the AP exam is not Gaussian as is reported for most national standardized exams such as the SAT or ACT. While those at the center of the distribution can be thought of as earning close to the mean score, those at the 4 level or above (35%) or at the 5 level (16%) occur with much greater frequency than students at one standard deviation above (16%) or two standard deviations above (2%).

22. Regular signifies that the highest level of course taken in high school was neither AP nor honors level. Honors signifies an honors course. AP courses are broken down by AP exam scores of 1 through 5 or the category of students not taking an AP examination. Means are reported for the SAT math or SAT verbal scores (or the equivalent from the ACT), the fraction completing a calculus course in high school, the last grade in mathematics, the mean grade in the relevant science preparatory course (if taken), the mean grade in the AP course (if taken), and the average of science grades in fields other than the college science course of interest.

also reported higher grades in their AP courses, on average. Students in our sample who chose not to take the AP exam, even after taking the AP course, reported academic performance measures similar to the students in the AP 3 category.

In our analysis, we were particularly concerned with the academic performance in college of students who reported receiving AP Exam scores of 3, 4, or 5. We wondered if the students in our sample were particularly weak in these science disciplines despite their high AP exam scores. Might a weakness in their academic background have led them to pursue additional course work at the introductory college level? We imagined that students with high AP exams, but relatively weak academic backgrounds may not be fully fluent in English, or may have weaker math backgrounds, or may have earned lower overall high school grades. To address these possibilities, we analyzed these various background measures and did not find that, compared with other students, high AP exam scorers exhibited any weaknesses in language proficiency (SAT Verbal score), mathematics proficiency (SAT Math score, high calculus course taking patterns, high school mathematics course grades), or overall high school grades (grades in various high school science courses including AP, regular, or honors). On the contrary, we found in our sample that students who reported an AP exam score of 5 also averaged 1300 on their SAT's, with three-quarters having taken calculus in high school, and also reported earning an "A" on average in their high school science and mathematics courses.

Who are the group of students who have taken an AP course in high school and have earned an AP exam score greater than 3? Why are they taking an introductory course in college? Were they advised or required to take this course by their college? To gain a better understanding of possible motivations, we contacted 44 students who both scored well on AP exams and took introductory college science courses in these disciplines. These students had volunteered their contact information in response to a question asking if they would agree to future contacts in light of survey related queries. Among the responses, fifteen students provided detailed replies. While these responses should not be considered representative, they do offer some insight into the diversity of students' experiences and opinions. Five made the personal decision to take the introductory course even though they could have placed out, citing advice from more advanced students in the field or their academic advisors. Five took the course because of department requirements in their disciplines. Three reported AP exam scores of 4, while their college required a score of 5 for credit. Two reported low scores on departmental placement exams required by their colleges to earn course credit. All but one felt that their AP course prepared them well for introductory college science. However, eleven reported that taking the introductory college course as a very worthwhile experience that paid off for them later in their college science course work. Four noted, in retrospect, after taking advanced courses, that they should have moved directly to higher-level course work.

As one of those students noted:

- *I very much should have been [granted] credit. The class was pretty much a waste of my time.*

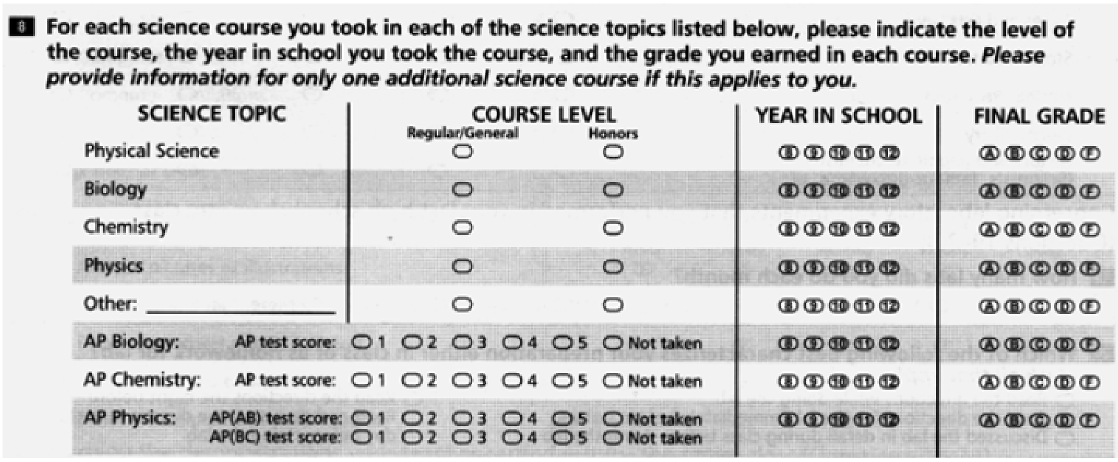
In contrast, others offered more positive reports of their college science course experiences:

- *It was worthwhile and I am glad I chose to take the intro level course. I recommend it to freshman coming in who face a similar situation. Taking the introductory course refreshes my memory on the subject, provides deeper insight into the subject that just*

Table 1: Comparison of High School Performance Measures by Group

Highest HS Course or AP Score	Sat Math	SAT Verbal	HS Calculus	Last Math Grade	Other Science Grades	Grade in Honors or Regular	AP Grade
not taken	475	454	27%	86.0	88.9		
regular	529	510	32%	87.7	89.4	88.8	
honors	596	564	60%	89.4	90.8	90.4	
AP, no exam	610	574	70%	89.6	91.1	92.4	87.6
AP 1	543	508	47%	85.2	88.4	90.0	85.0
AP 2	592	541	54%	90.1	90.5	91.5	88.1
AP 3	595	568	61%	90.1	90.9	91.8	88.9
AP 4	631	628	64%	91.0	92.1	93.9	92.2
AP 5	666	654	75%	92.5	93.0	94.0	94.1

Figure 2: Survey Questions Relating to High School Course Taking in Science and Mathematics



8 **For each science course you took in each of the science topics listed below, please indicate the level of the course, the year in school you took the course, and the grade you earned in each course.** ***Please provide information for only one additional science course if this applies to you.***

SCIENCE TOPIC	COURSE LEVEL		YEAR IN SCHOOL	FINAL GRADE
	Regular/General	Honors		
Physical Science	○	○	8 9 10 11 12	A B C D F
Biology	○	○	8 9 10 11 12	A B C D F
Chemistry	○	○	8 9 10 11 12	A B C D F
Physics	○	○	8 9 10 11 12	A B C D F
Other: ________	○	○	8 9 10 11 12	A B C D F
AP Biology:	AP test score: ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ Not taken		8 9 10 11 12	A B C D F
AP Chemistry:	AP test score: ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ Not taken		8 9 10 11 12	A B C D F
AP Physics:	AP(AB) test score: ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ Not taken AP(BC) test score: ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ Not taken		8 9 10 11 12	A B C D F

wasn't there in high school, and prepares me for higher level courses in the subject.

- *It was definitely a worthwhile experience to retake the classes, as it gives you a second chance to absorb the material—which can never hurt.*
- *I was at least familiar with many of the topics covered in my college class. Many concepts were covered with greater detail in college. I felt well prepared for the college class, but I am glad I took the class instead of testing out.*

The student comments indicate that both motivations and experiences associated with introductory college science courses differ widely.

The Survey

The survey instrument was constructed to collect data concerning academic experiences and performance measures closely associated with high school science course work that might predict performance in introductory college science classes. In the form of a 4-sheet booklet, students were asked to complete these surveys while in their college classes. Using the names and ID numbers provided by students on the cover, college instructors entered students' course grades. After this process was completed, the perforated cover with student identification was torn away and discarded, leaving the surveys anonymous, except for any contact information students may have volunteered in response to a "further contact" request. Students also reported on a variety of demographic variables (parents' education, profession, etc). Home ZIP codes allowed us to match the surveys with existing data on median income, home value, and average educational background of each home locale. The survey questions themselves were carefully designed to fit on the front and back of the remaining three sheets. See Figure 2.

Two major concerns loom for any large-scale self-report study, accuracy and reliability. These concerns stem from a report by Bradburn, Rips, and Shevell (1987) that appeared in the journal *Science*. The study concluded that self-report data were inaccurate and unreliable. However, in research that has followed in the intervening years, which includes further research by Bradburn himself, these conclusions have been qualified and mollified to portray a complex and more responsible conclusion: that the accuracy and reliability of self-report likely depend on several factors including context, relevance, and survey clarity Pace, Barahona & Kaplan, 1985; Bradburn, 2000; Niemi & Smith, 2003). In particular, self-reports of course taking, grades earned, and standardized test scores made by college students tend to be highly accurate (Baird, 1976; Anaya, 1999). Enrollment reports are especially accurate for courses with unambiguous names and for high-achieving students (Sawyer, Laing & Houston, 1989). In a recent review of existing research on self-report, Kuncel, Credé, & Thomas (2005) conclude that self-report may be characterized as reasonably accurate in samples where the surveys address issues relevant to the respondents. Our study surveyed introductory college science students in their fall semester college science classes where reflection on their prior experience would be commonplace. In addition, the students' own instructors administered the surveys.

We entered this project with previous survey research experience in which we surveyed over 2000 college physics students in over 20 different college and universities (Sadler & Tai, 2001). The design of the current survey included the use of student focus groups to comment on question formats, two pilot surveys that included 304 college science students, a review by an advisory panel of college professors and high school teachers, and a separate reliability study where 113 college chemistry students took the survey twice, two weeks apart and where their responses were compared. Much can be done to enhance accuracy and recall: careful design, contextual cues, and participant relevance all play important roles.

In large-scale surveys, missing responses are not uncommon. In dealing with missing responses, list-wise deletion is the simplest and most commonly applied option. However, list-wise deletion assumes that the data are missing at random. Should a large percentage of the sample exhibit missing responses, this approach may not be the most appropriate option. We analyzed the percentage of missing responses and found that, for the variables of concern, only about 1-2% were missing for each. Given this characteristic, we chose to use list-wise deletion. The two exceptions were that no standardize test score (SAT or ACT) was entered for 347 cases (quantitative) and 386 cases (qualitative). A mean value for such scores was substituted in these cases since we did not want to eliminate students from the dataset who had not taken these exams.

Choice of controls

Our initial analysis compares the mean college grades for students with different levels of high school coursework in biology, chemistry and physics. This particular analysis only shows the predictive value of AP exam scores on college course grades, quantifying the magnitude of the difference between students who take AP or not, and those who score at different levels on the AP exam. This first step describes the differences that need to be explained by competing hypotheses (Dougherty, Mellor & Jian 2006a).

However, attempting to account for the potential of taking an AP course requires controls for demographic and academic preparation differences among students as well. Far from being *tabula rasa*, students' skill and knowledge prior to, or unrelated to, taking an AP course should be included. Table 2 shows the Pearson correlations between academic preparation variables and college course grade.

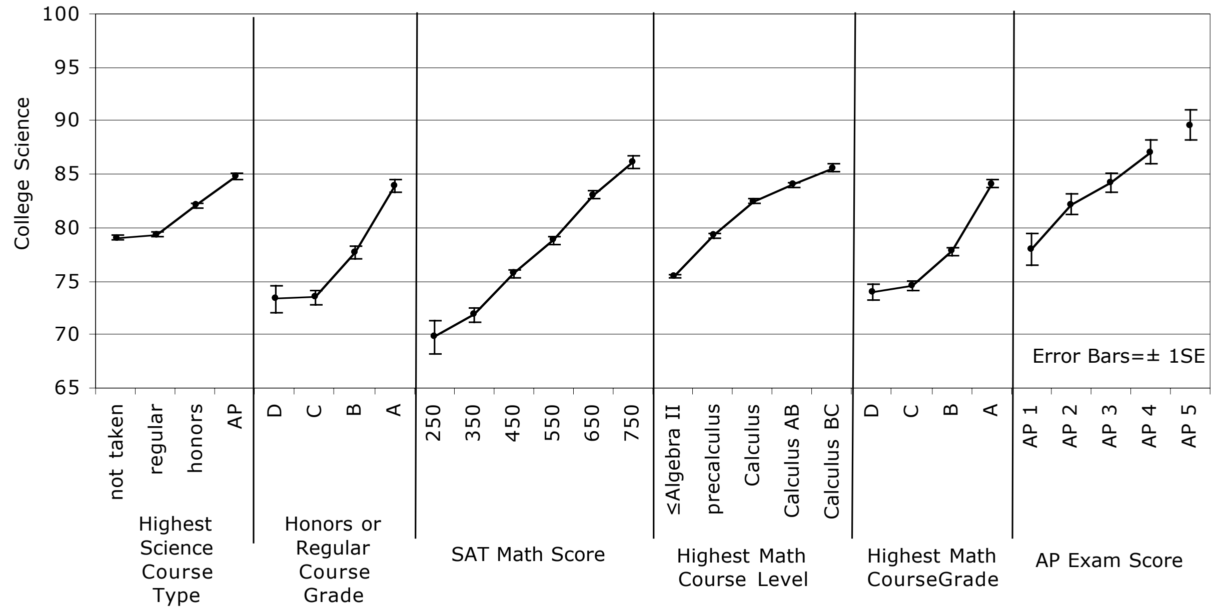
Among the nine academic performance measures, AP Score exhibits the highest correlation with College Grade in introductory science courses, though other measures with similar correlations include: SAT Math, "other" science GPA, and last high school math grade. A central issue faced by those investigating the association between AP exam performance and college performance is that many other measures of academic performance are also positively associated with college grades.

Figure 3 offers a graphical representation using the results produced from our sample. College science grades vary with AP exam score, but they also vary with the other five measures of academic preparation. As expected, more rigorous levels of high school science and mathematics courses match with higher college science grades. Each graph shows a clear monotonically increasing association between college grades and the six high school measures of academic achievement. It should be noted that the average grade attained by students who earn an AP score of 3 is equaled by students: earning a grade of "A" in

Table 2: Correlation Coefficients of Academic Preparation Covariates

	College Grade	SAT Math	SAT Verbal	Last Math Grade	Last Eng Grade	Science GPA	Regular or Honors Grade	Honors	Calculus	AP Score
College Grade	1.000									
SAT Math	0.315	1.000								
SAT Verbal	0.200	0.330	1.000							
Last Math Grade	0.317	0.288	0.108	1.000						
Last Eng Grade	0.219	0.134	0.200	0.343	1.000					
Science GPA	0.306	0.254	0.190	0.387	0.378	1.000				
Regular or Honors Grade	0.269	0.238	0.166	0.306	0.265	0.464	1.000			
Honors	0.128	0.199	0.140	0.107	0.098	0.129	0.063	1.000		
Calculus	0.244	0.415	0.203	0.192	0.159	0.247	0.196	0.281	1.000	
AP Score	0.328	0.311	0.404	0.216	0.189	0.277	0.207	0.004	0.157	1.000

Figure 3: Course Grade and Values of Key Covariates.



an honors or regular science course, scoring a 750 on their quantitative SAT, who take an AP course in calculus, or whose last high school mathematics course grade was an "A."

Results

Descriptive statistics

Table 3 displays a cross tabulation between college science grades and the various academic performance levels displayed in Table 1. Approximately 13% of students in introductory courses did not take a corresponding course in high school. Roughly half of the students in our sample reported taking only the regular level high school course in the discipline corresponding to college science course, with a quarter enrolling at the honors level in high school and 11% enrolling in AP courses of the corresponding science. Of the students who reported enrolling in AP courses, 433 took the AP exam and reported their scores.

Inferential Analysis

The inferential analysis offers a series of five multiple linear regression models. Model A includes only the AP status of student: if they took AP, if they did not take the exam, or exam scores if they did take the exam. It explains only 2% of the variance in college grade. This number is to be expected since the AP status is only one factor in the background of students. Model B includes a dummy variable for each professor (each separate college course) and accounts for differences in grading severity between courses. This shows the predicted grades as if all the students enrolled in the same university course. This model shows a small upward change in the predicted mean grade earned by each AP grouping. Model C adds in students SAT scores, or ACT scores converted using a concordance table (Dorans, Pommerich & Lyu, 1999) for both mathematics and composite scores. Model D includes variables that account for general high school performance in science, mathematics, and English, typical of many studies accounting for high school grade point average (HSGPA) or its correlate, high school rank-in-class. Model E adds the kind of course and performance in any preparatory course prior to taking the AP course in the subject. This model seeks to control for the influence of different levels of preparation in biology, chemistry or physics. Model E can be interpreted as presenting the difference in college performance that is related to taking an AP course or to a certain performance on the AP test between hypothetically identical students in all other variables, i.e. between students who had the same preparation and background in prior years. It generates an estimate of the incremental effect of taking an AP course or having a certain performance on the AP test on the college grade in the same subject. Included for each model are the standardized beta (β) coefficients that allow a comparison between the magnitude of the relative contribution of each variable to the

Table 3: Distribution of College Grades by Most Rigorous High School Course in the Subject

College Science Grade	Prerequisite HS Course			Advanced Placement Course							total
	not taken	regular	honors	no exam	took exam	AP1	AP2	AP3	AP4	AP5	
55 F	106	356	117	12	12	4	3	4	0	1	603
61 D-	4	36	11	0	3	1	1	1	0	0	54
65 D	102	381	127	22	11	3	4	2	2	0	643
68 D+	12	83	21	0	5	0	3	1	0	1	121
71 C-	48	163	57	5	15	2	5	5	3	0	288
75 C	226	957	382	78	56	9	17	23	6	1	1699
78 C+	46	245	61	12	22	3	8	5	5	1	386
81 B-	56	225	83	22	28	7	8	5	7	1	414
85 B	243	934	518	134	79	6	17	38	12	6	1908
88 B+	44	226	102	27	56	5	16	14	10	11	455
91 A-	36	202	90	22	39	1	6	15	10	7	389
95 A	179	659	447	165	101	4	17	31	27	22	1551
98 A+	1	31	40	5	6	0	1	1	2	2	83
total	1103	4498	2056	504	433	45	106	145	84	53	8594
Mean	78.79	79.18	81.98	85.21	84.35	78.11	82.11	84.30	87.15	89.77	80.42
SE	0.36	0.17	0.25	0.43	0.47	1.63	0.96	0.78	0.88	1.05	0.12

explained variance. Note that the final regression model accounts for 35% of the overall variance. The lack of high correlation coefficients in Table 2 as well as other tests are evidence of a lack of colinearity in the models.

A graphical comparison of Model A and Model E is shown in Figure 4. (Error bars represent one standard error from the mean for each category.) The two models show the comparison of mean college science grades by simple calculation from raw grades in Model A and by controlling for the collection of important covariates in Model E. The course mean is represented by a horizontal dashed line. The magnitudes of coefficients for AP experience are clearly muted when important background factors not associated with AP are accounted for in subsequent models. The average slope of the data relating AP score with course grade is considerably reduced as shown in the third panel of Figure 4. Accounting for covariates leads to a reduction in the magnitude of all AP-related coefficients.

Model E, deemed the "full model," is particularly interesting, given that it accounts for several factors not easily disaggregated from students' and teachers' experiences. This model includes four significant demographic variables that together may be thought of as measures of socio-economic status (SES). Adjustments for such variables can never insure that SES is controlled for fully, since it cannot be measured without error:[23]

- Race – Accounting for the self identified race or ethnicity of students: White: Black, Hispanic, Asian, multi-ethnic, or other.
- Hipared – the highest level of education of either parent: less or equal to a high school diploma, some college, 4-year degree, graduate or professional degree.
- Hstype' – The kind of high school last attended: private, public or other.
- Mean Ed Level – The average educational level of those in the student's home ZIP code, based on the categories use for Hipared.

Discussion

The initial model which does not control for student backgrounds and academic performance prior to, and independent of, any AP course experiences, shows large differences between students who enroll in AP and others, earning college course grades about four points higher than the course average. There is little difference between those reporting taking the AP exam and those who choose not to in AP courses. However, when demographics and prior academic achievement are accounted for, the apparent advantage held by students with AP experiences in high schools are roughly cut in half. This reduction is similar to findings by Dougherty *et al.* (2006a) in which the introduction of controls in their regression analysis

23. This point suggested by an anonymous reviewer of this paper.

Table 4: Regression Models Explaining Variance in College Science Grade

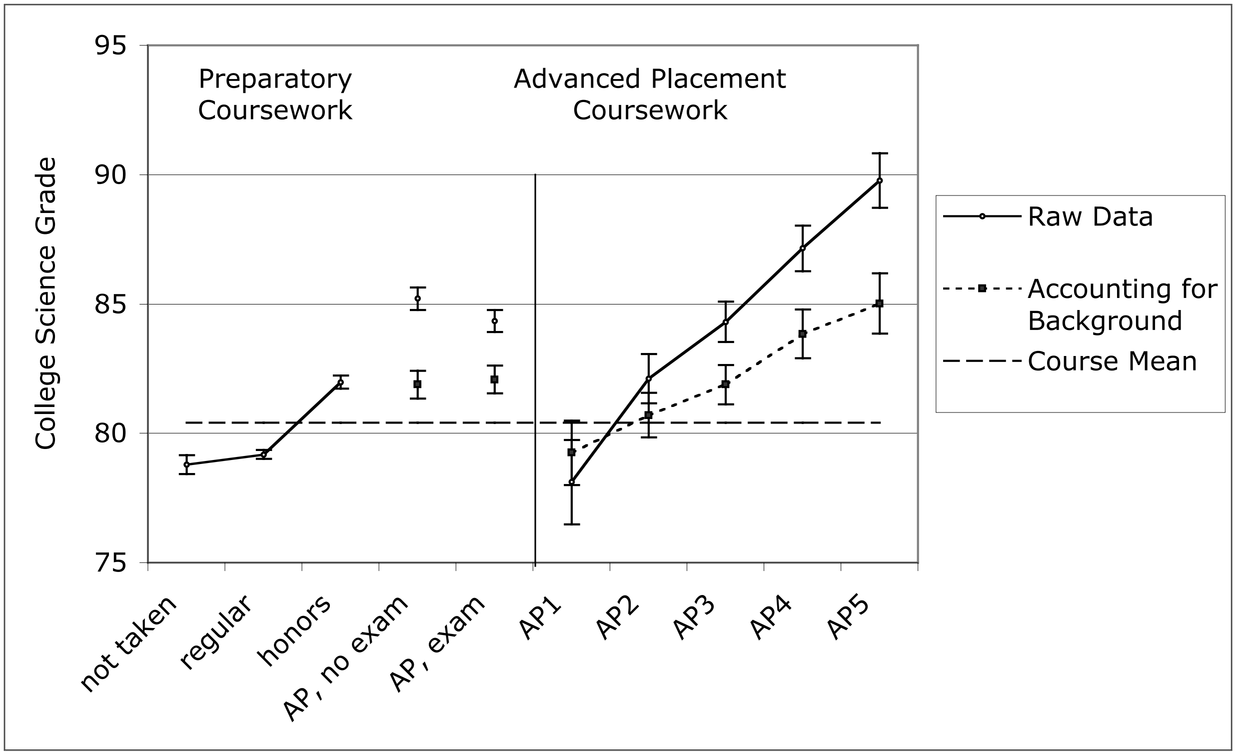
	DoF	Model A		SE	Model B		SE	Model C		SE	Beta	Model D		SE	Beta	Model E		SE	Beta
Constant	1	83.79	***		84.76	***	0.45	59.07	***	1.26		37.73	***	1.61		42.84	***	1.68	
Prof ID	123					***			***				***				***		
Race	5								***				***				***		
hipared	3								***				***				***		
Hstype'	2								**				***				***		
Mean Ed Level'	1								-				***			0.61	***	0.19	0.03
SATM_CAL'	4							0.01	***	0.00	0.09	0.01	***	0.00	0.09	0.02	***	0.00	0.13
SATV_CAL'	1							0.03	***	0.00	0.27	0.01	***	0.00	0.06	0.01	***	0.00	0.06
himath	4												***		0.13		***		0.11
≤Algebra II												-2.69		0.28		-2.31		0.28	
precalculus												-1.33		0.19		-1.16		0.19	
Calculus												0.36		0.25		0.27		0.25	
Calculus AB												1.38		0.23		1.17		0.23	
Calculus BC												2.28		0.37		2.03		0.37	
Math Grade	1											2.51	***	0.15	0.17	2.27	***	0.15	0.16
English Grade	4											0.93	***	0.19	0.05	0.75	***	0.19	0.04
Science Grades	1											2.99	***	0.23	0.14	2.13	***	0.24	0.10
Prep Course	4																***		0.08
not taken																0.19		0.29	
not reported																-0.41		0.39	
C or lower																-1.63		0.33	
B																-0.28		0.22	
A																2.12		0.21	
Honors	1															0.97	***	0.26	0.04
AP Status	6		***			***			***		0.19		***		0.14		***		0.15
no AP		-3.91		0.45	-4.66		0.44	-3.57		0.41		-2.69		0.39		-2.26		0.40	
No AP Exam		1.42		0.61	0.14		0.60	0.07		0.56		0.11		0.53		0.16		0.53	
AP 1		-5.68		1.49	-5.22		1.41	-3.32		1.33		-2.19		1.26		-2.47		1.25	
AP 2		-1.68		1.02	-1.34		0.97	-0.80		0.91		-0.94		0.86		-1.02		0.86	
AP 3		0.51		0.90	0.78		0.86	0.36		0.81		0.18		0.76		0.16		0.76	
AP 4		3.36		1.13	4.22		1.07	2.93		1.01		2.22		0.95		2.12		0.94	
AP 5		5.98		1.38	6.08		1.32	4.32		1.25		3.31		1.18		3.30		1.18	
Level of AS		**Raw Data**														**Full Model**			
no AP		79.88		7657	80.11		7657	79.05		7657		80.18		7575		79.46		7575	
No AP Exam		85.21		504	84.90		504	82.69		504		82.98		502		81.88		502	
AP 1		78.11		45	79.55		45	79.30		45		80.68		45		79.25		45	
AP 2		82.11		106	83.42		106	81.82		106		81.93		105		80.70		105	
AP 3		84.30		145	85.54		145	82.98		145		83.05		145		81.88		145	
AP 4		87.15		84	88.98		84	85.55		84		85.09		84		83.84		84	
AP 5		89.77		53	90.84		53	86.94		53		86.19		52		85.02		52	
Total Cases		8594			8594			8594				8594				8594			
Cases Missin		0			0			0				86				86			
Cases Used		8594			8594			8594				8508				8508			
r^2=		0.022			0.153			0.252				0.336				0.347			

resulted in reductions in the modeled outcome variable, college graduation rate, from 39% to 26% for students passing an AP exam. Hence, in that study, a third of the apparent effect of passing an AP exam was explained by the two covariates used: eighth-grade mathematics achievement and high school characteristics. Our findings are similar to those of Willingham & Morris (1986) who also found that half of the "AP advantage" was accounted for by matching students by background.

Analyzing college grade by different levels of AP exam score is also illuminating (Figure 4). We found that students who are awarded an AP exam score of 1 earned college science grades no better than did those students who entered the college course after having taken a regular course in the subject. Similarly, students who have an AP exam score of 2 do no better than those who have completed an honors course. Since students usually complete either a regular or honors course prior to enrolling in an AP course, it appears as if students reporting low AP exam scores may have gained no benefit from their AP course. This finding contrasts with the belief espoused by many that it is always better to take a more rigorous course and earn a low grade than an easier course and earn a high grade (Adelman, 1999).

Based on the results from the Full Model, those students who enroll in AP and have not taken the AP exam perform in college science about as well as AP 3 students: about 1.5 points higher than the student average. AP 4 students are predicted to earn college science grades 3.4 points above average. AP 5 students are predicted to earn grades 4.6 points higher than average. Though these predicted scores appear to be fairly impressive, we should con-

Figure 4: Comparison of Model A ("Raw Data") and Model E ("Full Model")



sider that the student average sits just above 80 points, amounting to about a grade of B-. An additional 4.6 points would only elevate the course grade to roughly 85 points, suggesting that the average AP 5 student in introductory college science would be predicted to earn a letter grade of B if one accounts for those background variables that are different from most students in their science course.

Next we considered the relative importance of variables in the explained variance of the full model based on the standardized β (in units of the fraction of a standard deviation of college grade earned for one standard deviation change in the independent variable). As one might expect, AP exam score has a substantial role ($\beta = 0.15$). Yet, four other variables have similar ratings: high school math grade ($\beta = 0.16$), SAT Math ($\beta = 0.13$), highest math course taken in HS ($\beta = 0.11$), and average science grade (in other subjects ($\beta = 0.10$). Taken as a group, these are arguably stronger predictors of college grade than performance in high school courses preparatory to an AP course ($\beta = 0.08$). This may reflect the importance of mathematical knowledge and reasoning ability for college students in the sciences (Burton & Ramist, 2001).

Poor performance in AP courses is often attributed to a concern over the "lack of academic readiness" for advanced placement courses, particularly a deficit in the "possession of prior relevant knowledge (Dougherty *et al.,* 2006b, p. 6)." While one may assume that such knowledge could be acquired from a preparatory course (e.g. regular or honors chemistry for an AP chemistry course), the relatively low standardized coefficients related to theses courses (0.04 for honors level and 0.08 for the grade in this course) tend to invalidate this claim. More important to success in college science appear to be general academic skills, particularly in mathematics. AP exam score had a correlation of 0.404 with SAT verbal scores (Table 2). This can be interpreted as evidence that high verbal functions (both in reading and writing) are necessary to get the most from an AP course (particularly reading the textbook) and to perform well on the AP exam.

There are many possible explanations for why AP exam high-scorers do not consistently attain levels of performance commensurate with stated College Board expectations. We discuss some of these below:

- **Non-equivalency of the AP exams and college science attainment measures.** AP exams may not fully reflect content of the college courses in this study, even though the College Board's efforts are designed to accurately reflect the college curriculum. Certainly there is content at the college level that is uniquely presented by college professors that cannot show up in a standardized AP test. Also, more goes into a college course grade than a single final exam. Students often take additional tests (e.g. mid-terms), are graded on their laboratory work and problem sets, and may be required to pursue a project or paper. No matter how close the AP exam is to the college-level exam, there may be missing elements that do not make the AP test a valid substitute for a college grade in science, however similar it may be to the final exam.

- **Over-generosity in AP Exam Scoring.** "*In general, an AP grade of 3 or higher indicates sufficient mastery of course content to grant a student exemption from a college course, credit, or both (College Board, 2004, p. 6.)"* The College Board is clear in its guidelines, adjusting the cut-points (how exam scores are assigned) of each of its AP examinations so that *"… the lowest AP 5 is equivalent to the average "A" in college, the lowest AP 4 is equivalent to the average "B," and the lowest AP 3 is equivalent to the average "C" (College Board, 2004, p. 11)."* The so-called "composite score boundaries" set the cut-offs for AP exam "grades" of 1, 2, 3, 4, or 5 from the raw scores or "formula scores." The setting of these boundaries is explained on the AP website.[24] An example is given for the Physics C (Electricity and Magnetism Exam). The cut-off score over which students earn a "5" is 19.7 out of 35 problems correct or 56%. The cut-off score over which students earn a "4" is 40% and over which they earn a "3" is 31%.[25] While assigning final "grades" is a function of item difficulty, such a low percentage of items correct translating into a top score is unusual for a test. As a criterion referenced exam, the criteria may be set too low (Lichten, 2000).
- **Lack of Sufficient Content Coverage in AP Exams.** A fundamental limitation of all standardized exams is the amount of content each exam may include. Although exams may be "comprehensive" in some ways, a single Advanced Placement exam has no hope of including the degree of comprehensive content coverage typically included in college coursework. As a result, AP exams are at a fundamental disadvantage, depending largely on an assessment of student content knowledge based on a sampling of student performance. This approach is based on the same premise as this large-scale study; however, the difference lies in its application. While we make no claims as to the performance of individuals based on our analysis, but rather limit our conclusions to generalizations for large groups, AP exams do make the claim of providing a highly accurate appraisal of an individual student's knowledge and ability. The accuracy of these claims must be clear and unequivocal, since decisions for an individual are made on the weight of a single AP exam score. A decision to bypass an introductory course does have potential repercussions for individual students.
- **Weak Methodology in AP Score Validation.** For the process of validation carried out by the College Board, items from the relevant AP exam are administered to college science students at the end of their course. Any laxity of effort on the part of these students to do well would invalidate the equating of AP scores to college grades (Lutkus, Weiss, Campbell, Mazzeo & Lazer 1999). The College Board is silent in its technical manual as to what motivation these students have to do their best on these exams. Also, extensive preparation of AP students with practice exams familiarizes them with the format and probable content. Participating college students enjoy no such "prepping" for the questions that they answer drawn from AP exams and may be unaccustomed to their formulation, earning them lower scores. There is a lack of parity since the College Board does not administer actual college science final exams to AP students while in high school to see if they score at levels reflective of high college grades.
- **Sample Size of the Current Study.** While 8594 students were included in this study, only 937

24. "An AP grade may also indicate that the student should receive credit for a college course that he or she has not taken. To check the validity of such decisions, studies are conducted in which the AP Exam is administered to students taking the college courses in question. The college students' grades on the AP Exam are then compared with their grades in the college course." (http://apcentral.collegeboard.com/article/0,3045,152-167-0-2052,00.html#key)
25. http://apcentral.collegeboard.com/members/article/1,3046,154-181-0-11583,00.html

reported taking the relevant AP course in high school. Of these students 433 reported taking the exam with 282 reporting AP exam scores of 3 or higher. While this study is smaller than several concerning Advanced Placement, it is the largest investigation of its kind that controls for relevant background factors such as prior mathematics and science course-taking and performance. This study offers a snapshot of the actual performance of students in introductory college science in U.S. colleges and universities. While studies with larger samples of AP students abound, most control for few or no background variables. These results offer *little to no information* about the *disaggregated* relationship of Advanced Placement experiences and student performance in college. After controlling for several covariates, our study found that AP course participation and performance on AP tests were indeed related to performance in college science courses, but that the magnitude of the effect was about half that found without controls. Moreover, we do not know if AP experiences help students or if better students help AP outcomes. Given the clear disparities of educational resources in American education, the latter may not be further from the truth than the former.

- **Missing students who "place out" in the current study.** This study, by its very design, does not include any students who have earned high enough AP exam scores to receive credit for an introductory college science course and, hence, do not take it. Many, perhaps most, of the students in colleges that we study who earn a 3, 4, or 5 are simply not in our sample. There is no direct way to know how these students would fare in an introductory course, one can only analyze the grades of those who are actually enrolled in an introductory course, as we have. However, we have established that these high scoring AP students have no apparent deficiencies in other areas. Through our follow-up survey we have also found that most thought retaking the introductory course was beneficial. These students strive for high grades and work as hard or harder than their classmates.
- **AP students were bored retaking introductory courses.** One explanation may be boredom among AP exam takers since the college course may merely be a review of concepts covered in high school. With this issue in mind, we queried students in our survey who offered "further contact" information and who reported high AP exam scores. Our follow-up inquiries found that a majority of the students contacted reported the "re-taking" of the college science course to be beneficial. In addition, we did not find significant differences (at $p \leq 0.05$ for χ^2 test) in level of effort reported in college science between AP students who passed or failed the AP exam in high school.

Conclusions

This study seeks to disentangle two issues: the validity of Advanced Placement exams as predictors of college science performance and the "value-added" of taking AP science courses in high school. AP courses act as both a filter for the most motivated and brightest high school students, while offering an opportunity to learn science in a rigorous, fast-paced environment modeling an introductory college course. We have chosen to examine the performance of college students taking introductory college biology, chemistry, and physics. These students have varied academic backgrounds that allow for the use of multiple regression to model the degree to which specific variables explain differences in college performance.

We find that the students in our sample who reported AP exam scores of 3 or higher earned college grades that were higher than the student average. However, many AP 5 students performed at levels below the College Board's claims of excellence even after taking a semester of the college science. While 73% of AP 4 students earned college science grades above the College Board's stated *"mid-level B performance in college,"* about half of the AP 5 students missed the A level performance cited by the College Board. We should be clear on one very important point, all the students in our survey experienced a semester of actual introductory college science course work. Thus, given that high school AP courses are intended to stand in place of college courses, it can be argued that AP students have taken the introductory college course twice, and despite this clear advantage, their performance is by no means indicative of the benefit many would have us believe Advanced Placement courses would impart to high school students who take them. As a measure of introductory college science content knowledge mastery, the performance of a considerable fraction of AP takers does not appear to adequately reflect the professed notions of the College Board. In fact, in our qualitative probe, many students feel that while AP is a good preparation for college science, many also feel they benefited from taking the actual college science course.

Based on our analysis, it appears that about half of the advantage attributed to AP experience can be accounted for by variables representing the academic abilities and experiences possessed by AP students prior to, or independent of, their AP course experiences. Students' backgrounds, particularly their mathematical and verbal skills, appear to contribute mightily to their performance in college science courses. While the AP examination program is an elaborate system, with its professional development program rivaling its assessment program in size and productivity, the Advanced Placement exams themselves appear to fall short of the predictive validity claimed by the College Board. Based on the findings from our study, AP exams scores of 3 do not appear to warrant the granting of college credit over those students who take an AP course in high school, but do not take the exam. In addition, we found that an AP exam score of 1 or 2 offered little evidence of any benefit derived from the AP course work experienced by the students in high school. In addition, students passing an AP exam might well consider retaking introductory courses in college to more completely master the content, as many colleges and universities already require.

... our study has found indications that the AP program, while certainly of value to many students, may lack some of the evidence necessary to support its claim of academic rigor equal to that of introductory college and university courses in science.

Our findings suggest that the educators, teachers, parents, and students would be better served by:

–An acknowledgment from the College Board that evidence exists indicating the non-equivalence of AP courses and college courses in biology, chemistry, and physics.

–A more balanced mix of both supportive and critical research at the Advanced Placement website.[26]

–The College Board clearly stating their role in funding research studies cited as evidence for AP course effectiveness.

–AP exam scoring made more stringent so that the highest exam score truly represents achievement at the highest level of college work.

–The College Board acknowledging that the later performance of students who pass the AP exams appears to be explained as much by their prior background as their AP course enrollment.

For high schools seeking to evaluate their AP science courses, a preponderance of AP exam scores of 1 or 2 may be interpreted to mean that the AP course in their school offered little or no benefit to students beyond regular or honors level science courses. Schools in which the majority of students score in this range should examine whether exposure to AP-level rigor holds any benefits apart from the "academic padding" coined by Lichten that aids them in the college admission process.

Advanced Placement has become one of the most highly respected "brand-names" in secondary/post-secondary education. However, our study has found indications that the AP program, while certainly of value to many students, may lack some of the evidence necessary to support its claim of academic rigor equal to that of introductory college and university courses in science. While this current study certainly suffers from limitations, as do all research studies, what is clear from our findings is that the claims of the College Board regarding the interpretation and validity of AP exam scores are problematic. Our findings indicate the need for further investigations of greater size, scope, and intensity, especially given the potential being ascribed to the Advanced Placement program as part of the cure for what ails American science education.

References

Adelman, C. (1999) *Answers in the Toolbox: Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment*. Jessup, MD: Educational Publications Center.

Anaya, G (1999) Accuracy of Self-Reported Test Scores, *College and University*, 75(2), 13-19.

Baird, L. (1976) *Using Self –Reports to Predict Student Performance*. Research Monograph No. 7., New York, NY: College Entrance Examination Board.

26. http://apcentral.collegeboard.com/apc/public/colleges/research/index.html

Bradburn, N., Rips, L. and S. Shevell (1987) Answering autobiographical questions: The impact of memory and inference on surveys. *Science*. 236(4798)**:** 157–61.

Bradburn, N. (2000). Temporal representation and event dating. A. A. Stone, J.S. Turkan, C.A. Bachrach, J.B. Jobe, H.S. Kurtzman, V.S. Cain (Eds.), . *The Science of Self-Report*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 49-61.

Burdman, P. (2000). Extra credit, extra criticism. *Black Issues in Higher Education* 17(18), 28-33.

Burton, N. W. and Ramist, L. (2001) *Predicting Success in College: SAT Studies of Classes Graduating Since 1980*. New York, NY: College Board.

Cahow , C. R., Christensen, N. L., Gregg, J. R., Nathans, E. B., Strobel, H. A. & Williams, G. W. (1979) *Undergraduate Faculty Council of Arts and Sciences, committee on Curriculum, Subcommittee on Advanced Placement Report*. Durham, NC: Duke University.

Chamberlain, P., Pugh, R. and Shellhammer, J. (1978) Does Advanced Placement Continue Through the Undergraduate Years? *College & University*, 53, 195-200.

College Board. (2004) *Interpreting and Using AP Grades*. New York, NY.

College Board (2006) *Advanced Placement Report to the Nation*. New York, NY.

Dodd, B., Fitzpatrick, R. and Jennings, J. (2002) *An Investigation of the Validity of AP Grades of 3 and a Comparison of AP and Non-AP Student Groups*. New York, NY: College Board.

Dorans, N. (1999) *Correspondence Between Act and SAT I Scores* (College Board Research Report 99-1), New York: College Board.

Dougherty, C., Mellor, L. and Jian, S. (2006a) *The Relationship Between Advanced Placement and College Graduation*. National Center for Educational Accountability.

Dougherty, C., Mellor, L. and Jian, S. (2006b) *Orange Juice or Orange Drink? Ensuring that "Advanced Courses" Live Up to Their Labels*. National Center for Educational Accountability.

Geiser, S. and Santelices, V. (2004) *The Role of Advanced Placement and Honors Courses in College Admissions*. Center for Studies in Higher Education. Paper CSHE-4-04. http://repositories.cdlib.org/cshe/CSHE-4-04/

Hershey, S. (1990) "College Admission Practices and the Advanced Placement Program." *Journal of College Admission*, 128, 8-11.

Juillerat, F., Dubowsky, N., Ridenour, N., McIntosh, W., and Caprio, M. (1997) Advanced placement science courses: high school--college articulation issues, *Journal of College Science Teaching,* 27(1), 48-52.

Klopfenstein, K. and Thomas, K. (2005) *The Advanced Placement Performance Advantage: Fact or Fiction?* Accessed from http://www.aeaweb.org/annual_mtg_papers/2005/0108_1015_0302.pdf

Klopfenstein, K. (2004) The advanced placement expansion of the 1990s: How did traditionally underserved students fare? *Education Policy Analysis Archives, 12*(68), 1- 14.

Kuncel, N., Credé, M., & Thomas, L. (2005) The validity of self-reported grade point averages, RICs, and test scores: A meta-analysis and review of the literature. *Review of Educational Research,* 75(1), 63 – 82.

Lichten, W. (2000) Whither Advanced Placement*? Education Policy Analysis Archives,* 8(29). http://epaa.asu.edu/epaa/v8n29.html

Lichten, W. (2007) Equity and Excellence in the College Board Advanced Placement Program, *Teachers College Record,* 1/16/07.

Lutkus, A., Weiss, A., Campbell, J., Mazzeo, J., and Lazer, S. (1999) *NAEP 1998 Civics Report Card for the Nation*, Washington, DC: National Center for Educational Statistics, 132.

MacVicar, R. (1988) *Advanced Placement: Increasing Efficiency in High School--University Articulation*. Phoenix, AZ: Arizona Board of Regents. (ED306835)

National Research Council (2002) *Learning and Understanding: Improving Advanced Study of Mathematics and Science in US High Schools*, Committee on Programs for Advanced Study of Mathematics and Science in American High Schools. J. P. Gollub, M. W. Bertenthal, J. B. Labov, and P. C. Curtis, Jr., eds. Center for Education, Washington, DC: National Academy Press.

Niemi, R. and Smith, J. (2003) The accuracy of students' reports of course taking in the 1994 National Assessment of Educational Progress. *Educational Measurement: Issues and Practice*, 22(1), 15-21.

Pace, C., Barahona, D., and Kaplan, D. (1985) *The credibility of student self-reports*. Los Angeles, CA: UCLA Center for the Study of Evaluation

Phillips Academy (1952) *General education in school and college; a committee report by members of the faculties of Andover, Exeter, Lawrenceville, Harvard, Princeton, and Yale*. Cambridge, MA: Harvard University Press.

Pushkin, D. (1995) The AP Exam and the Introductory College Course. *The Physics Teacher,* 33(8), 532-535.

Ruch, C. (1968) A Study of the Collegiate Records of Advanced Placement and Non-Advanced Placement Students. *College & University*, 44(2), 207-210.

Sadler, P.M. & Tai, R.H.(2001) Success in College Physics: The Role of High School Preparation, *Science Education*. 85(2), 111-136.

Sawyer, R., Laing, J. and Houston, W. (1989) Accuracy of Self-Reported High School Courses and Grades of College-Bound Students. *College and University,* 64(3), 288-299.

Willingham, W. and Morris, M. (1986). *Four Years Later: A Longitudinal Study of Advanced Placement Students in College*. New York: College Board Publications.

Wood, W.B. (2000) Advanced High School Biology in an Era of Rapid Change: A Summary of the Biology Panel Report from the NRC Committee on Programs for Advanced Study of Mathematics and Science in American High Schools. *Cell Biology Education*, 123-7.

Philip M. Sadler is director of the science education department at the Harvard-Smithsonian Center for Astrophysics and F.W. Wright Senior Lecturer in Harvard's Department of Astronomy, Cambridge, MA 02138. Correspondence concerning this article may be sent to psadler@cfa.harvard.edu.

Robert H. Tai is assistant professor of science education, Curry School of Education, University of Virginia, Charlottesville, VA 22904Acknowledgements

The authors would like to acknowledge the people who helped make this large research project possible: Janice M. Earle, Finbarr C. Sloane, and Larry E. Suter of the National Science Foundation for their insight and support; James H. Wandersee, Joel J. Mintzes, Lillian C, McDermott, Eric Mazur, Dudley R. Herschbach, Brian Alters, and Jason Wiles of the FICCS Advisory Board for their guidance; and Nancy Cianchetta, Susan Matthews, Dan Record, and Tim Reed of our High School Advisory Board for their time and wisdom. This research has resulted from the tireless efforts of many on our research team: Michael Filisky, Hal Coyle, Cynthia Crockett, Bruce Ward, Judith Peritz, Annettte Trenga, Freeman Deutsch, Nancy Cook, Zahra Hazari, Jamie Miller, Marc Schwartz, and Gerhard Sonnert. Matthew H. Schneps, Nancy Finkelstein, Alex Griswold, Tobias McElheny, Yael Bowman, and Alexia Prichard of our Science Media Group constructed of our dissemination website (www.ficss.org). We also appreciate advice and interest from several colleagues in the field: Michael Neuschatz of the American Institute of Physics, William Lichten of Yale University, Trevor Packer of the College Board, Saul Geiser of University of California, Paul Hickman of Northeastern University, William Fitzsimmons, Marlyn McGrath Lewis, Georgene Herschbach, and Rory Browne of Harvard University, and Kristen Klopfenstein of Texas Christian University. We are indebted to the professors at universities and colleges nationwide who felt that this project was worth contributing a piece of their valuable class to administer our surveys and their students' willingness to answer our questions. This work has been carried out under a grant form the Interagency Educational Research Initiative (NSF-REC 0115649). Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, The U.S. Department of Education, or the National Institutes of Health.

Austin M. Hitt, J. Scott Townsend

Getting to the Core Issues of Science Teaching: A Model-Based Approach to Science Instruction

The authors present analogies as useful instructional tools that improve students' and teachers' understanding by anchoring abstract concepts to concrete and familiar entities.

Introduction

Science teachers are responsible for guiding students through the process of learning and understanding science content. In order to carry out this charge, they need to present information in a manner that is congruent with the way students learn. Educational Psychologist Jerome Brunner summarized this issue when he wrote, "The task of teaching a subject to a child at any particular age is one of representing the structure of that subject in terms of the child's way of viewing things (Brunner, 1960)." Bruner stated that students must be able to "represent" phenomena at three distinct levels: the enactive, iconic, and symbolic (Brunner, 1966). The enactive level pertains to phenomena that can be observed. This is the world of experiences that includes what we learn about the world through our five senses. The iconic level refers to internal representations or models that students use to visualize phenomena. Finally, the symbolic level refers to the abstract scientific symbols, definitions, and formulas that are used to explain phenomena. According the Brunner, in order for students to truly understand a concept they need to have experience with it at all three levels.

Three Conceptual Levels of Scientific Understanding

Macroscopic Level

In order to specifically address the unique perspectives in chemistry and the physical sciences Johnstone (1991) proposes a parallel model that is amenable to these subjects. We refer to this model as the Three Conceptual Levels of Scientific Understanding. Johnstone's conceptual levels are parallel to Brunner's but he renames them the macroscopic, sub-microscopic and symbolic. The macroscopic level refers to phenomena that can be observed directly. At the macroscopic level students experience the phenomenon through demonstrations or hands-on activities. These types of observable demonstrations help students to build connections between experiences and science concepts (Bransford 2004; Bransford & Donovan, 2005).

Sub-microscopic / Particle Level

The sub-microscopic level refers to phenomena that occur at the level of molecules and atoms. In order to more accurately reflect the learning issues in chemistry,

Gabel (1999) asserts that this level be redefined as the particle level. The term particle is generic and can be applied to diverse species such as atoms, ions, or molecules. In order to accurately explain chemical phenomena students need to identify the relevant particles and describe the interactions between the particles.

By constructing their own models, the students can see, touch, and manipulate abstract information which help them develop a deeper understanding of the scientific concepts.

For example, research has revealed that students, at diverse developmental levels, have naïve or incomplete understandings of the process of burning (Driver, Squires, Rushworth, & Wood Robinson, 1994; Gabel, Monahan, MaKinster, & Stockton, 2001). Studies have shown that students recognize that oxygen is necessary but they frequently do not understand the role of oxygen in burning. Because students can not conceptualize how oxygen interacts at the particle level, they tend to envision burning as a process in which the starting material is destroyed.

At the macroscopic level, this is a reasonable conclusion based on common experiences such as burning a log or newspapers. The starting material breaks apart and the overall product, ash, appears to have less mass because of the loss of material in the form of smoke, carbon dioxide and other products. In reality, burning involves the addition of oxygen to materials which results in new products that have a greater cumulative mass than the starting material. However, students cannot observe how the particles of the starting material interact with oxygen. What they observe is the ash that looks and feels lighter than the original log or newspaper. Based on their macroscopic observations, it is relatively easy for students to develop the misconception that burning results in the destruction and consequent disappearance of matter.

In order to devise an accurate scientific explanation for burning or any other chemical phenomena, students need to be able to visualize the interactions between particles. Because they can not see these particles they need a tool for bridging the gap between their experiences and interactions at the particle level, models. By constructing their own models, the students can see, touch, and manipulate abstract information which help them develop a deeper understanding of the scientific concepts (Gilbert, 1991; Robinson, 2000; Gilbert & Ireton, 2003; Mclachlan, 2003; Hitt & Townsend, 2004).

Models and model building also helps teachers assess their students' knowledge and understanding. Research reveals that model-based instruction can be used to identify students' misconceptions about science content (Greca & Moreira, 2002; Bunce & Gabel, 2002; Taber, 2003) and the nature of scientific inquiries (Niaz, 2001; Justi & Gilbert, 2000). By assessing students' models, teachers can focus instruction on content that is the most difficult for the students to understand.

Because models are so critical to understanding concepts at the sub-microscopic/particle level, we propose that it be reconceived as the "modeling" level (Hitt & Townsend, 2004). We have two reasons for our proposal. First, because all scientific fields use models, in some form, there is a clear connection between the three conceptual levels of scientific understanding and diverse science subjects. This means the modeling concept is compatible with other scientific fields such as biology and meteorology.

Second, redefining the sub-microscopic / particle level aligns it with phenomena at significantly different scales. For example, a high school biology class can address concepts that range from processes at the molecular level, such as the replication of DNA, to ecosystems which encompass thousands of biotic and non-living variables. Students can not directly observe these concepts in the classroom but they can visualize them by creating models (Robinson, 2002; Mclachlan, 2003).

Symbolic Level

The symbolic level refers to scientific formulas, equations, and definitions. Students have the greatest difficulty understanding science content at the symbolic level because it is unfamiliar to them (Johnstone, 1991, Gabel, 1987; 1999; Hitt & Townsend, 2004). Unfortunately, science instructors tend to focus on the symbolic level exclusively which can result in students becoming disinterested and simply memorizing definitions and formulas (Gabel, 1993; 1999). Research on learning demonstrates that memorization is ineffective because it does not result in connections between concepts and the information is stored relatively briefly (Bransford, 2004; Bransford & Donovan, 2005).

Research on learning demonstrates that memorization is ineffective because it does not result in connections between concepts and the information is stored relatively briefly.

In this article we present an analogy, *Apple Activity*, designed to help science teachers and students reflect upon science instruction and learning at the three conceptual levels of scientific understanding. Analogies are useful instructional tools that improve individuals' understanding by anchoring abstract concepts to concrete and familiar entities (Lakoff & Johnson, 1980). This analogy provides a simple and familiar situation designed to help science teachers reflect on their students' perspectives. Too often science instructors, including the authors, teach science through traditional lectures that mainly address the symbolic and macroscopic levels. This approach is logical for science teachers who are experts but it can be confusing for students who are relative novices. The *Apple Activity* creates an analogy that helps science instructors to "see" the differences between experts and novices. After we introduce the *Apple Activity* we discuss how it connects the Three Conceptual Levels of Scientific Understanding to science content and effective teaching practices.

Activity Setting

This activity is presented in the form of a dialogue between Dr. Smith, the professor for a secondary science methods course, and his pre-service science teachers. We decided to present the activity as a vignette in an attempt to provide readers with a student's perspective. When we present this activity to teachers and students we assume the role of Dr. Smith.

Scenario

Dr. Smith starts class by stating, "Good morning class. I can tell you are eager to learn how to teach science. Let's start class with a simple concept." He then places a transparency on the overhead (Figure 1). "Now I want you to create a list of all the ideas that are relevant to this concept."

蘋果

Figure 1: Grammatical symbols for a common object

After waiting a few moments Dr. Smith looks around the room and observes the confused expressions on his students' faces. He sardonically responds, "Oh this is ridiculous. Everyone in this class has seen this before! How can we even begin if you don't know such a basic concept?"

After waiting a bit longer, Dr. Smith decides to help the class out. "Okay class I know that it is early in the semester and you may need help in getting started so I will provide you with a different representation of the same concept." He replaces the first overhead with a second overhead (Figure 2).

사과

Figure 2: Second set of grammatical symbols for a common object

Dr. Smith asks, "Now can anyone give me some information about this concept?" The students continue to stare at the overhead and all the coaxing that Dr. Smith can muster fails to get responses from the class.

Dr. Smith finally concedes that the students are not familiar with the concept and that he must start at a more basic level. "Maybe I'm just introducing this problem the wrong way. Perhaps we need to approach this concept from a different perspective." He reaches into a brown grocery bag and distributes bright red apples to the class (Figure 3).

Figure 3 : Image of an apple

"What can you tell me about apples?" The students look at Dr. Smith wryly. "Now I want everyone to write down a list of observations that you can make for your apple." Initially the class is reluctant but with some encouragement they use their senses to investigate the apples.

After discussing the students' observations Dr. Smith distributes a collection of plastic apples (Figure 4). "These are plastic models of real apples that I purchased at a hobby store. Notice they look nearly identical to the real apple. I now want you to look at your list of observations for the real apple and cross out any observations that do not apply to the plastic model."

Figure 4: Image of a plastic model of an apple

Dr. Smith then leads a class discussion about the similarities and differences between the plastic models and the real apples. "Now class you probably noticed that the models have some similarities and some differences compared to real apples. This is because models are not exact replicas but are conceptual representations of the real phenomenon or target."

"You have probably seen many types of models such as physical models, computer models, mathematical models, and graphical models just to name a few. Despite their different appearances, all models have the common function of explaining and connecting concepts associated with the target. In our case we are using a simple physical model that addresses the concepts associated with apples. For example the idea that apples are 'red' is a concept that is built into our model."

"Now class does anyone have any difficulty relating this model to a real

Despite their different appearances, all models have the common function of explaining and connecting concepts associated with the target.

apple?" The students quietly indicate that they have no difficulty with this concept. "Good! Now we will examine two more models of apples." Using the overhead, Dr. Smith shows the class a two dimensional model of a red apple (Figure 5).

"Notice that this model is more abstract then the plastic model of the apple. In order to understand it you need to have a deeper understanding of apples and the concepts associated with apples. Now mark off all of the observations for the real apple that do not pertain to this model."

Figure 5 Colored two dimensional model of an apple

After the class discusses the common characteristics of the apple diagram and real apples, Dr. Smith places a simple black and white model of an apple on the overhead (Figure 6). He again instructs the students to mark off any observations of the real apple that do not apply to the new model.

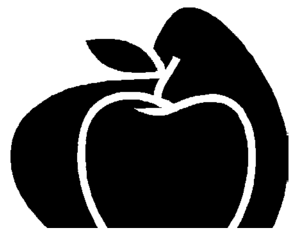


Figure 6: Two dimensional two-toned model of an apple

Finally, Dr. Smith is ready to share with the class the most abstract representation of an apple (Figure 7). "Now I want you to eliminate all of the attributes of the real apple that the word 'apple' lacks."

APPLE

Figure 7: Symbolic representation of an apple

Dr. Smith asks the class to share any observations they recorded for the real apple that apply to the word "apple." After another moment of silence Mr. Smith asks, "So none of you have any observations for this symbolic representation? Does that mean it has no meaning to you? Of course not! I bet the word apple conjures up a lot of images."

"The word 'apple' is a symbol or abstract representation of an apple. In fact no physical or conceptual information about an apple is observable. It is greatly simpler than a model and if you understand it you also have a firm understanding of the physical and diagrammatic models of an apple." Dr. Smith again shows the overheads for the word "apple" written in the unfamiliar languages. "I bet you can recall many images and ideas that are connected to these symbols or models." The students nod in agreement.

"Now I want to change the focus of our discussion. Imagine you are a student in a science class and one of the first things you are expected to learn is a definition, formula, or symbol. You are told that it is a basic concept and you must learn it because all future information will be built upon it. You are not familiar with this elementary concept so it might as well be in a foreign language. Can you imagine how you might feel at this moment?"

Dr. Smith walks over to the board and writes a formula for a chemical reaction (Figure 8). "Imagine that you are in Chemistry I and the topic is chemical reactions.

$$Cu\ (S) + 2AgNO_3(Aq) \rightarrow 2Ag\ (s) + Cu(NO_3)_2\ (aq)$$

Figure 8: Symbolic formula for the reaction of copper wire and silver nitrate solution

Based on what you have gleaned from the *Apple Activity* can you tell me how I could introduce this chemical reaction without using this relatively abstract formula?"

Discussion

Science teachers at all levels live for those "AHA" moments when the light goes on and their students grasp a concept. By sharing the *Apple Activity* with pre-service and inservice teachers and college faculty, we have observed many "Aha" moments. The *Apple Activity* provides educators at all levels with a fresh perspective on the problems their students face in their classes.

The *Apple Activity* also helps educators focus on the attributes of scientific models. Frequently science educators use a variety of models to

Science teachers at all levels live for those "AHA" moments when the light goes on and their students grasp a concept.

explain phenomena but they do not explain the assumptions built into the models. Research in science education reveals that students maintain naïve conceptions about scientific models (Grosslight, Unger, Jay, & Smith, 1991). Without explicit reflection on the similarities and differences between the model and its target, students can internalize misconceptions. For example, we have used ball-and-stick models in order to teach high school students about molecular structures. The models were effective at teaching them the structure of molecules but the models also introduced misconceptions. After questioning our students, we found that many of them believed that chemical bonds were "sticks" or "beams" that held atoms together. These misconceptions could have been avoided by having our students explicitly reflect on the nature of models (Eichinger, 2005).

Another use of the *Apple Activity* is to help students reflect on the process of learning science. Many students find science intimidating because of all of the scientific jargon used to describe phenomena. The *Apple Activity* is a useful tool for addressing their apprehensions by informing them how the science content will be presented in class.

After completing the *Apple Activity* a chemistry instructor could tell her class, "The way we are going to approach chemistry this semester is analogous to the *Apple Activity*. The lab is like investigating the 'real apple' because you are making observations and inferences for real chemical reactions. Initially our lectures will be like examining the 'apple models' because we will be building molecular models and drawing cartoon-like images of molecules. Later, the lecture will be similar to identifying the 'word for apple' when we translate our observations and models into chemical symbols and formulas. By the end of this course you will have 'seen' chemical phenomena, created models in order to visualize the phenomena, translated your observations and models into a symbolic language, and used this scientific language to communicate with others."

Finally, the *Apple Activity* is a useful introduction to our main objective, improving classroom instruction by using the Three Conceptual Levels of Scientific Understanding (Johnstone, 1991; Gabel, 1999; Hitt & Townsend, 2004).

All three of the levels are interconnected and are essential for promoting student learning (Figure 9). Omitting a step disrupts conceptual learning and prevents students from making connections between real world phenomena, models, and scientific concepts.

The easiest level for students to understand is the macroscopic level which refers to observations derived from the five senses. In the *Apple Activity* this level equates to the students using their senses to observe the real apple.

The particle / modeling level consists of constructing models in order to visualize phenomena. They can be physical models, diagrams, verbal descriptions, computer models etc. Students can see and manipulate the models and effectively internalize scientific concepts. An example of this level in the *Apple Activity* would be the plastic apple and the apple diagrams.

The most abstract and difficult level for students to master is the symbolic which consists of the definitions, symbols, and formulas that describe phenomena. This level is difficult because it is comparable to learning a new language. The symbols used to describe the phenomenon are complete abstractions and provide no clues about the phenomenon. In the *Apple Activity*, the symbolic level is represented by the various words for apple.

In order to demonstrate the levels of scientific understanding in our science methods courses we select a concept such as a chemical reaction. For example we use the reaction listed in the scenario which involves a copper wire placed in a dilute solution of silver

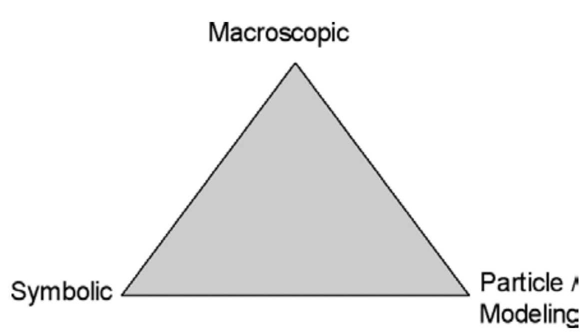


Figure 9: Three conceptual levels of scientific understanding

nitrate. We have students combine the reactants and then observe what happens at the macroscopic level. The copper wire and silver nitrate react to produce a silver precipitate and a bluish copper/nitrate complex (Figure 10).

Figure 10: Copper wire and silver nitrate reaction

First students observe the characteristics of a chemical reaction at the macroscopic level such as a color change, precipitate formation, and temperature changes. We then point out that all of the macroscopic characteristics of chemical reactions can also be observed during physical changes. For example: color changes occur when a dye is added to water, sugar precipitates out of tea when the tea becomes saturated, and temperature changes occur when water evaporates.

We emphasize that observations are only the first steps to understanding a phenomenon. By definition a chemical reaction is the breaking and forming of chemical bonds which results in the rearrangement of atoms and molecules. In order to visualize a chemical reaction students need to construct particle models (Figure 11). The model building process can produce diverse results due to the developmental levels of the students and the requirements of the course.

The critical requirements for a satisfactory model for a middle school science class or high school physical science course may be limited to balancing the equation and including the appropriate atoms and molecules (Figure 11a). After constructing the model the students can discuss how the particle model is different from the real chemical reaction. They can note differences such as the model of the Copper (II) nitrate molecule, $Cu(NO_3)_2$ and the model of Silver Nitrate $AgNO_3$ is a solid particle but in the actual chemical reaction the copper, silver and nitrate ions are aquated in solution, $2Ag^+ + 2NO_3^-$ and $Cu^{2+} + 2NO_3^-$. This process of comparing the model to its target helps the students recognize the limitations of their models and facilitates the development of more accurate scientific models (Figure 11b) (Eichinger, 2005).

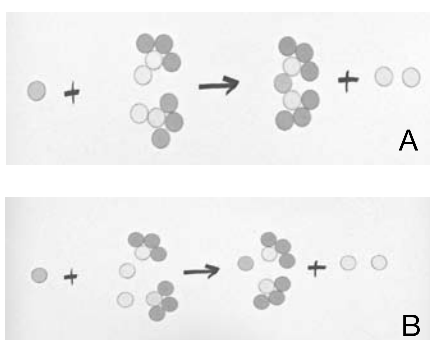


Figure 11: Student models for the reaction between solid copper and silver nitrate. (A) Lower level model that shows the Cu and Ag atoms covalently bonded to the NO_3 molecules. (B) Relatively more advanced model depicting the aqueous nature of the reaction. The interaction between Cu and Ag with NO_3 is not covalent but represents an interaction between opposite charges.

The most abstract and difficult level for students to master is the symbolic which consists of the definitions, symbols, and formulas that describe phenomena.

In contrast, high school or college-level general chemistry students will be expected to produce more complex models and more detailed critiques of their models. For example, they may be required to discuss the structure or construct a model that accurately displays the blue species of copper $[Cu(H_2O)_6]^{2+}$ or create a model that represents a skeleton equation in which NO_3 is simply a spectator ion. After the model is constructed the process is the same as for younger students. The students compare and contrast their model to the target and adapt their model to meet the appropriate scientific explanation.

Regardless of the developmental level, we have found that when students construct models it helps them to visualize and communicate their ideas. Students, ranging from middle level to college-graduates, frequently tell us that this is the first time that they can "see" what occurs during a chemical reaction.

Finally, the students create a symbolic formula in order to describe a chemical reaction (Figure 6). The formula for a chemical reaction is a compilation of symbols that do not resemble the actual reactants and products. Because the terms in the formula are not intuitive and composed of symbols we prefer to designate the chemical formula as a symbolic representation. In our experiences we have discovered that some students and educators prefer to interpret the formula as a model or envision it as a hybrid of symbols and a model. In any case, the chemical formula is more abstract than the particle model and discrepancies over the nature of a chemical formula can engage students and educators in interesting scientific discussions.

Final Thoughts

The three conceptual levels of scientific understanding can be used to improve science instruction by aligning it with the way students learn science. In order to convince secondary inservice and preservice teachers of the efficacy of this model we utilize the *Apple Activity*. The activity has received positive feedback from both groups. Based on our experiences, we believe that it can also be used with inservice and pre-service teachers at the elementary to middle school levels. When science instruction is based on the three levels of scientific understanding it becomes compatible with the way students learn. As a result science teachers can help their students master the "core" ideas in science.

References

Bransford, J. D. (2004). *How people learn: brain, mind, experience, and School*.Washington, D.C.: National Academy Press.

Bransford, J. D. & Donovan, M. S. (2005). Scientific inquiry and how people learn. In J. Bransford & M. Suzanne Donovan (Eds.), *How Students Learn History, Mathematics, and Science In The Classroom* (pp. 397-421). Washington, D.C.: National Academy Press.

Bruner, J. S. (1960). *The process of education*. New York: Vintage Books.

Bruner, J. S. (1966). *Toward a theory of instruction*. New York: W.W. Norton & Company Inc.

Bunce, D. M. & Gabel, D. (2002). Differential effects on the achievement of males and females of teaching the particulate nature of matter. *Journal of Research in Science Teaching 39*(10): 911-927.

Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (1994). *Making Sense of Secondary Science*. New York: Routledge / Falmer.

Eichinger, J. (2005). Using models effectively: how to guide students through age-appropriate, critical analyses of instructional models. *Science and Children 42*(7): 43-45.

Gabel, D., Samuel, K.V. & Hunn, D. (1987). Understanding the particulate nature of matter. *Journal or Chemical Education 64*(8): 695-697.

Gabel, D. (1993). Use of the particle nature of matter in developing conceptual understanding. *Journal of Chemical Education 70*(3): 193-194.

Gabel, D., Briner, D. & Haines, D. (1992). Modeling with magnets, a unified approach to chemistry problem solving. *The Science Teacher 59*(3):58-63.

Gabel, D. (1999). Improving teaching and learning through chemistry education research: a look to the future. *Journal of Chemical Education*, *76*(4): 548-554.

Gabel, D., Monahan, D. L., MaKinster, J. G., & Stockton, J. D. (2001). Changing children's conceptions of burning. *School Science and Mathematics 101*(8):439-449.

Gilbert, S. W. (1991). Model building and a definition of science. *Journal of Research In Science Teaching*, *28*(1): 73-79.

Gilbert, S. W. & S., Ireton, S. W. (2003). *Understanding models in earth and space science*. Arlington: National Science Teachers Association Press

Greca, I. L., & Moreira, M. A. (2002). Mental, physical, and mathematical models in the teaching and learning of physics. *Science Education, 85*(6), 106-21.

Grosslight, L., Unger, C., Jay, E., Smith, C. L. (1991). Understanding models and their use in science; conceptions of middle and high school students and experts. *Journal of Research In Science Teaching*, *28*(9) 799-822.

Hitt, A. & Townsend, S. (2004). Models that matter. *The Science Teacher,* 71(3), 29-31.

Johnstone, A. H. (1991). Why is science so difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning* 7(1), 75-83.

Justi, R. & Gilbert, J. (2000). History and philosophy of science through models: some challenges in the case of the atom. *International Journal of Science Education, 22*(9): 993-1009.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Mclachlan, J. C. (2003). Using models to enhance the intellectual content of learning in developmental biology. *International Journal of Developmental Biology 47*(2): 225-229.

Niaz, M. (2001). A rational reconstruction of the origin of the covalent bond and its implications for general chemistry textbooks. *International Journal Of Science Education, 23*(6): 623-641.

Robinson, W. R. (2000). Learning about atoms, molecules, and chemical bonds: a case study of multiple-model use. *Journal of Chemical Education 77*(9): 1110-1111.

Taber, K. S. (2003). Mediating mental models of metals: acknowledging the priority of the learner's prior learning. *Science Education*, *87*(10):732-758.

Austin M. Hitt is assistant professor of secondary science education, Spadoni College of Education, Coastal Carolina University, Conway SC 29528. Correspondence concerning this article may be sent to amhitt@coastal.edu.

J. Scott Townsend is assistant professor of science education, Eastern Kentucky University, Richmond, KY 40475.

Robert H. Tai, Christine Qi Liu, Xitao Fan

Factors Influencing Retention of Mathematics and Science Teachers in Secondary Schools—A Study based on SASS/TFS

Using a sample of mathematics and science teachers extracted from the 1999-2001 SASS/TFS surveys, connections between teacher and school district characteristics regarding retention are investigated to offer insight into how mathematics/science teacher recruitment might be focused.

Introduction

Shortages of qualified science and mathematics teachers are a ubiquitous problem in the United States. Many schools face both sides of the problem: recruitment and retention of those teachers. Since bringing new teachers in and keeping them are equally important, where should school districts look for teachers that are most likely to stay? While many school districts cannot afford to be too choosy in where they look, their limited resources also mean that they cannot look everywhere. Therefore, information on who are more likely to stay may be valuable for districts to focus their resources for maximum impact.

Background

While recruitment and retention are clearly linked, these topics represent very different aspects of school staffing. With respect to recruitment, mathematics and science education research appears to be focused on attracting individuals with mathematics and science backgrounds into teaching as a profession (e.g. Moin, Dorfield, & Schunn, 2005; Tomanek & Cummings, 2000; Wang, 2004). With respect to retention, the focus appears to be on issues of work environment and new teacher induction (Ingersoll, 2001; Ingersoll, 2003; Luft & Patterson, 2002; Luft, Roehrig, & Patterson, 2003; Patterson, Luft, & Roehrig, 2003; Smith & Ingersoll, 2005). Others suggested that professional preparation prior to recruitment may play a more important role (Kirby & Grissmer, 1993; Murnane, 1987; Reynolds, Ross, & Rakow, 2002; Rhoton & Bowers, 2002; Weld, 1998).

Teacher certification has evolved from the familiar college and university-based teacher education programs to a multitude of different forms and formats.

With the proliferation of alternative teacher certification programs, the backgrounds of individuals entering teaching have shifted. Teacher certification has evolved from the familiar college and university-based teacher education programs to a multitude of different forms and formats. Many programs have structures designed to attract particular groups of young college graduates or working professionals. Other programs have been designed specifically to dovetail with the life demands faced by individuals who have science backgrounds and an interest in teaching. For example, *Teach for America* tends to attract recent college graduates with little or no teaching experience, while community-based alternative certification programs offering evening and weekend classes tend to attract people who hold full-time jobs, but wish to transition into teaching. Some mid-career programs are more time intensive and require full-time enrollment, while still others place pre-service teacher in schools with limited teaching loads and mentors (e.g. *New York City Teaching Fellows Program*).

Given our focus, a comparison of traditional and alternative certification programs was not our purpose. Rather, we have chosen to analyze various teacher and school characteristics to provide a snapshot of the interactions among these factors and their association with the likelihood of teachers to remain in their original schools. We expect that the current study will offer some insight to those seeking to maximize their recruitment efforts by effectively keeping the mathematics and science teachers.

For this analysis, we have selected to study the influence of the following teacher characteristics: age, educational background, salary satisfaction, teacher experience; along with the following school characteristics: school-related earnings, sector (public or private), and urbanicity (urban, suburban, or rural). We have also controlled for the demographic background variables, gender and race/ethnicity. Our analysis included both mathematics and science teachers.

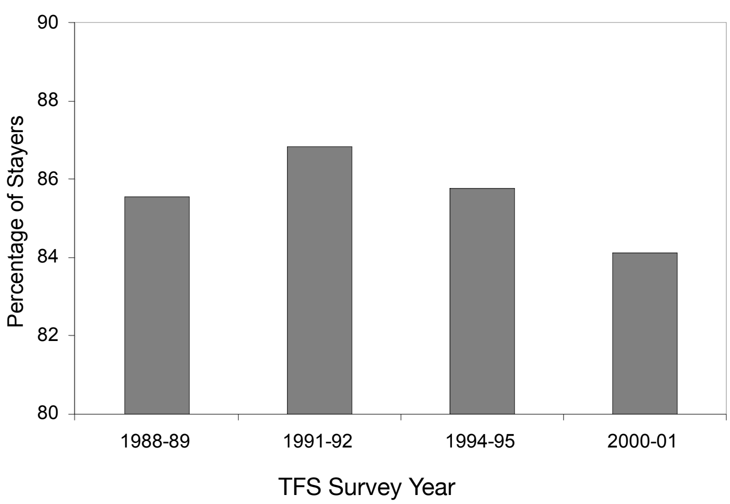
Methods

Data Source

The data from our study was obtained by linking two large-scale educational surveys from the *National Center for Educational Statistics* (NCES), the *Schools and Staffing Survey (SASS) of 1999-2000* and the *Teacher Follow-up Survey of 2000-2001*. The fourth in this series, the *SASS* data, collected by U.S. Census Bureau for *NCES* from a random sample of schools, were stratified by state, public/private sector, and school level (NCES, 2004). After one year, the same schools were again contacted and a representative sample of teachers was asked to complete the *Teacher Follow-up Survey (TFS)*, including those in the original sample who had left their teaching jobs. The survey also obtained information relevant to their departures. In earlier studies by Ingersoll et al. (2001; 2003; Smith & Ingersoll, 2005), teachers were placed into three categories: "leavers" for those who left the teaching profession; "movers" for those who moved to other schools; and "stayers" for those who remained at their original schools. However, for the purpose of this study, teachers not remaining in their original institutions are the primary concern, so we have chosen to group "leavers" and "movers" together into a single group "non-stayers", and compare them to the "stayers". Figure 1 shows a graph of the percentage of stayers over the four SASS/TFS surveys.[1] The most recently available *SASS* 1999-2000 collected data from approximately 52,000 teachers in 12,000 schools. *TFS* 2000-2001 obtained information on 5788 teachers, among them 3473 non-stayers and 2315 stayers.

The results suggest that overall, rural districts struggle most in retaining teachers, both new and experienced.

Figure 1: Percentage of stayers, teachers remaining in their original schools, from the four SASS/TFS surveys*



* Data obtained from Luekens, Lyter, & Fox (2004), p. 8. These are percentages of stayers regardless of subject areas.
** SASS/TFS Surveys were not performed in 1997-98.

Analyses

The selected sample for analysis included 916 mathematics/science teachers who completed both *SASS* and *TFS* in 1999 – 2001. Excluding teachers who retired (n = 137) and 34 others who were not regular teachers (i.e. substitutes, administrators, staff), the final sample included 745 regular mathematics and science teachers. The sample included 304 stayers and 441 non-stayers, with 327 male and 418 female, 394 in mathematics and 351 in science.

The independent variables used in this analysis came from the 1999-2000 SASS teacher questionnaires. As a measure for educational background differences, we included variables identifying teachers who held Advanced Degrees in science/math-

1.Note that the spacing of the first three SASS/TFS surveys was carried out in 3 year intervals; however, no SASS/TFS surveys were carried during the interval beginning in 1997.

ematics and those who held Advanced Degrees in education in general. For Salary Satisfaction, we included teachers' response to a question asking them to rate their degree of satisfaction in salary on a 4-point Likert-type scale. For teacher experience, we compared teachers with 3 or fewer years of experience (New Teacher) to those with more than three years of experience (Experienced Teacher). Four teacher age groups were used: less than 30 years old; 30 – 39; 40 – 49; 50 years and older. The teacher-reported school-related earnings (School Earnings) were grouped into 4 categories: 1) less than 20,000; 2) 20,000 – 29,999; 3) 30,000 – 39,000; and 4) 40,000 or more. Sector accounted for differences between Public and Private schools. Finally, Urbanicity classified by Federal Information Processing Standards as used by the U.S. Census (NCES, 2004) was included to account for differences in the geographical locations of schools in proximity to population centers ranging from rural to suburban to urban. The analysis also included tests for interactions among these predictors. The binary format of the outcome comparing Stayers to Non-stayers indicated that logistic regression was most appropriate for this analysis. We used the logistic regression module available in SPSS 14.0.

Table 1: Binary logistic regression model with interactions (N = 745)

	B	S.E.	Odds Ratio
Constant	-0.72	0.76	0.49
Demographic Background			
Male	0.05	0.16	1.05
Asian	-0.95*	0.47	0.39
Black	0.49	0.40	1.64
Hispanic	-0.07	0.43	0.93
Native American/Am. Indian	-1.06	0.83	0.35
Educational Background			
Advanced Math/Science Deg.	-0.62**	0.24	0.54
Advanced Education Deg.	-0.32	0.23	0.73
School Characteristics			
School Earnings	0.38***	0.10	1.46
Private	0.24	0.18	1.28
Urbanicity	-0.89*	0.35	0.41
Teacher Characteristics			
Salary Satisfaction	0.31***	0.08	1.37
New Teacher	-1.20*	0.57	0.30
Age	-0.31	0.25	0.74
Interactions			
New Teacher * Urbanicity	0.63*	0.26	1.88
Age * Urbanicity	0.31**	0.12	1.37
Nagelkerke Pseudo R^2	0.15		

***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$.
B: regression coefficient; S.E.: standard error; odds ratio: odds ratio of independent variables = e^B.
School sector is a dichotomous variable with private school = 1 and public school = 0.
Urbanicity is an ordinal variable with higher values indicate more urban areas.
Teacher earning from school (an ordinal variable with higher values indicate higher salary), year at the current school, whether they have an advanced degree in math/science (yes = 1), whether they have an advanced degree in education (yes = 1), whether they are new teachers with less than 3 year experience (yes = 1). An advanced degree was defined as a degree beyond a baccalaureate.
Satisfaction in the salary (higher value indicates higher satisfaction), satisfaction in general (higher value indicates higher satisfaction), and self-rated intention to remain in teaching (higher value indicates more inclined to stay in teaching).

Results and Discussion

Through our analysis we arrived at the regression model shown in Table 1. The model with a Nagelkerke R^2 of 0.15 included the main effects variables we listed in the previous section and two significant interactions: *New Teacher* by *Urbanicity* and *Age* by *Urbanicity*.

We approach the discussion in two parts, beginning with a discussion of the significant main effects and then proceeding to a discussion of the interactions. Applying this approach separates the significant effects accordingly, main effects only:[2] *Advanced Math/Science Degree*, *School Earnings*, and *Salary Satisfaction*; and interactions: *New Teacher by Urbanity*, *Age* by *Urbanicity*. Here, school earnings and salary satisfaction have a very low correlation of 0.12 which allowed them to be entered into the logit model together.

2.The *Asian* Race/Ethnicity grouping included only 27 teachers. Given the size of all non-white Race/Ethnicity groupings, the inclusion of these variables were as controls, rather than as robust means of uncovering differences. As a result, we caution against the direct interpretation of the significance of the *Asian* Race/Ethnicity main effect.

We begin by considering educational background of teachers. The logit model coefficient for *Advanced Math/Science Degree* is negative, indicating that those who held advanced mathematics/science degrees are less likely to stay in their original schools than those who did not hold these degrees. Calculating the reciprocal of the odds ratio reported in Table 1 yields a value of 1.85, indicating that advanced mathematics/science degree holders were 1.85 times more likely not to stay (i.e. move to another school or leave teaching all together) than those who did not hold these degrees. The odds ratio result for *Advanced Education Degree*, however, was not significant.

School Earnings account for the dollar amounts teachers were paid by schools in the logit model and were found to be the most statistically significant factor in the model. The findings indicate that after controlled for other variables including teacher experience, teachers in a higher earning bracket were 1.46 times more likely to stay than a lower one (e.g. teachers earning $30,000 - $39,000 compared to teachers earning $20,000 - $29,000). While this result is certainly not surprising, arguments continue to be made for why teacher salaries should not be increased.

Salary Satisfaction is a rating of teachers' satisfaction with their current earnings. To some degree, this variable offers some insight into teacher pay and local area cost of living. For example, earning $50,000 in Ruckersville, VA, is very different from earning the same amount in Brooklyn, NY. In addition, this variable reflects an individual's psychological perception, which can be different from the actual salary level. Here the results indicate that teachers who reported a higher salary satisfaction are 1.37 times more likely to stay in the same school than teachers who reported lower satisfaction (e.g. teachers who were "strongly satisfied" compared to "somewhat satisfied").

... the findings suggest that for urban and suburban school districts, after teacher experience has been taken into account, recruitment might best be focused on mid-career pre-service teachers.

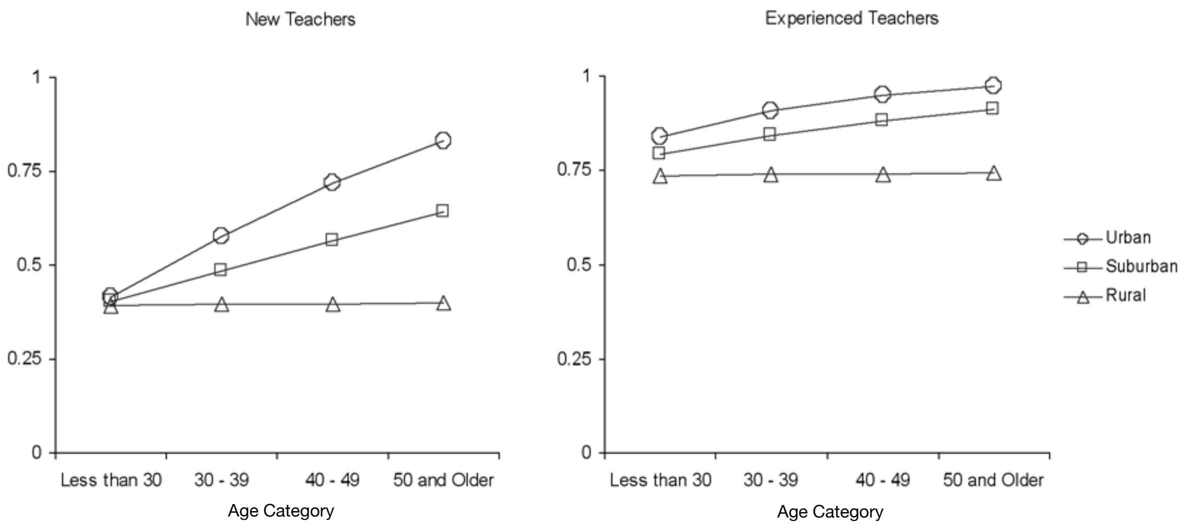
Given the complexity introduced by the interactions, a graphical display (Figure 2) of the estimated probabilities offers a clearer picture of the results. The interactions involve three variables: teacher experience, teacher maturity, and proximity to urban population centers. Figure 2 shows two panels for comparisons between *New Teachers* and *Experienced Teachers*. The x-axis of the graphs displays differences across *Age*. The y-axis represents the estimated probability of teachers staying in their original schools, with values ranging from 0 to 1. Finally, the three trajectories juxtapose the differences across *Urbanicity*. Here, we can see that the estimated probabilities for teachers to stay in their original schools varied across the three different predictors involved in the interactions. For rural schools, the estimated probabilities appear to be fairly flat, indicating little variation across *Age*, for both *New Teacher* and *Experienced Teacher* groups. The results suggest that overall, rural districts struggle most in retaining teachers, both new and experienced. For suburban and urban schools, the findings suggest that among *New Teachers*, the highest estimated retention rates appear to be for teachers in the older cohorts. Here the estimated probabilities indicate that older teachers are predicted to have greater likelihoods of remaining in the same school.3 Among *New Teachers* in suburban schools, teachers in the "Less than 30" group have an estimated probability of 40% of staying, while those in the "50 years and older" group stand at 64%. For urban schools, the difference is even greater, 42% compared to 83% for "Less than 30" versus "50 years and older", respectively. A similar outcome was also found among *Experienced Teachers*, where the estimated probabilities for rural districts remain at about 74% across *Age*. While estimated probabilities of staying for teacher in the "Less than 30" group are 79% and 84%, suburban and urban, respectively; the probabilities for *Experienced Teachers* in the "50 years and older" age groups are above 90%.

Conclusions

Much of the existing research on recruitment has focused on bringing new teachers with science and mathematics backgrounds into teaching. Much of the existing research on retention has focused on work environment and induction as a means of keeping these new teachers in the schools. However, we cannot ignore the economic and social realities that these teachers are faced with when they continue to make the choice of teaching as a profession.

3.Recall that we combined movers and leavers for this analysis and therefore do not distinguish between these groups of non-stayers.

Figure 2: Comparison of the estimated probabilities for teachers to stay in their original schools across Age Categories, Urbanicity, and Teaching Experience, where New Teachers are defined as those with 3 or fewer years of experience



The findings from this study confirmed some commonly understood connections with some new revelations. The connections between earnings *or* salary satisfaction and teacher retention are not new to teaching. Yet here earnings and salary satisfaction were found to be barely correlated and seemed to contribute independently to teacher retention. Another interesting finding was that the results with respect to *Urbanicity* show that while rural districts were consistently below suburban and urban districts in estimated probabilities to retain teachers, urban districts were actually predicted to have higher retention rates when other teacher characteristics and school characteristics are controlled for.

Older teachers are found to be more likely to stay in both urban and suburban districts in this study. Therefore, the findings suggest that for urban and suburban school districts, after teacher experience has been taken into account, recruitment might best be focused on mid-career pre-service teachers. However, for rural districts, the predictions are flat across *Age* for both *New Teachers* and *Experienced Teachers*. The results suggest that while *New Teachers* are predicted to have a roughly 40% likelihood to stay in their original schools, *Experienced Teachers* are predicted to have a 74% likelihood. This differential between rural school teachers with less-than-3-years of teaching experience versus more-than-3-years is very impressive. But does this result suggest that these districts should seek to fill their classrooms with practicing teachers from other districts? In the corporate world, this practice is called "poaching" and in the end, while beneficial for individual teachers, districts and the students they serve suffer. While most high profile efforts have concentrated on large urban districts, these results suggest that for rural districts recruitment faces problems of a different nature.

Should teachers who have accumulated a wealth of knowledge and experience leave teaching, it would be a loss not only to the students and the school, but also to the profession.

The finding that teachers with advanced mathematics/science degrees were more likely to leave the profession or move to another school should not be taken casually. This certainly does not suggest that school districts should stay away from these advanced degree holders. Actually, this result may be partly due to school districts' efforts to get the "cream of the crop": advanced mathematics/science degree holders are more likely to be approached by another school or industry with higher pay or better work environment and thus more likely to leave or move. While higher credentials do not necessarily translate into more knowledgeable about the subject matter or better pedagogy, it is at least equally important for schools to keep these teachers as to recruit them.

The existing literature (e.g. Ingersoll, 2001, Luft, et al. 2004) is very clear on good practices to retain new mathematics and science teachers. While school districts strive to provide better work environment for these teachers, it is beneficial for both teachers and schools to find a good compatibility between the two. Should teachers who have accumulated a wealth of knowledge and experience leave teaching, it would be a loss not only to the students and the school, but also to the profession.

References

Kirby, S. N. & Grissmer, D. W. (1993). *Teacher attrition: Theory, evidence, and suggested policy options*. Santa Monica, CA: Rand Corporation. (ERIC Document Reproduction Service No. ED364533)

Ingersoll, R. M. (2001). Teacher turnover and teacher shortage: An organizational analysis. *American Educational Research Journal, 38(3),* 499-534.

Ingersoll, R. M. (2003). The teacher shortage: Myth or reality? *Educational Horizons, 81(3),* 146-152.

Luekens, M.T., Lyter, D.M., and Fox, E.E. (2004). *Teacher Attrition and Mobility: Results from the Teacher Follow-up Survey, 2000–01* (NCES 2004–301). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Luft, J. A., Roehrig, G. H., & Patterson, N. C. (2003). Constrasting landscapes: A comparison of the impact of different induction programs on beginning secondary science teachers' practices, beliefs, and experiences. *Journal of Research in Science Teaching*, *40 (1)*, 77-97.

Luft, J. A. & Patterson, N. C. (2002). Bridging the gap: Supporting beginning science teachers. *Journal of Science Teacher Education, 13(4)*, 267-282.

Moin, L. J., Dorfield, J. K., & Scunn, C. D. (2005). Where can we find future K-12 science and math teachers? A search by academic year, discipline, and academic performance level. *Science Education, 89*, 980-1006.

Murnane, R. J. (1987). Understanding teacher attrition. *Harvard Educational Review*, *57*, 177-182.

National Center for Educational Statistics. (2004). *1999-2000 Schools and Staffing Survey (SASS) data file user's manual*. Washington, DC: Author.

Patterson, N.C., Roehrig, G. H., & Luft, J. A. (2003). Running the treadmill: Explorations of beginning high school teacher turnover in Arizona. *The High School Journal, April/May*, 14-22.

Reynolds, A., Ross, S. M., & Rakow, J. H. (2002). Teacher retention, teaching effectiveness, and professional preparation: A comparison of professional development school and non-professional development school graduates. *Teaching and Teacher Education*, *18(3)*, 289 -303.

Rhoton, J., & Bowers, P. (2002). *Science teacher retention: Mentoring and renewal*. Arlington, VA: National Science Teachers Association. (ERIC Document Reproductive Service No. ED472332)

Smith, T. M., & Ingersoll, R. M. (2005). What are the effects of induction and mentoring on beginning teacher turnover? *American Educational Research Journal, 41(3),* 681-714.

Tomanek, D, & Cummings, K. E. (2000). The use of secondary science classroom teaching assistant experiences to recruit academically talented science majors into teaching. *Science Education*, *84*, 212-227.

Wang, H.–H. (2004). Why teach science? Graduate science students' perceived motivations for choosing teaching as a career in Taiwan. *International Journal of Science Education, 26(1)*, 113-128.

Weld, J. (1998). Attracting and retaining high-quality professionals in science education. *Phi Delta Kappan*, *79(7)*, 536-540.

Robert H. Tai is an associate professor of science education in Curry School of Education, University of Virginia. His science education research and teaching agenda spans the K-Graduate range and includes large-scale statistical analysis as well as interview and observational studies. His current interests include research to inform national science education policy. Correspondence concerning this article may be sent to rht6h@virginia.edu

---

Christine Qi Liu is currently a doctoral student in the Department of Leadership, Foundations, and Policy at the Curry School of Education of the University of Virginia. Her research interests include applications of statistical methods in educational research and large scale analysis of various educational policy issues. With background in neurogenetics and cognitive psychology, Christine is also involved in research in science education and educational/psychological measurement.

Xitao Fan is a professor of education in Curry School of Education, University of Virginia. He is quantitative methodologist with research interests on quantitative methods, validity and reliability issues in measurement. Substantively, he has used multiple national databases to address issues concerning student academic achievement, gender differences in academic achievement, adolescence self-report response validity, etc.

Carmen Sorge

What Happens? Relationship of Age and Gender with Science Attitudes from Elementary to Middle School

This study examines the attitudes of 1008 students from rural New Mexico in elementary and middle schools from ages 9 through 14. A large decrease in science attitudes between the ages of 11 and 12 years, corresponding with the move from elementary to middle school was observed.

Introduction

This research evolved from the examination of the results of the implementation of a school based science enrichment program involving students from ages 9 through 14. One of the instruments given to the students is designed to measure science attitudes. Over the course of the study it has become obvious that a significant drop in science attitudes is taking place between the ages of 11 and 12 for these students. This research examines the relationship of age and gender on science attitudes.

Why do Science Attitudes Matter?

The study of the relationship between science attitudes and achievement has intrigued many researchers. Meta analysis of 43 studies (Willson, 1983) and 66 studies (Steinkamp & Maehr, 1983) found significant correlational relationships between attitude and achievement in science. Individual studies examining the relationship between attitude and achievement in science have also been examined: for longitudinal relationships (Simpson & Oliver, 1990), for high school students (Reynolds & Walberg, 1992), for eighth graders by gender and ethnicity (Catsambis, 1995), for 10th graders (Young, Reynolds, & Walberg, 1996), by impacts of instructional strategies (Houtz, 1995), for differences across gender over time (Mattern & Schau, 2002), and in Cypress (Papanastasiou & Zembylas, 2002). All found correlational relationships between attitudes and achievement in science.

The study of the relationship between science attitudes and achievement has intrigued many researchers.

Some of the research also narrows the focus by examining the influence of age and/or gender on both science attitudes and achievement. A meta study (Fleming & Malone, 1983) of research utilizing studies from kindergarteners through 12th grade spanning 1960 through 1983 concluded that as age increased the relationship with achievement increased and with attitude decreased. Gender differences indicated a weak superiority of male attitudes and achievement over females. A more recent meta analysis (Weinburgh, 1995) of 18 studies found males to have more positive attitudes than females. The correlation between attitude and achievement was positive for both genders but stronger for females. Catasabis found that female achievement was equal to male achievement but that females had less positive attitudes toward science (Catsambis, 1995). Conversely in China, Boone found that females had more positive attitudes than males (Boone, 1997). The impact of science enrichment programs on attitudes found more changes for females (Stake & Mares, 2001). Utilizing a structural equation model Mattern and Schau found that for males increases in positive attitudes toward science did not lead to greater achievement or more positive attitudes in the future. For females achievement was also not impacted by attitude. However, for females previous positive attitudes toward science did indicate more positive attitudes in the future (Mattern & Schau, 2002).

The impact of age and gender on science attitudes and achievement is plainly an area of interest for those in the field. Most of the research is correlational and does not hypothesize the direction of causality between attitude and achievement in science. It is apparent that positive attitudes and achievement are related.

Methodology

Data Source

The students are located in a rural area outside of Albuquerque, New Mexico. Six elementary schools and three middle schools in two districts are represented. A total of 1080 students (595 females and 485 males) ranging in age from 9 years old to 14 years old are included in the sample.

The data for the present analysis is from a school based science program collected over a three-year period from 2003 to 2005 at the beginning of each fall semester. School begins the middle of August and the data were collected from the end of August through the beginning of October each year. The overall study seeks to examine changes in various attitudes and knowledge for the program. Pretests and posttests are collected each year. The data used here represents only the pretest data, before the students are exposed to the program. Data were collected only once from each student, each age group represents a different group of students, not the same students over a five-year period.

Science Attitude Survey

The Science Attitude scale consists of 10 items designed to evaluate the participant's feelings and attitudes about science such as "Science is fun" and "I would like to learn more about science". These items were culled

The impact of age and gender on science attitudes and achievement is plainly an area of interest for those in the field.

using item analysis after two years of administering a longer 25 item instrument. All questions are scored on a Likert scale from 1 to 6 ranging from Very False to Very True. The total score for the Science Attitude scale is the mean score for the 10 items. Thus the minimum possible score for the Science Attitude scale is 1 and the maximum is 6. All questions are renormed so that a higher value indicates a more desirable outcome. Cronbach's alpha for the science attitude instrument is $\alpha = .96$ using the overall sample. Reliability greater than $\alpha = .7$ is considered good (Litwin, 1995; Nunnaly, 1978).

Results

The initial consideration of possible relationships between the variables developed when results from the first year of data indicated that age was a confounding variable for measuring science attitudes. At first this was thought to be an anomaly, perhaps based on the sample used for that year. However, after collecting pretest data for three years it became unambiguous that the decline with age was non- linear. The mean science attitude scores by age and gender are plotted in Figure 1 below. Statistics for science attitude by age and gender are reported in Table 1.

A two way factorial ANOVA with Gender and Age as fixed factors and Attitude toward Science as the dependent variable was conducted. This method was used rather than regression due to the non-linear relationship between the variables which

Figure 1: Science Attitude Score by Age and Gender

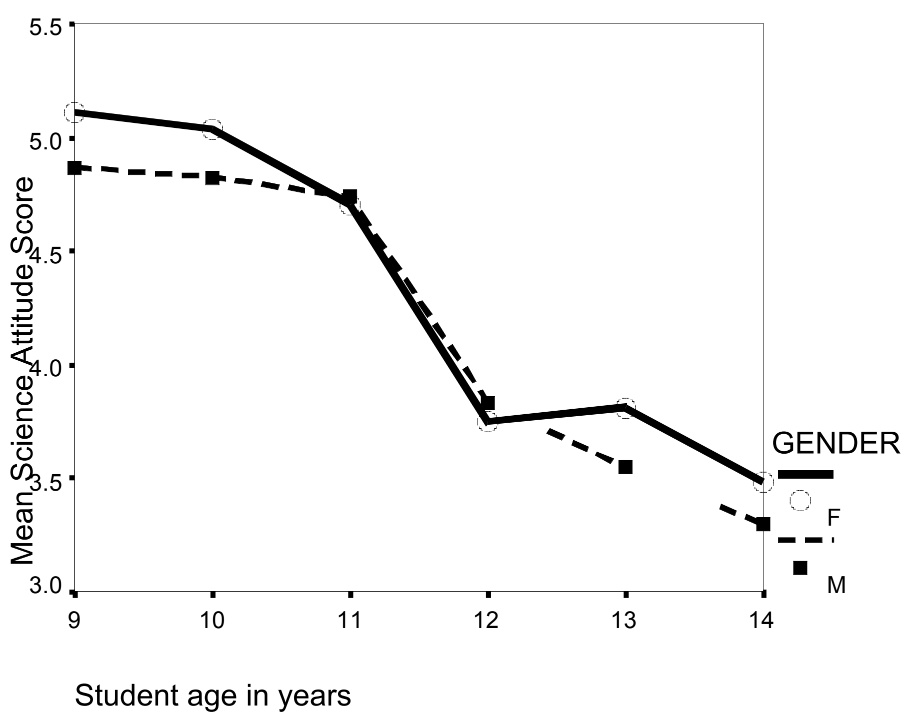


Table 1: Means, Standard Deviation and N for Science Attitude by Age and Gender

Student age in years	Gender	Mean Science Attitude Score	Std. Deviation	N
	Female	5.11	1.07	61
9	Male	4.87	1.13	59
	Total	*4.99*	*1.10*	*120*
	Female	5.03	0.88	86
10	Male	4.82	1.00	76
	Total	*4.93*	*0.94*	*162*
	Female	4.70	1.11	148
11	Male	4.74	1.12	120
	Total	*4.72*	*1.11*	*268*
	Female	3.75	1.33	189
12	Male	3.83	1.33	146
	Total	*3.78*	*1.33*	*335*
	Female	3.81	1.26	81
13	Male	3.55	1.38	69
	Total	*3.69*	*1.32*	*150*
	Female	3.48	1.50	30
14	Male	3.30	1.69	15
	Total	*3.42*	*1.55*	*45*
	Female	4.30	1.33	595
Total	Male	4.28	1.34	485
	Total	*4.29*	*1.34*	*1080*

was observed when plotting the data. Results are presented in Table 2.

No significant interaction between gender and age was found F(5,1068)=.815, p=.539. The relationship of gender was also not significant F(1,1068)=1.92, p=.167. The relationship of age on attitude toward science was significant F(5,1068)=46.88, p<.001, partial η2 =.18, η =.42. This represents a large effect size (Cohen, 1988).

A trend analysis indicated that the data were fit by a linear model (p<.001, η^2 =.15) with a quadratic component (p=.003, η^2 =.01).

Post hoc analyses of the differences across grades were tested with the Games-Howell multiple comparison procedure due to unequal variances (Games & Howell, 1976). Post hoc tests reveal that the 9, 10 and 11 year olds are not significantly different from each other. The 12, 13 and 14 year olds are also not significantly different from each other. However, these two groups of younger and older students do differ significantly from each other. The two homogeneous subsets of science attitude are comprised of the 9, 10 and 11 year olds and the 12, 13 and 14 year olds.

Discussion

In these school districts the move from elementary school to middle school occurs between 11 and 12 years of age. Thus, for the vast majority of students (the exceptions being accelerated students or those held back) science attitudes at age 11 reflect those of the last year of elementary school and those at age 12 represent the first year of middle school. Unmistakably a precipitous drop in science attitudes takes place between elementary school and middle school. As the students have only been attending middle school for a maximum of six weeks before they take the instrument this result is a cause for great concern. The effect

Table 2: Two Way Analysis of Variance for Science Attitude Scores as a Function of Age and Gender

Variable and Source	df	Mean Square	F	Partial η^2
Science Attitude				
Age	5.000	68.794	46.883**	0.180
Gender	1.000	2.811	1.916	0.002
Age * Gender	5.000	1.195	0.815	0.004
Error	1068.000	1.467		

**p<.001

Table 3: Means for Homogeneous Subsets of Science Attitude

Age Group	Mean	N	Std. Deviation
9, 10 and 11 year old students	4.84	551	1.07
12, 13 and 14 year old students	3.73	532	1.35
Total	4.29	1083	1.33

size for this change is large, representing a major drop in attitudes toward science over the course of one year. This descent occurs for both males and females at an equal rate. The students do not recover their previously higher levels of science attitude in the later middle school years.

The students do not recover their previously higher levels of science attitude in the later middle school years.

Limitations of this study are evident. New Mexico has a different demographic structure than much of the rest of the United States. The area including the students studied includes a larger proportion of persons of Hispanic or Latino origin (43 %) than the general population of the US (14%). Native Americans also comprise a larger percentage (10%) than the general US population (1%) and African Americans a smaller percentage (2% vs 13%). The number of persons living below the federal poverty line in the area is 17%, compared to 13% in US overall ("U.S. Census Bureau Data," 2007). As cultural factors have been shown to impact science attitudes (Ato & Wilkinson, 1983; Fisher & Waldrip, 1999; Lee & Burkam, 1996; Osborne, Simon, & Collins, 2003; Reynolds & Walberg, 1992) the relationships found in this study may not be applicable to a more general population.

Maturation of the subjects may also have played a role in the student's changes in attitude. Changes in development may impact attitudes as well as exposure to science over time.

Further research on the relationship of the move from elementary school to middle school to changes in science attitudes is merited. It would be helpful to determine if maturation or cultural influences are factors in this change. Examination of students from other cultures as well as consideration of students who do not change schools from elementary to middle school would yield further information about this relationship. Clearly for this group of students a change in their attitudes is concurrent with the move to middle school. Whether this change is more generally applicable requires further investigation.

Reference

Ato, T., & Wilkinson, W. J. (1983). Factors related to secondary school students' attitudes to science in Benue State of Nigeria. *Research in Science & Technological Education, 1*(2), 209-220.

Boone, W. J. (1997). Science attitudes of selected middle school students in China: A preliminary investigation of. *School Science & Mathematics, 97*(2), 96.

Catsambis, S. (1995). Gender, race, ethnicity, and science education in the middle grades. *Journal of Research in Science Teaching, 32*(3), 243-257.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* Hillsdale, NJ: Lawrence Earlbaum Associates.

Fisher, D. L., & Waldrip, B. G. (1999). Cultural factors of science classroom learning environments, teacher-student interactions and student outcomes. *Research in Science & Technological Education, 17*(1), 83-96.

Fleming, M. L., & Malone, M. R. (1983). The relationship of student characteristics and student performance in science as viewed by meta-analysis research. *Journal of Research in Science Teaching, 20*(5), 481-495.

Games, P. A., & Howell, J. F. (1976). Pairwise multiple comparison procedures with unequal n's and/or variances: a Monte Carlo study. *Journal of Educational Statistics, 1*, 113-125.

Houtz, L. E. (1995). Instructional strategy change and the attitude and achievement of seventh- and eighth-grade science students. *Journal of Research in Science Teaching, 32*(6), 629-648.

Lee, V. E., & Burkam, D. T. (1996). Gender differences in middle grade science achievement: Subject domain, ability level, and course emphasis. *Science Education, 80*(6), 613-650.

Litwin, M. (1995). *How to measure survey reliability and validity: The survey kit.* Thousand Oaks, CA: Sage.

Mattern, N., & Schau, C. (2002). Gender differences in science attitude-achievement relationships over time among White middle-class students. *Journal of Research in Science Teaching, 39*(4), 324-340.

Nunnaly, J. C. (1978). *Psychometric Theory*. New York: McGraw-Hill.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: a review of the literature and its implications. *International Journal of Science Education, 25*(9), 1049.

Papanastasiou, E. C., & Zembylas, M. (2002). The Effect of Attitudes on Science Achievement: A Study Conducted Among High School Pupils in Cyprus. *International Review of Education, 48*(6), 469-484.

Reynolds, A. J., & Walberg, H. J. (1992). A structural model of science achievement and attitude: An extension to high school. *Journal of Educational Psychology, 84*(3), 371-382.

Simpson, R. D., & Oliver, J. S. (1990). A summary of major influences on attitude toward and achievement in science among adolescent students. *Science Education, 74*(1), 1-18.

Stake, J. E., & Mares, K. R. (2001). Science enrichment programs for gifted high school girls and boys: Predictors of program impact on science confidence and motivation. *Journal of Research in Science Teaching, 38*(10), 1065-1088.

Steinkamp, M., & Maehr, M. (1983). Affect, Ability, and Science Achievement: A Quantitative Synthesis of Correlational Research *Review of Educational Research, 53*(3), 369-396.

U.S. Census Bureau Data. (2007). Retrieved July 28, 2007, from http://quickfacts.census.gov/qfd/states/35000.html

Weinburgh, M. (1995). Gender differences in student attitudes toward science: A meta-analysis of the literature from 1970 to 1991. *Journal of Research in Science Teaching, 32*(4), 387-398.

Willson, V. L. (1983). A meta-analysis of the relationship between science achievement and science attitude: Kindergarten through college. *Journal of Research in Science Teaching, 20*(9), 839-850.

Young, D. J., Reynolds, A. J., & Walberg, H. J. (1996). Science achievement and educational productivity: A hierarchical linear model. *Journal of Educational Research, 89*(5), 272-278.

Carmen Sorge is an educational and statistical consultant with Leiden LLC. Correspondence concerning this article may be sent to csorge@nmia.com or sorgec@gmail.com

Frackson Mumba, Vivien Mweene Chabalengula, Cynthia J. Moore, William J. F. Hunter

Mathematics and Science Teaching Fellows' Instructional Planning for K-12 Classrooms

This study explored a cohort of the National Science Foundation (NSF) GK-12 Teaching Fellows' instructional planning for K-12 classrooms.

Introduction

Instructional planning has increasingly been recognized as an essential element for good teaching at different levels of education (Calderhead, 1984; Hewson & Hewson, 1987; Lubinski, 1993; Shavelson, & Stern, 1981; Tobin, Tippins, & Gallard, 1995). Instructional planning constitutes many things which teachers consider and do when they are planning for teaching (Sanchez & Valcarcel, 1999). It provides a roadmap that assists the teachers to focus on achievable objectives and learners' needs (Clark & Yinger, 1987), and to create a flow of events with starting, middle, and ending points in a lesson. As such, its primary purpose is to provide students with meaningful learning experiences (Bruce, Bruce, Conrad, & Huang, 1997).

Several studies have examined and reported teachers' instructional planning practices. For example, Aikenhead (1984) conducted a case study in which he explored the decision-making that occurred when teachers planned for teaching. Aikenhead found that teachers made decisions within a framework that holistically integrated the subject matter and practical classroom knowledge. According to Aikenhead, practical classroom knowledge embraces the basic beliefs of a teacher and the socialization of the learners. Duschl and Wright (1989) investigated the manner and degree to which teachers considered the nature of the subject matter in their decision-making as they planned their lessons. They found that the teachers' decisions on content selection, implementation, and development of instructional tasks were dominated by consideration of three factors: student development, curriculum guide objectives, and pressure of accountability from school administrators. Aguirre, Haggerty, and Linder (1990) also found that teachers' instructional practices are influenced by factors such as previous schooling experience, teaching experience, context, and role models. Other studies report that when teachers are planning for teaching, they create, arrange, organize, and design events that should occur in the classroom (Freiberg & Driscoll, 2000), and consider the nature of instructional approaches and activities to be used in a lesson (Lederman, Gess-Newsome, & Latz,1994; Lubinski, 1993). Mellado (1998) also states that teachers use their conceptions about teaching as frameworks for adopting teaching strategies, and these in turn influence the quality of instruction and outcome of student learning. Bellon, Bellon and Blank (1992) found that instructional planning occurs in three phases: Pre-active, Active (Interactive), and Post-active. In the Pre-active phase, teachers make decisions about instructional goals, objectives, teaching strategies, grouping and seating arrangements, and use of teaching materials. The Active phase involves the interactions between the teacher and students during the actual teaching. The Post-active phase involves evaluating the instructional episode through self-reflection practices and examination of student data sources (e.g. tests, papers, and projects).

There is also paucity of research that compares trained teachers' instructional planning practices with those of scientists and mathematicians in training who are involved in mathematics and science teaching in K-12 schools.

Although several studies have investigated instructional planning processes, studies of this nature have only focused on pre-service and in-service teachers. However, it is also of value to explore instructional planning practices of university graduate students in traditional mathematics and science degree programs, who are providing content knowledge and instructional support to K-12 teachers in schools through university-school partnership programs. The traditional mathematics and science degree programs are programs that prepare mathematicians and scientists, respectively. The graduate students in these programs have content backgrounds and career goals that differ greatly from K-12 mathematics and science pre-service and in-service teachers.

There is also paucity of research that compares trained teachers' instructional planning practices with those of scientists and mathematicians in training who are involved in mathematics and science teaching in K-12 schools. Yet, such a knowledge base could be essential in designing outreach programs that would be successful in helping graduate students or scientists and mathematicians acquire desirable instructional planning practices. The assumption is that these graduate students or scientists and mathematicians would translate the desirable instructional planning practices into effective instructional practices in K-12 classrooms, and this in turn would improve the quality of mathematics and science learning among K-12 students.

Our study explored university mathematics and science graduate students' instructional planning practices for K-12 classrooms. All the participants were serving in a National Science Foundation (NSF)-funded GK-12 project. The GK-12 program is a mathematics and science education outreach program that uses the content knowledge and skills of graduate students and faculty at US universities to improve teaching and learning in schools. The acronym GK-12 is used to mean that the Fellows provide content knowledge and instructional support in classrooms from kindergarten to twelfth grade.

The successful instructional planning demonstrated by the Fellows in this study does not refute the value of undergraduate pre-service teacher education, but does provide supporting data for other models of teacher education.

The graduate students who work in the NSF GK-12 program are called Teaching Fellows. They are offered Fellowships as compensation for their work in schools. The Fellows, however, are not training to become full-time teachers, nor do they become certified to teach as a result of their involvement in this program. The Fellows are training to be scientists and mathematicians. Therefore, in addition to their involvement in this program, they are engaged in authentic scientific and mathematical research for their graduate degree programs and professional development. At the request of partner K-12 teachers, the Fellows plan and teach hands-on lessons in participating schools. The Fellows also help partner teachers infuse new content into the curriculum, act as role models for students, and learn about how K-12 schools and teachers operate.

Our study is driven by three questions:

1. How do the Fellows plan for K-12 classroom teaching?
2. What do the Fellows think about their instructional planning?
3. How do their instructional planning practices relate to instructional planning best practices as revealed in the current literature?

Context of the study

This study was conducted as part of a GK-12 project at a medium-sized university (21,000 students) in the Midwest of the USA. The GK-12 project is an eight-year-long university-school partnership project involving the departments of Biological Sciences, Chemistry, and Mathematics, and more than ten school districts within a one hour drive from the university. The project started in 2001 and this study was conducted in the fourth year of the eight-year plan. The project has four interrelated goals: to use the science content knowledge and skills of university graduate students and faculty to increase scientific literacy among high school students; to enhance K-12 teachers' science content and pedagogical knowledge; to enhance teachers' knowledge and skills for conducting action research; and to enhance the existing partnerships and create new ones among the university and local schools. Within the second goal, the project regards good instructional planning as an essential element for effective teaching practices by the Fellows in schools. The schools and districts within which the project occurs are in small towns of 2000-5000 people, and in small cities of 50,000 people, with student populations of 300-500 and 1700-2500 respectively. At the beginning of a school year the Fellows are matched to schools and teachers by the project directors. Later in the year, as the Fellows establish stronger working relationships with teachers in participating schools, additional self-selected matching occurs. In some schools, the Fellows work with a single teacher. In other cases, a pair of Fellows works with a single teacher or with a pair of teachers. The Fellows have been told that their roles are to plan collaboratively among themselves and with the teachers, and to deliver inquiry-based mathematics and science lessons to classrooms.

Methodology

Participants

We studied a cohort of fifteen Fellows who were serving in the project between Fall 2004 and Spring 2005 semesters. Subject areas and degree programs for the Fellows are shown in Table 1. None of the Fellows had formal teacher training and teaching experience at the K-12 level before joining the project. However, the majority of them had one year of teaching experience in undergraduate university courses through teaching assistantships.

Data collection

Data were collected through semi-structured interviews, lesson plans, reflective journals and minutes of project meetings. Semi-structured interviews were administered to Fellows on separate days. The interview duration ranged from 30 to 45 minutes. Sample questions from the interviews were: How do you plan for your teaching in schools? What factors do you consider when planning lessons or units? What do you think about your planning? Depending on the responses provided by the Fellows, follow-up questions were asked to probe further about emerging issues. The interviews were conducted and transcribed by the first two authors. The other sources of data were the lesson plans the Fellows prepared and taught in schools during the data collection period. Ten lesson plans (four in Chemistry, three in Biology and three in Mathematics) were randomly selected from a group of thirty multi-day lessons. The Fellows also kept reflective journals in which they documented their project activities, including how they planned for teaching in schools. Fifteen journals, one from each participant, were examined. Three of the authors attended and took the minutes of the bi-weekly project meetings that focused on how the Fellows planned and taught lessons in schools.

Data analysis

Data were analyzed using the procedure proposed by Strauss and Corbin (1998). The procedure involves first reading the text line by line, coding the text to identify recurring themes and descriptors, and then establishing regularities which override individual differences, thus defining the representative profile of the group studied. An attempt was made to ensure objectivity during data analysis by continually revisiting the data to verify the elements of emerging themes. The first two authors conducted the analysis; themes that emerged from the analysis are presented in the results section below.

Results

Six recurring themes emerged from the analysis: Planning procedure; Antecedent conditions; Scope and depth of content; Content selection and sequencing; Resources; and Assessment of instructional planning.

Planning procedure

In most cases, the Fellows' lesson and unit planning involved the following procedures: consulting partner teachers, examining existing curriculum used by the teachers, identifying topics, formulating goals and objectives, identifying appropriate State Learning Standards, developing activities and assessment tools, and gathering necessary materials for activities. However, their planning procedure was not linear but rather recursive, as illustrated by the interview excerpt below:

> I start by identifying a topic, formulate objectives and align them with Learning Standards … but this is not the case all the time … . because I also moved from one section of the lesson to another just to make sure that all the relevant aspects have been considered (John).

According to John, lesson planning is a continuous process where one has to think about many factors as

Table 1: Profiles of the Participants

Subject area	Number of Fellows	Gender		Degree program	
		Female	Male	MS	PhD
Chemistry	5	2	3	5	0
Biology	6	4	2	3	3
Mathematics	4	3	1	1	3

he or she plans for teaching. Indeed, most lessons and units we examined showed that the Fellows revisited them several times.

Antecedent factors

The Fellows considered several antecedent factors such as age of students, their abilities, grade level, class size, type of students (regular or advanced placement classes), location of the schools, and facilities available in schools. The Fellows also made references to these factors during project meetings. For example, in one of the meetings, Sara stated some factors she considered when planning for her teaching.

> I consider their ages, prior knowledge and ability to handle the content and perform activities. I don't make assumptions for different classes. It has been helpful to consider a lot of these things when planning for teaching in schools (Sara).

Sara implies that when teachers plan for teaching, they should consider several instructional factors if they are to be successful in their lessons. The extent to which these antecedent conditions were considered varied from one lesson or unit to another and from Fellow to Fellow. When asked how they knew or determined the antecedent factors to consider in their planning, the Fellows said that they relied on classroom observations, informal conversations with students and partner teachers, and students' performance records. However, a small number of reflective journals and lesson plans showed that some Fellows prepared and administered tests that were aimed at eliciting students' prior knowledge.

Scope and depth of content

To a large extent, most antecedent conditions outlined above influenced the scope and depth of content in the lessons and units. The Fellows also said they used the goals and objectives of the lessons to determine the scope and depth of the content to be addressed in a particular lesson or unit. As indicated in the planning procedure, the Fellows consulted the partner teachers before they started developing full lessons. For example, in an interview excerpt below, Peter, biology Fellow, talks about the importance of the discussions with the partner teachers.

> I discuss with my partner teachers before I decide on the nature of the content to be addressed in a lesson or unit … because they have curriculum from the school districts to complete in a school year. I make sure that the content is not off the curriculum. However, I include content and activities that are challenging to students (Peter).

Peter consulted partner teachers and developed lessons and units using curriculum guides given to teachers by their school districts. This was the expected practice, because the Fellows were mainly there to provide content knowledge and instructional support to the teachers. It was during the discussions with the teachers that the scope and depth of content to be covered in a lesson or unit were agreed upon. The discussions also provided the Fellows with opportunities to suggest science and mathematics concepts to be included in the lesson or unit that were not part of the existing curriculum.

Content selection and sequencing

When asked about how they selected and sequenced the content in the lessons and units, most Fellows said they used a systematic approach for introducing the concepts in their lessons. For example, in a reflective statement below, Jaime, a mathematics teaching Fellow, talks about how she sequenced content in her lessons and units.

> I start with simple or obvious concepts and then build on them as the lesson or unit progresses. So far, this approach has worked very well (Jaime).

Jaime implies that it is better to start with basic concepts and build on them as the lesson progresses. A majority of the Fellows said that this approach helped students participate, become interested, and understand the concepts in the lesson. Indeed, the lesson plans and units we analyzed showed that most Fellows presented the concepts in this manner.

Resources

The Fellows used several resources during their instructional planning, such as existing lessons, technology and textbooks they found in schools. They also used any resources that were available in the library and departmental resource centers at the university, such as college textbooks, science and mathematics journals, curriculum units, biological species and teaching models. According to the Fellows, resources at the university provided content, activities and other relevant information that were not available in the K-12 schools. In an interview excerpt below, Tina, a chemistry Fellow, names the resources she used.

> I consult several sources when I plan for my teaching … such as science education websites,

> journals, and textbooks. I also use the lessons from other teachers … modify them or make them more challenging and interesting by including more activities (Tina).

Tina used both electronic and paper-based resources when planning her lessons and units. She also used resources and lessons from teachers, and made them more activity-based and interesting to students. The Fellows also said that they found technology a useful tool in their planning for teaching in schools. Several science lessons that were analyzed had laboratory activities that involved the use of computers for simulations and data gathering and processing. A majority of mathematics lessons also involved the use of calculators and probes for data collection, numerical manipulation and graphing.

Assessment of instructional planning

Most Fellows said that they assessed their instructional planning practices through students' and teachers' feedback and their own self-reflection practices. Generally, the Fellows said that they were satisfied with their instructional planning practices. They attributed their satisfaction to successful instructional experiences in schools, collaboration with partner teachers, information searching skills and opportunity to plan and teach more lessons. In the excerpts below Andrew, a biology Fellow and Charles, a mathematics Fellow, explain why they were satisfied with their instructional planning and how they assessed themselves.

> I am pretty satisfied with my planning strategy … because my lessons go on well in schools…and students give me positive feedback. My partner teachers like the way the lessons are planned and taught. The more I teach the better I teach. (Andrew).

> I ask myself questions during planning and after teaching. Are the activities challenging? How helpful is this activity to students? How did the lesson go? Did I achieve the objectives? After answering them I have an idea about my planning and teaching (Charles).

Andrew, Charles and the rest of the Fellows made value judgments about their instructional planning based on feedback from students and teachers and their self-reflections on the lessons and teaching.

Conclusions and Discussion

These findings demonstrate that most of the factors the Fellows considered during instructional planning were similar to those reported in previous studies that involved pre-service and in-service teachers (Aikenhead, 1984; Bellon et al., 1992; Duschl & Wright, 1989; Lederman, Gess-Newsome, & Latz, 1994; Lubinski, 1993; Sanchez & Valcarcel, 1999; Shavelson & Stern, 1981). In particular, Fellows were concerned with antecedent conditions, and consideration of scope and depth, and identification of content; and their planning procedures were similar to those of pre-service teachers. Their collaborative planning practices were also consistent with those that can lead to effective teaching and learning in schools when non-teachers act as experts and bring resources to the classroom (Bruce et al., 1997). Clearly, the Fellows were also generally satisfied with their instructional planning – perhaps even as a result of the layered considerations they brought to instructional planning.

Although the cohort of the Fellows in this study did not receive formal training in instructional planning, they learned how to plan for instruction through practice and interactions with partner teachers. As dedicated professional teacher educators, the authors have a vested interest in the use of direct instruction for pre-service teachers, but for these Fellows, direct instruction is not vital. Rather, on-the-job training under the guidance of an expert can be an effective method of teaching instructional planning.

The successful instructional planning demonstrated by the Fellows in this study does not refute the value of undergraduate pre-service teacher education, but does provide supporting data for other models of teacher education. The findings of this study also reinforce our professional observations in science and mathematics teacher education that some teachers tend to learn about instructional planning more effectively outside the traditional academic courses.

Future research should focus on whether there is a relationship between the Fellows' instructional planning practices and their knowledge base of science and mathematics teaching. In addition, most of the instructional planning themes in this study were found to be in Pre-active phase of the model developed by Bellon et al. (1992). This result should not be too surprising considering that the purpose of the study was to explore the Fellows' instructional planning practices - most of which occurred in the Pre-active phase. Therefore, a study should be conducted to examine the Fellows' considerations in Pre-active, Active, and Post-active instructional phases. This will lead to a better understanding of what the Fellows do in each of the three phases of instructional planning.

Finally, one limitation of this study was that only the GK-12 Fellows' instructional planning practices were explored. There were no data collected from the partner teachers who worked with the Fellows in schools. We suggest that future studies should attempt to collect data from the Fellows and partner teachers to provide contrast or additional support to conclusions reported in this study.

Implications

Our data support the hypothesis that direct mentoring by practicing teachers can be an effective means of modeling and communicating appropriate instructional planning skills. This approach may be useful in alternative teacher training programs, where students (in similar fashion to GK-12 Fellows) may receive minimal formal instruction in lesson planning.

Since the Fellows in this type of project are themselves graduate students in science and mathematics, these findings suggest that GK-12 Fellows could be useful in the training of graduate teaching assistants. The Fellows learned a great deal about how to plan for teaching through practice and through interactions with teachers and K-12 students. Therefore, academic departments at universities that have GK-12 projects can use the Fellows' expertise in teaching assistantship programs through presentations on instructional planning and/or mentoring teaching assistants.

Acknowledgement

This study was supported by National Science Foundation Grant # DGE-0338188. The opinions expressed in this article are solely those of the authors and are not those of the Foundation.

References

Aikenhead, G. S. (1984). Teacher decision making: The case of Prairie High. *Journal of Research in Science Teaching, 21*; 167-186.

Aquirre, J. M., Haggerty, S.M., & Linder C., J. (1990). Student-teachers' conceptions of science, teaching and learning: a case study in pre-service science education. *International Journal of Science Education, 12*, 381-390.

Bellon, J.J., Bellon, E.C., & Blank, M. A. (1992). *Teaching from a research knowledge base: A development and renewal process*. New York: Macmillan.

Bruce, B. C., Bruce, S. P., Conrad, R. L., & Huang, H (1997). University science Fellows as curriculum planners, teachers, and role models in elementary school classrooms. *Journal of Research in Science Teaching, 34* (1); 69-88.

Clark, C. M., & Yinger, R. J. (1987). Teacher planning. In J. Calderhead (Ed.), *Exploring teachers' thin*king (pp.84-103). London: Cassell Educational.

Duschl, R. A., & Wright, E. (1989). A case study of high school teachers' decision making models for planning and teaching science. *Journal of Research in Science Teaching, 26*; 467-501.

Freiberg, H. J., & Driscoll, A. (2000). *Universal teaching strategies*. Boston: Allyn and Bacon.

Lederman, N. G., Gess-Newsome, J. & Latz, M. S. (1994). The nature and development of pre-service science teachers' conceptions of subject matter and pedagogy. *Journal of Research in Science Teaching, 31*, 129-146.

Lubinski, C. A. (1993). More effective teaching in mathematics. *School Science and Mathematics*, *93*(4), 198-202.

Mellado, V. (1998). The classroom practice of pre-service teachers and their conceptions of teaching and learning science. *Science Education, 82*, 197-214.

Sanchez, G, & Valcarcel, M.V (1999). Science teachers' views and practices in planning for teaching. *Journal of Research in Science Teaching, 36*(4), 493-513.

Shavelson, R. J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions and behavior. *Review of Educational Research, 51*; 455-498.

Tobin, K. Tippins, D. & Gallard, A. (1995). Research on instructional strategies for teaching science. In D. Gabel (Ed.). *Handbook of Research on Science Teaching and Learning*. New York: MacMillan, 45-93.

Frackson Mumba is assistant professor of science education and director of teaching excellence in mathematics and science, Southern Illinois University Carbondale, IL 62901. Correspondence concerning this article may be sent to Frackson@siu.edu

---

Vivien M. Chabalengula is a lecturer of science education, Southern Illinois University Carbondale, IL 62901.

Cynthia J. Moore is associate professor of biological sciences and director of biology teacher education, Illinois State University, Normal, IL 61790.

William J.F. Hunter is associate professor of chemistry and associate director of the Center of Mathematics, Science and Technology, and director of Chemistry Teacher Education at Illinois State University, Normal, IL 61790.

Joseph A. Taylor, Pamela Van Scotter, Doug Coulson

Bridging Research on Learning and Student Achievement: The Role of Instructional Materials

The authors describe results from two field-test studies conducted by BSCS, both of which indicate a strong relationship between fidelity of curriculum implementation and student learning gains.

For decades the National Science Foundation has been funding the development of instructional materials whose design is based upon the recommendations of educational research. These recommendations include the idea that learning be sequenced and organized using an experiential learning cycle (see Atkin & Karplus, 1962; Piaget, 1975) or an instructional model such as the Biological Science Curriculum Study (BSCS) 5E Instructional Model (see Bybee, 1997). More recent research studies such as those synthesized in *How People Learn* (NRC, 2001) suggest that instruction address students' prior knowledge, help students connect new knowledge to a rich framework of big ideas, and support students in monitoring and taking control of their own learning. Work by Pellegrino, Chudowsky, and Glaser (2001) make complementary recommendations such as the following:

- Instruction should be organized around meaningful problems and goals.
- Instruction must provide scaffolds for solving meaningful problems and supporting learning for understanding.
- Instruction must provide opportunities for practice with feedback, revision, and reflection.
- The social arrangements of instruction must promote collaboration and distributed expertise as well as independent learning.

Over time, the work of curriculum developers in response to these recommendations has resulted in an extensive portfolio of research-based instructional materials that span the sciences disciplines. However, few researchers have systematically explored how these materials influence student learning and the role that implementation fidelity plays in the materials' ultimate impact. The purpose of this article is to address this issue by exploring data collected in two BSCS research studies. In the following sections we briefly describe these studies, their respective findings, and make recommendations for teacher professional development.

BSCS Field-Test Research: A Brief Overview

The BSCS, as a routine part of its curriculum field-testing process, studies the impact of the materials on student learning. In 1995, BSCS conducted case studies of four teachers who were field-testing a new high school science program. The learning experiences in this curriculum program are structured and sequenced using the BSCS 5E Instructional Model (Engage, Explore, Explain, Elaborate, and Evaluate) which is based upon the research-based learning cycles of Atkin and Karplus (1962) and Piaget (1975).

The four case studies uncovered distinct differences in the pre/post learning gains of students whose teachers implemented the program as designed as opposed to those of students whose teachers implemented the program with considerably less fidelity. Student learning was measured using a 20-item subset of questions from the National Science Teachers Association (NSTA)/National Association of Biology Teachers (NABT) biology exam, administered at the beginning and end of the school year. Fidelity was measured through classroom observations conducted by BSCS curriculum development staff. These findings are illustrated in Figure 1 and Table 1.

The case studies suggested a relationship between fidelity and student learning, and BSCS sought to explore

Figure 1: Pre/Post Results for NABT/NSTA Biology Exam

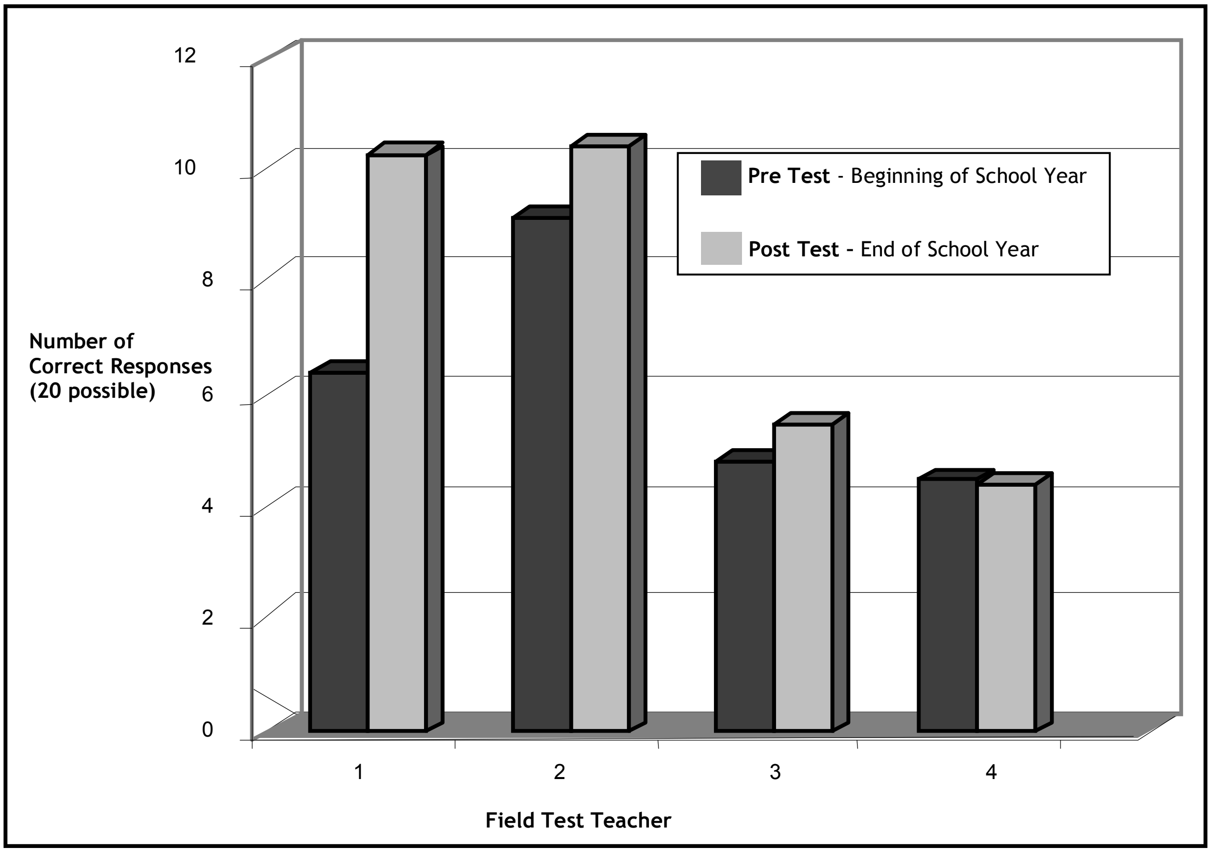


Table 1: Student Learning Gains by Teacher

Teacher	Pre-test Avg.	Post-test Avg.	Avg. Gain
1	6.4	10.3	3.9
2	9.2	10.4	1.2
3	4.8	5.5	0.7
4	4.5	4.4	0

Teacher 1 - Field tested instructional materials for two years with a high level of fidelity
Teacher 2 - Field tested instructional materials for two years with a medium level of fidelity
Teacher 3 - Field tested instructional materials for one year with a medium level of fidelity
Teacher 4 - Field tested instructional materials for one year with a low level of fidelity

this relationship with larger numbers of teachers and students in future field-test studies. This opportunity came in 2002 as BSCS began field-testing a newly-developed high school science program also structured using the BSCS 5E Instructional Model.

The 2002 Study: An Overview

One of the questions that the 2002 study sought to answer was whether the magnitude of student learning gains were in any way related to how closely teachers adhered to the intended use of the instructional materials. Specifically, the study asked: *how do the learning gains of students whose teachers are successful in implementing the program as designed (with fidelity) compare with the learning gains of students whose teachers are less successful in implementing the program as designed?*

In this study, the learning gains of 326 ninth-grade students (across 15 teachers) were determined by administering an identical chapter test before and after instruction. Implementation fidelity was measured by a research team composed of external evaluation and curriculum development staff using an observation protocol adapted from the *Horizon, Inc. Classroom Observational Protocol* (see HRI, 2000). The implementation section of this protocol included 11 individual rating scales that addressed the use of specific teaching strategies consistent with the BSCS 5E Instructional Model. These rating scales were in Likert format with scores ranging from 1 to 5; one meaning "not at all," 3 meaning "to some extent," and 5 meaning "to a great extent." For each scale, the team recorded specific notes as evidence and justification for the ratings. Specifically, the individual rating scales were intended to quantify the extent to which teachers encouraged students to: engage in metacognitive activity; communicate their understanding of concepts, and apply their understanding to new situations. These rating scales were also used to quantify the extent to which teachers used embedded assessment and student self-assessment, as well as provided students with appropriate feedback to help them construct their understanding of targeted concepts.

When student learning gains were examined in light of implementation fidelity, a strong relationship emerged.

Using these 11 ratings of teachers' use of 5E-based strategies and learning sequences, BSCS then holistically classified each teacher's fidelity of use as either "low," "medium," or "high" (see Table 2). These holistic ratings alone were used to divide teachers into low, medium, or high levels of fidelity because trichotomizing teachers on the holistic rating resulted in a clearer division than averaging the other 11 individual rating scales. The averaging process tended to cancel out the variation, making it problematic to create fidelity categories.

Findings from the 2002 Study

When student learning gains were examined in light of implementation fidelity, a strong relationship emerged. An analysis of covariance (ANCOVA) was conducted across the three fidelity groups (i.e., low, medium, and high) on post test using the pre test as a covariate. As expected, the covariate (pre-test) was significantly related to the dependent variable (post-test). There was a main effect associated with the implementation fidelity grouping ($F = 7.51$; $df = 2, 322$; $p < .01$). This analysis, summarized in Table 3, suggests that we reject the null hypothesis that the post-test scores are equal across fidelity groups.

Post hoc paired comparisons using Fisher's Least Significant Difference (LSD) method were conducted in order to identify which of the three implementation fidelity groups were statistically different. These three comparisons showed that the post-test scores (when adjusted for pre-tests) of the low fidelity group were on average statistically lower than both the medium and high level fidelity groups and that the adjusted post-test scores for the medium and high fidelity groups were not statistically different. Table 4 and Figure 2 below summarize these findings.

Discussion

The data presented in this study suggest that the students whose teachers used the instructional materials with medium or high levels of fidelity scored statistically higher on the post-test achievement measure than students of low fidelity teachers. Using the operational definitions of medium

Table 2: Fidelity Levels as a Continuum of 5E-Based Teaching

	Holistic Rating of Fidelity Level		
	Low	**Medium**	**High**
Observed use of strategies and learning sequences that are consistent with the 5Es	**None**	**Basic**	**Extensive**
Number of Teachers Rated at each Fidelity Level	**4**	**7**	**4**

Table 3: ANCOVA Output

Source	DF	Mean Square	F-Ratio	Probability
Pre Test	1	2.8368	48.5	.0000
Implementation Fidelity Group	2	.43938	7.51	.0006
Error	322	.05849		
Total Adjusted	325			

Table 4: Adjusted Means Table

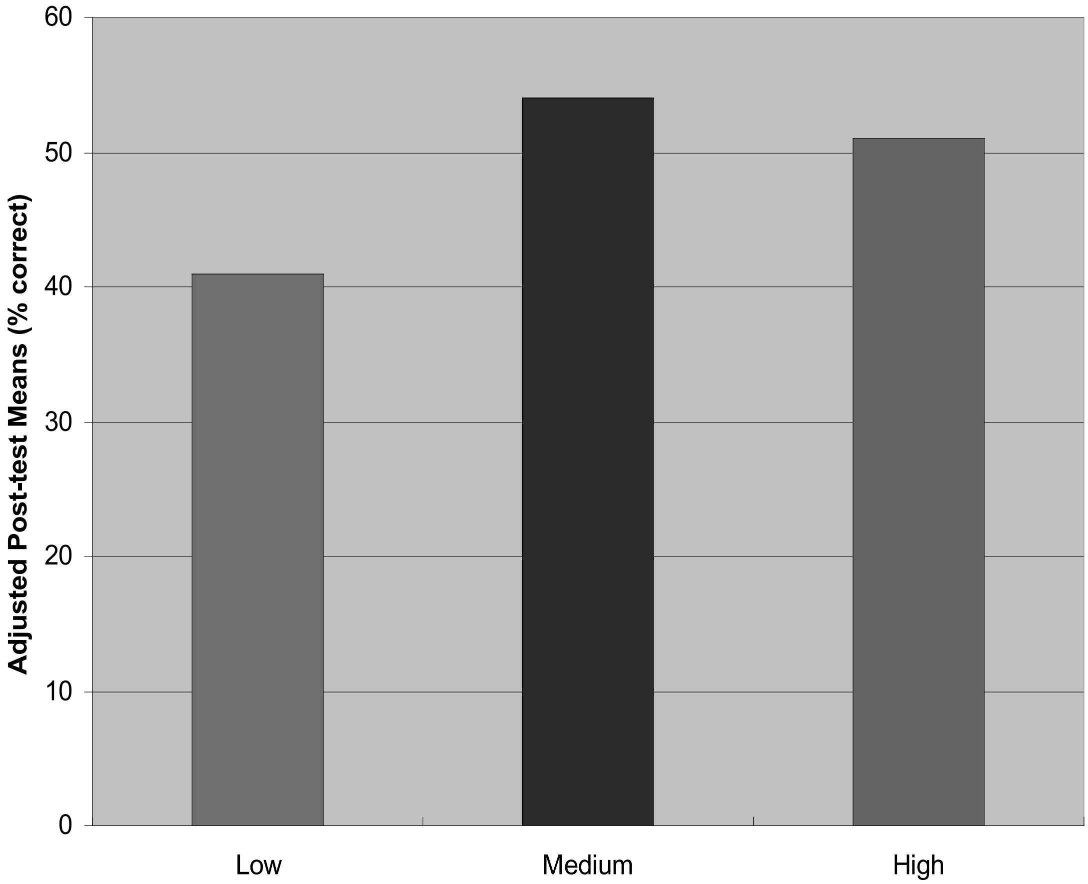
Teacher Group	#of Students	Adjusted Mean Post-test Score (% correct)	Standard Deviation	Fisher LSD (adjusted mean is significantly different from)
Low Level Fidelity	70	41	2.9	Medium and High Fidelity Group
Medium Level Fidelity	168	54	2.6	Low Fidelity Group
High Level Fidelity	88	51	2.0	Low Fidelity Group

and high levels of fidelity, we can then state that there is a statistical link between superior student achievement and ***basic*** or ***extensive*** use of strategies and learning sequences consistent with the 5Es. The observation that marked differences in achievement begin at even basic use of the 5Es makes a powerful statement about the effectiveness of the instructional model. Complimentary findings can be found in comparative studies--conducted across multiple science disciplines and grade levels--suggesting that research-based instructional models likely promote greater gains in student achievement than more didactic teaching approaches (e.g., Ates, 2005; Ebrahim, 2004; Lord, 1997).

The lower student achievement observed in low fidelity classrooms is not surprising. Too often in science education we see well-designed instructional materials, even those designed to organize everyday instruction, collecting dust on shelves or being pulled off shelves and used haphazardly as a mere resource or supplement. This often results in a patchwork approach to curriculum implementation, which leads to lack of coherence in the learning sequence for students (Rutherford, 2000; Taylor, et al., 2005). This is not to say that teacher-designed materials tend to be of poor quality. However, most classroom teachers' ability to field-test their materials with large numbers of diverse students or to have them reviewed by content experts is clearly limited. A haphazard approach to curriculum implementation does not take advantage of the thoughtful work of the science education research and curriculum development communities. The prevalent notion of teacher as curriculum developer or curriculum "hunter and gatherer" must be challenged at both the K-12 and higher education levels. The data in this study directly confront this common notion and suggest an approach to curriculum implementation where fidelity is valued.

It is interesting to speculate about why extensive use of the 5Es did not yield student achievement that was significantly different from basic use of the instructional model. It is possible that the instructional materials, which make the 5Es explicit for both teachers and students, are ***fidelity-resilient*** (within limits). That is, even when a teacher implements the material with only medium fidelity, the inherent design of the instructional materials likely contributes to enhanced un-

Figure 2: Adjusted Post-test Means by Level of Fidelity



derstanding, and thus achievement, for students.

Other hypotheses around this finding center on how the data were analyzed. For example, some of the teachers who were rated holistically at medium levels of fidelity were rated so because their use of 5E-based strategies and learning sequences was not consistent across a continuum from non-use to extensive use. It is possible that a large percentage of these teachers made extensive use of the very strategies that were most effective for students. To address this issue in future research we would like to conduct an analysis of covariance across selected individual fidelity rating scores from the observation protocol, again using pre-test scores as covariates. This analysis would help us make more direct connections between student achievement and the use of specific instructional strategies.

We also hypothesize that in some cases, teachers made departures from the design of the instructional materials that optimized student learning and were in the ***spirit*** of the instructional model but were noted by observers as showing less fidelity—resulting in lower fidelity ratings. This conclusion would suggest that it is indeed possible for teachers with a nuanced understanding of the 5Es to enhance the impact of 5E-based instructional materials with well-informed pedagogical decisions.

The question then becomes: how do we help teachers develop sophisticated understandings of instructional materials and the instructional models that drive their design? Suggested approaches to professional development are articulated in other studies that establish a link between fidelity of implementation and student learning (e.g., Cohen & Hill, 2002; Darling-Hammond, 1997). In many of these studies, high levels of fidelity are attributed in part to on-going, comprehensive professional development that is focused almost solely on helping teachers understand the design of the instructional materials. The suggestion is quite simple: if you want higher fidelity implementation, spend some professional development time and resources on helping teachers understand the instructional materials. Further, since the design of research-based instructional materials can look quite different from that of traditional instructional materials and embody instructional models of which many teachers are unfamiliar, the importance of professional development is emphasized.

A haphazard approach to curriculum implementation does not take advantage of the thoughtful work of the science education research and curriculum development communities.

In the 2002 study, the translation of this simple suggestion was to provide teachers with a professional development program that included workshops to help them understand the goals and purpose of the instructional model (5Es) that organizes and structures the learning for students. The specifics of this approach were influenced by scholars such as Loucks-Horsley, et al. (2003) who suggested that professional development mirror the instructional methods to be used with students. Therefore, teachers in this study were engaged as learners of science in investigations where the facilitator modeled exemplary use of the 5E model. To help teachers apply their science learning experience to learning about science teaching, the investigations were discussed afterward in terms of what the learner and the facilitator were doing in each stage of the 5Es.

The professional development program also engaged teachers in other foundational work to help them understand the instructional materials. For example, BSCS conducted workshops to help teachers better understand inquiry and effective ways to create a climate of inquiry in their classrooms, as well as sessions on *Understanding by Design* (Wiggins and McTighe, 2005), which provided key principles for the design of the program. The professional development program also included focused sessions on the science content since it was observed in the past that low levels of fidelity can often result from unfamiliarity or discomfort with the science content.

In summary, the data in this study suggest that research-based instructional models are most effective when they are taught with at least a basic level of fidelity. We also hypothesize that well informed modification of research-based instructional materials could optimize their impact on student achievement. However, regardless of whether the goal is to have teachers use research-based instructional materials with optimal fidelity or to modify the materials in appropriate ways, it is critical that professional development focus on helping teachers develop a vision of implementation that is consistent with the designer's intent.

References

Ates, S. (2005). The effectiveness of the learning-cycle method on teaching DC circuits to prospective female and male science teachers. *Research in Science and Technological Education* 23(2): 213-227.

Atkin, J. M., & Karplus, R (1962). Discovery or Invention, *The Science Teacher, 29 (2),* 121-143.

Bybee, R. W. (1997). *Achieving Scientific Literacy: From purposes to practices*. Portsmouth, NH: Heinemann.

Cohen, D. K., & Hill, H. C. (2002). *Learning policy: When state education reform works*. New Haven, CT: Yale University Press.

Darling-Hammond, L. (1997). *Doing what matters most: Investing in quality teaching*. New York: National Commission on Teaching and America's Future.

Ebrahim, A. (2004). The effects of traditional learning and a learning cycle inquiry learning strategy on students' science achievement and attitudes toward elementary science (Kuwait). *Dissertation Abstracts International* 65(4): 1232.

Horizon Research, Inc. (2000). *Horizon Research-2000-2001 Local systemic change classroom observation protocol*. (Found at www.horizon-research.com).

Lord, T. R. (1997). A comparison between traditional and constructivist teaching in college biology. *Innovative Higher Education* 21(3): 1127-1147.

Loucks-Horsley, S., Love, N., Stiles, K., Mundry, S., & Hewson, P. (2003). *Designing professional development for teachers of science and mathematics*, 2nd ed. Thousand Oaks, CA: Corwin Press.

National Research Council. (2001). *How People Learn*. Washington, DC: National Academies Press.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of assessment*. Washington, DC: National AcademyPress.

Piaget, J. (1975). *The Development of Thought*. New York: Viking Press.

Rutherford, F. J. (2000). Coherence in high school science. In *Making sense of integrated science: A guide for high schools*. Colorado Springs, CO: BSCS.

Taylor, J., Powell, J., Bess, K., & Lamb, T. (2005). Examining the professional growth of out-of-field physics teachers: Findings from a pilot study. *Journal of Physics Teacher Education Online, 2(4),* 16-22.

Wiggins, G., & McTighe, J. (2005). Understanding by design, 2nd ed. Alexandria, VA: Association for Supervision and Curriculum Development.

Joseph A. Taylor is director, BSCS Center for Research and Evaluation, 5415 Mark Dabling Blvd., Colorado Springs, CO 80918. Correspondence concerning this article may be sent to jtaylor@bscs.org

Pamela Van Scotter is director of the Center for Curriculum Development at BSCS in Colorado Springs, Colorado.

Doug Coulson is senior partner at PS International, an evaluation-consulting group in Annapolis, Maryland, and adjunct faculty in education statistics at the University of Maryland, College Park, Maryland.

Patricia Mills Shane, Brenda Shumate Wojnowski

Technology Integration Enhancing Science: Things Take Time Revisited

A revisit of a long-term professional development effort to imbed technology into the K-8 curriculum suggests that lasting changes such as retention of effective existing practices and pedagogical security result when teachers are given sufficient time and duration to carry out the innovation.

Project TIES (Technology Integration Enhancing Science), a four-year K-8 Technology Literacy Challenge project, combined technology as a tool for teaching and learning with earth and environmental science education. This project provided teachers in two North Carolina school systems with professional development as well as technology equipment and materials during the project years, 1998-2002. These resources enabled teachers to make the transition from traditional classroom methodologies to the use of technology as an imbedded and integral part of teaching and learning.

In the TIES project, teachers participated in professional development involving science content, the inquiry process, student-centered projects, and the use of technology as a tool for teaching and learning. Over the course of the project, many TIES teachers assumed leadership roles within their school systems and in state professional organizations, thereby assuring that the expertise and leadership needed to sustain the project for four years resided within the school. One of the goals of the project was to sustain the appropriate use of technologies in classrooms after the project was terminated. The authors were interested in whether the leadership had remained in place and whether these teacher leaders had been able to maintain their gains four years after funding for the project ended. The authors visited the school systems in the original project ending in 2002 again in 2006 to see if this goal of sustainability of the project had been met as well as to see what changes may have occurred in the use of technologies in classrooms since the completion of their original work.

TIES Project schools represent a diverse middle grades student population with respect to ethnicity and economic background and are located in suburban and rural communities.

As the technology era entered the classrooms of the 1990's, it became clear that significant professional development was needed to help teachers understand how to incorporate various technologies as an authentic part of teaching and learning. To help meet this need, the Technology Literacy Challenge Fund (TLCF) was established as part of the Elementary and Secondary Education Act (ESEA). The purpose of the TLCF program was to provide assistance to states and districts to support the integration of technology into school curricula with the goal of improving teaching and learning and enabling all students to become technologically literate. As a result of this ESEA legislation, Project TIES became a reality; and a four-year saga of change and innovation began. [See Shane and Wojnowski article in Science Educator, 14 (1).]

Things Take Time

Change is not easy. For long-lasting pedagogical change to occur, teachers must be afforded the opportunity to learn new teaching methodologies, incorporate those methodologies into their classroom practices, modify any practices that do not work for them, and retest the modifications. "It is clear that, for science and mathematics professional development to be effective, experiences for teachers must occur over time, provide ample time for in-depth investigations and reflection, and incorporate opportunities for continuous learning….[T]he idea of building new understandings through active engagement in a variety of experiences over time, and doing so with others in supportive learning environments, is critical for effective professional development" (Loucks-Horsley, S., Love, N., Stiles, K.E., Mundry, S., and Hewson, P.W., 2003, p. 81-82). Although the project was nearing completion as this caveat was published, Project TIES reflected this precept. TIES allowed teachers the time to assimilate new pedagogies and implement them in their classrooms.

For this particular technology-based project, it is accurate to add the admonition that "Things Take Materials." The intention was to provide sufficient resources for teachers to make the transition from traditional practice to classrooms where science and technology were imbedded and integral parts of teaching and learning. This was accomplished through a two-pronged approach. First, make the technology equipment available to teachers in sufficient quantity for easy student access within the classroom setting. Second, provide for acquisition of the concomitant abilities needed to use the technology in an appropriate and authentic manner. This approach permitted students and teachers to use technology on a regular and frequent basis and allowed for integrated, project-based instruction. The combination of new knowledge and behaviors as a result of professional development, combined with the needed equipment, helped to provide profound and lasting change.

Project Description

The overarching goal of the TIES Project was to produce a successful, innovative, and replicable model for inquiry- and project-based instruction that used technology to integrate science with other curricula. To attain this goal, teachers developed long-term inquiry-based science projects appropriate for their elementary and middle grades students. Underlying these projects, as well as other classroom instruction, was the seamless blending of technology with science content and project-based instruction. The ensuing professional development not only incorporated project-designed activities, but also a wide array of nationally recognized curriculum materials and activities. The National Science Education Standards were issued at about the same time and became an integral part of the project as well. These programmatic components were phased in over the project's first three years, with full implementation achieved in Year 4.

Project TIES had these objectives:

- provide professional development in:
  - technologies in the context of authentic projects
  - the Internet as a tool to support classroom learning
  - strategies and techniques for integrating technologyinto the curriculum
  - science content for K-8 teachers
- acquire adequate technology hardware and software for partner schools to insure student access
- provide opportunities for TIES participants to learn to utilize their school grounds to enhance their instruction in the context of the science curriculum and technology tools
- provide opportunities for TIES leaders to share their expertise with new TIES teachers, as well as other teachers in their schools
- form a collaboration of partner schools to enhance and support each other

A continuing part of the project was the attainment of assured sustainability for the model. This priority was accomplished by way of five strategies. First, TIES implemented a process of collaborative team efforts utilizing the leadership of experienced TIES teachers. Year-1 and Year-2 teachers became mentors for teachers who entered the project in Year 3. Second, experienced teachers assumed leadership roles as they participated in providing professional development sessions in Years 3 and 4. Third, the technology equipment was housed in teachers' classrooms. Fourth, teams of TIES teachers disseminated knowledge gained and lessons learned from the project as they presented TIES at science and technology conferences and at parent and faculty meetings. Finally, participating schools included TIES in their school-based budgets. This article revisits the success of these five long-term sustainability strategies four years after completion of the funding phase of the project.

Collaborations

The TIES Project was built on the strong collaborations of four schools in two school systems, the Center for Mathematics and Science Education in the University of North Carolina at Chapel Hill (CMSE), the North Carolina Department of the Environment and Natural Resources (DENR), LEARN NC (a statewide technology network), the North Carolina Department of Parks and Recreation, the Eisenhower Consortium at SERVE, and the GLOBE Program. In addition, an external evaluator was recruited to help determine the extent to which the objectives were achieved. Interestingly, when TIES was revisited after four years, the authors found that although the Eisenhower Consortium is no longer an active entity, the schools and the remaining partners continue to be engaged.

TIES Project schools represent a diverse middle grades student population with respect to ethnicity and economic background and are located in suburban and rural communities. The CMSE brought strong leadership capabilities in grant administration and professional development, as well as technical guidance in developing and implementing educational models. The DENR brought expertise in assessing and understanding the environmental resources of TIES school sites. LEARN NC, a statewide network of educators using Internet technologies, provided teaching resources, lesson plans keyed to the North Carolina Standard Course of Study, and an online outlet that allowed TIES teachers to share their expertise with other educators.

Implementation

Technology can be a powerful entity in classroom instruction when adequate resources are seamlessly incorporated into instructional approaches and strategies. One way to accomplish this is to provide teachers and students with a vehicle for instruction that brings applications to the world beyond the classroom. To implement these real-world projects successfully, teachers must develop skills in integrated instructional strategies, have exposure and experience with specific projects, and be proficient in the appropriate use of technology as a tool for instruction and learning. Administrative support and participation is crucial. Significant commitments of personnel, financial resources, and time are required for a single school to make improvements in these arenas. The need for collaboration is important so teachers, struggling for time to make improvements in their individual classrooms, do not waste time reinventing the wheel.

Technology can be a powerful entity in classroom instruction when adequate resources are seamlessly incorporated into instructional approaches and strategies.

To build and apply skills for using available infrastructure effectively, each year TIES classroom teachers, project support staff, and administrators participated in ten days of professional development, including two days at professional conferences. TIES professional development introduced authoring tools, word processing, databases, spreadsheets, and the effective use of the Internet. It also provided hands-on experiences for the understanding of science content –especially in the area of earth science, which blended well with the TIES “outdoors as a classroom” focus.

TIES teams implemented projects based on content and integrated instructional strategies developed during professional development sessions in their own classrooms. This implementation strengthened team building, leadership skills, and mentoring opportunities for TIES teachers and administrators. In TIES, the power of technology merged with a constructivist pedagogy in student-centered, project-based classrooms.

To support both curriculum and standards requirements, TIES project development used the following instructional aproaches.

Constructivist, Student-Centered Learning: Students learn best when they construct their own knowledge, based on multiple experiences with a concept or skill. Through active, hands-on experiences, they correct their misconceptions, extend what they know, and connect their knowledge to other concepts they understand. Student motivation is enhanced when students pursue answers to questions they have developed.

Collaborative Learning: Most students like to work with their peers and learn more from doing so. Working collaboratively is a required workplace skill for the Information Age. Many everyday activities are collaborative, with students working in small groups to solve a problem.

Authentic Learning: Students learn best when their learning is not artificial —when activities are authentic and connected to the world outside the classroom.

Student as Worker, Teacher as Facilitator: A teacher serves as a facilitator to student learning by arranging the environment so that students will ask important questions and discover ways to answer them.

Sustainability. There were two types of sustainability connected to this project: 1) intra-school sustainability within the school(s) after external funding was expended; and 2) inter-school sustainability attached to projects that served as models to be transferred to and used by other schools and districts. Follow-up questions concentrated on intra-school sustainability.

Intra-school sustainability requires having key elements of materials, equipment, personnel, and leadership in place in a school(s) so a project can continue after funding expires—to have a "life of its own," so to speak. Continued financial support to update equipment and replenish consumable materials is usually necessary as well.

When the authors revisited TIES, they found the leadership and personnel in place and more than willing to sustain and expand the gains made during the TIES project. The funding necessary to replace, update and/or repair equipment and to buy current more advanced technologies has been provided in one TIES system, but has been significantly more limited in the other. Disparities in levels of use of technologies in classrooms that were not evident in 2002 are very evident in 2006.

Great efforts were made with Project TIES to ensure it had the support needed to continue in current schools long after the conclusion of the grant period. Hardware and software were placed in classrooms, and professional development was provided to enable teachers to utilize the technology in an effective manner. In addition, extensive professional development was provided to enable participants to understand how to implement inquiry- and project-based instruction using technology as a tool. Returning TIES teachers emerged as leaders and provided on-going professional development to others in their schools and districts. Local school district budgets were modified to accommodate updates and repairs of project hardware and software.

When the authors revisited the TIES schools four years after the cessation of external funding, they found only three of the original Year 1 teachers still in the classroom in the original schools. Most had moved into administrative positions, retired, or moved away. Interestingly, seven of the teachers who entered in Year 2 and ten who entered in Year 3 were still in the classroom. This means that of the 47 original participants, twenty remained in the classroom four years later. In addition, the building-level technology specialist and science specialist in one school and technology director in another school system, all of whom were instrumental to the original development and implementation of TIES, were still in place.

In TIES, the power of technology merged with a constructivist pedagogy in student-centered, project-based classrooms.

When the grant period terminated, partnerships to enhance the grant had been put in place and continued to influence the schools. The project schools committed financial resources to support the project, and plans were put in place for continued funding of additional teachers and classrooms at each school. Experienced TIES teachers were poised to provide continued leadership at their schools. They had shown their leadership by being mentors to new TIES teachers, presenting at conferences, and by developing and presenting technology seminars. Four years later, TIES teachers continued as the technology proponents and leaders within their schools.

Obstacles

Each year, one of the most significant and challenging barriers reported by the project team was a difficulty inherent to any change effort—aversion to change or fear of the unknown. The change from a traditional to a technology-based pedagogical approach is very dramatic and met with resistance in some classrooms. Overcoming that resistance through a slow and on-going change process and reaching the levels of enthusiasm eventually seen in TIES classrooms are certainly two of the most important accomplishments of the project.

That same enthusiasm remained evident four years after the official end of the project. Teachers and administrators in all three TIES schools in both of the original school systems met with the authors, participated in focus groups, provided individual interviews and sent in surveys indicating the current levels of use of technology in their classrooms. All were very forthcoming and presented their current situations honestly and without embellishment. Funding for continued project-based instruction using technology was evident in one system and the lack of sufficient funding was just as obvious in the other. The disparity in funding is the result of differences in local tax bases and academic funding priorities. Teachers in both systems were enthusiastic and quick to tout their successes. They were just as quick to point out deficiencies in funding and technologies that have materialized over the past four years.

Successes

At the beginning of each year of the four years of TIES, teachers set goals and objectives, planned their projects, and proceeded to develop and implement them with the assistance of project staff. Each year, all TIES teachers met the objective of creating hands-on, technology-based projects within their classrooms. In addition, as the project progressed, TIES teachers became instructional leaders who took on responsibility for professional development and mentoring. TIES teachers who have remained in the systems have continued in those roles. Other successes that emerged from the evaluation of the project included positive attitudinal changes toward the objectives of the project.

Those positive attitudes remained very much in evidence as the teachers talked with the authors both individually and in focus group settings. The Levels of Use surveys (see Appendix 1) that participants completed indicated that all participants continue to use technology at least at the refinement level indicating that they continue to make changes in their use of technology in their classrooms to increase technology outcome measures. This level of use is quite remarkable given the equipment, materials and funding constraints under which some of the respondents are laboring.

At the close of the project, schools had strong technology and science resource help systems in place, including TIES mentors from previous years. In spite of time issues, participants who were in the project during the first two years were very helpful to the new project participants both in the technical aspects of how to use equipment and in the pedagogical aspects of using technology as a tool for effective instruction. Returning teachers were very willing to share classroom management techniques with teachers struggling to adapt their classrooms to a new mode of instruction. These returning teachers now work as peer coaches rather than mentors, but the camaraderie and willingness to help and to share expertise that was seen early on in the project is still in evidence.

Results

The overarching goal of the TIES Project was to produce a successful, creative, and replicable model for inquiry- and project-based instruction that used technology to integrate science and other curricula. Quantitatively, we saw an increase in competency rankings in technology knowledge and skills, as measured by a TIES Technology Expertise/Comfort Survey and on the Levels of Use of Technology in the Classroom scale (adapted from the CBAM research, 1987). Other evaluation strategies included site visits, workshop observations, interviews with project personnel, interviews with participants, and comment cards reflecting attitudinal changes from participants. Outcomes anecdotally reported by teachers included shifts in their beliefs and actions from instructionism to constructivism.

The overarching goal of the TIES Project was to produce a successful, creative, and replicable model for inquiry- and project-based instruction that used technology to integrate science and other curricula.

The Levels of Use of Technology in the Classroom self-report scale (Appendix 1, adapted from the CBAM research, 1987) was administered to all participants in the third and fourth years of the project and again in the follow-up phase. Initially, a clear distinction could be made between the levels of use of participants new to the project and those who had been with TIES for one or two years prior to the administration of the instrument. While new participants reported a wide range of levels of use, beginning at Level 0 (Nonuse) and continuing upward through Level IV (Refinement), no returning participant reported a level of use below Level III (Mechanical Use). Also of interest is the rapid movement of Year 3 participants up the Levels of Use scale, as compared to a more gradual movement for teachers who began the project in the first two years. Based on participant comments to a series of open-ended questions and on interview responses, this was presumed to be a result of mentoring provided by Year 1 and 2 teachers, as well as indirect exposure to the project before actually becoming a part of it. Year 4 participants showed limited growth; however, they were only in the project for one year, which is too short a period to allow for valid, reasonable conclusions to be drawn at that time. After five years, all of the participants reported use at Level IV (Refinement) or above. There was no discernible difference in Levels of Use among participants who entered the program in different years and all who were able to be reached were still actively engaged in the use of technology as an integral part of their work.

The project team noted some unanticipated beneficial outcomes. The comment cards used for formative evaluation indicated that the internal mentoring, support, and the coaching network were much stronger than proposers initially anticipated. Additionally, teachers reported that students wrote about their TIES projects with much less prodding than in traditional writing assignments.

The project team was also surprised, not that teacher attitudinal changes occurred, but by the extent of those changes, as evidenced in the comment cards. The magnitude of observed and anecdotally reported changes from a didactic to a student-centered teaching environment was much greater than proposers anticipated at the outset. Much to the delight of the proposers, these initial attitudinal changes as noted at the end of the project period, based on interviews and surveys, were still in evidence four years later.

Implications

"Fundamental beliefs are formed over time through active engagement with ideas, understandings, and real-life experiences….Deep change occurs only when beliefs are restructured through new understandings and experimentation with new behaviors" (Loucks-Horsley, S., et al., 2003, p. 49). For change to occur, things take time. This study exemplifies these beliefs. Initially, teachers who participated in the project for three or four years showed greater changes than those with only one or two years experience. Only participants who were in the project for more than two years reached Level V (Integration) or VI (Renewal) on the *Levels of Use of Technology* scale; not all veteran participants ever rose above Level IV (Refinement). After five years, all participants reported levels of use at Level IV or above, and no discernible difference was noted among participants by year of entry. The change literature, as well as our own experiences with this project, has led us to conclude that significant behavior changes require at least three to four years of implementation and on-going support to become institutionalized within the classroom and that institutionalization is retained as a part of regular classroom practice long after the end of the funding period.

See Appendix 1 on page 57

References

Archer, J. (October 1, 1998). Technology counts. *Education Week*. 6-10.

Bransford, J.D., Brown, A.L., and Cocking, R.R., Eds. (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

CBAM Project, Research and Development Center for Teacher Education, The University of Texas, 1987.

Jensen, E. (1998). Teaching with the brain in mind. Alexandria, VA: Association for Supervision and Curriculum Development. Science Teachers Association.

Loucks-Horsley, S., Love, N., Stiles, K.E., Mundry, S., and Hewson, P.W. (2003). *Designing professional development for teachers of science and mathematics* (2nd ed.). Thousand Oaks, CA: Corwin.

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

Shane, P.M. and Wojnowski, B.S. (2005). *Technology integraton enhancing science: Things take time*. Science Educator 14(1)

Sylwester, R. (1995). A celebration of neurons: An educator's guide to the human brain. Alexandria, VA: Association for Supervision and Curriculum Development.

Patricia Mills Shane is associate director, Center for Mathematics and Science Education, The University of North Carolina at Chapel Hill, NC. Correspondence concerning this article may be sent to pshane@email.unc.edu.

Brenda Shumate Wojnowski is the program officer, T-STEM Initiatives, Texas High School Project, Communities Foundation of Texas, Dallas, Texas 75225

See Appendix 1 page 58

Appendix 1

CBAM Levels of Use of Technology in the Classroom

Name_______________________________________

Year 1 (returning) _______

Year 2 (returning) _______

Year 3 (returning) _______

Year 4 (new) _______

Please circle the number that best reflects your current level of use of technology in your classroom.

Levels of Use	Behavioral Indices of Level
VI Renewal	I am seeking more effective alternatives to the already routinely established use of technology in my classroom.
V Integration	I am making deliberate efforts to help others to use technology in their classrooms.
IV Refinement	I am making changes in my use of technology in my classroom to increase outcomes.
III Mechanical Use	I am using technology in my classroom, but it is not always coordinated to my course of study.
II Preparation	I am preparing to use technology in my classroom.
I Orientation	I am seeking information on using technology in my classroom.
0 Nonuse	I am not taking any action in regard to using technology in my classroom.

(Adapted from the CBAM Project, Research and Development Center for Teacher Education, The University of Texas, 1987.)

Information for Authors

The *Science Educator* is directed to those concerned with science supervision at all levels. The journal is of primary interest to science supervisors, science department heads, science teachers, elementary science specialists, and college and university educators. Generally, articles are sought that aid practitioners in making more informed decisions about current or future operations, and which improve science teaching and learning in our nation's schools. Of particular interest are topics that deal with the following issues: curriculum and instruction, science education reform, inservice teacher enhancement, assignments/transfers/workload, teaching and learning, needs assessment/research and evaluation, resources and funding, school facility and equipment, textbook selection, and school board policies and practices.

1. All manuscripts should be sent to rhotonj@etsu.edu or in triplicate to Jack Rhoton, Editor, *Science Educator*, East Tennessee State University, P.O. Box 70684, Johnson City, TN 37614-1709.
2. Manuscripts should be typed double-spaced on standard 8.5 x 11 inch paper. Figures, tables and pictures must be submitted in camera ready condition. Tables should be typed on separate pages at the end of the manuscript.
3. The first page of the manuscript should contain the article title, author's name, affiliation, and address to which correspondence and proofs should be sent. The author should not place his/her name on the other manuscript pages so that anonymity can be maintained during the review process.
4. References cited in the manuscript should be listed at the end of the manuscript. Provide accurate and complete information. For questions of format follow the Publication Manual of the American Psychological Association (Fourth Edition) and Webster's College Dictionary.
5. Letters to the Editor are welcome and will be considered for publication if they respond to published articles or represent position statements for the purpose of raising issues or heightening awareness of alternative viewpoints regarding various phases of science supervision.
6. Authors who are not members of the National Science Leadership Association (NSELA) will be assessed a publication fee of $50.00 per article. Non members may avoid this charge by joining NSELA prior to their manuscripts being published. The institutional charge is $75, $30 individual journal-only. Membership dues are payable at the beginning of each calendar year. Information and application for membership can be obtained from Susan Sprague, NSELA Executive Director, 2011 Holly Drive, Prescott, AZ 86305.
7. Receipt of manuscripts will be acknowledged. All submitted manuscripts will be reviewed by at least three members of the Editorial Board. Authors will be notified by the Editor regarding the recommendation on publication of the manuscript. Transfer of the author's copyright to the National Science Education Leadership Association is a condition for publication.