Evaluating Science Research Experience For Teachers Programs and Their Effects on Student Interest and Academic Performance: A Preliminary Report of an Ongoing Collaborative Study by Eight Programs

Jay Dubner & Samuel C. Silverstein
Columbia University, New York, NY
Nancy Carey & Joy Frechtling
Westat, Inc., Rockville, MD
Tamra Busch-Johnsen
Business Education Compact, Beaverton, OR
Jeannie Han & George Ordway
University of Texas, Dallas, TX
Nancy Hutchison
Fred Hutchinson Cancer Center, Seattle, WA
Janet Lanza & Jim Winter
University of Arkansas, Little Rock, AR
Jon Miller
Northwestern University, Chicago, IL
Paul Ohme & James Rayford
Georgia Institute of Technology, Atlanta, GA
Kathryn Sloane Weisbaum
Consultant to Westat, Inc, Rockville, MD
Kaye Storm
Industry Initiatives for Science and Math Education, Cupertino, CA
Elda Zounar
Idaho National Engineering and Environmental Laboratory, Idaho Falls, ID

ABSTRACT

Formal programs that provide research experiences for teachers (RET) have been in existence for more than 20 years. Currently there are more than 70 formalized *Scientific Work Experiences Programs for Teachers* (SWEPTs) nationwide. [1] The underlying assumption of most RETs is that these intensive summer work immersion experiences, coupled with appropriate follow-up activities during the school year, expand teachers' professional skills and networks, and thereby improve the performance of their students. Many SWEPTs have collected anecdotal evidence indicating their program's positive impact on teachers. Missing from all SWEPT evaluations is quantitative evidence that teacher participation in these programs affects student interest and performance in the subject taught by the SWEPT teacher. As professional evaluators attest, it is difficult to differentiate the roles of teachers and teaching practices in changing student academic interest and performance from other factors (e.g., curriculum, school administration, non-random assignment of students, etc.).

This study controls for many of these factors by comparing interest and achievement of students in classes of SWEPT teachers with students in classes of

comparison teachers in the same school and teaching the same subject. The study's longitudinal design is commensurate with the philosophy and practices of the participating SWEPTs.

INTRODUCTION

Qualified science teachers are in short supply. SWEPTs provide a mechanism for elevating the performance of both new and experienced science teachers and for motivating experienced science teachers to remain in teaching. The findings that are emerging from this study will enable administrators, educators, policy makers, and corporations to determine whether further investments in SWEPTs can improve the quality of science education in U.S. schools, and thereby elevate student interest and achievement.

Teacher expertise is one of the most important factors in raising student achievement, [2] and investing in teachers is the most cost-effective way to accomplish this goal. [3] SWEPT Programs continue to receive feedback from their participants that their participation in science research experiences provides them with new insights, knowledge, and resources; that they are encouraged to implement more constructivist instructional practices; and that they are prepared to provide students with more up-to-date, relevant, and stimulating educational experiences in science. We describe below preliminary evidence obtained by Columbia University's SWEPT, and by this study, that these effects on teachers lead to improvements in student interest and achievement in science.

EXPERIMENTAL DESIGN

Description of the eight SWEPTs participating in this study: These SWEPTs are representative of most geographic regions of the U.S.: Northeast [NY], Southeast [GA], South [AR], Southwest [TX], Rocky Mountains [ID], West [CA], and Northwest [OR, WA]. They represent the four types of sponsoring organizations: government [ID], university [AR, GA, NY, TX], industry [CA, GA, OR], and independent research institutes [WA] (Table 1). They are among the nation's largest SWEPTs, have similar goals and participants (i.e., public high school science, math, and technology teachers), incorporate the full spectrum of "best practices" [4][5][6] are attentive to national standards-based education, [7] and have substantial experience with evaluation. They have been in continuous operation for 8-17 years. They are financially stable, governed by community organizations, universities, and/or research institutes, and committed to this study. While each of the eight SWEPTs has a unique "personality," all share the aim of providing hands-on experiences to science, math and technology teachers to improve the quality of science, math, and technology education for all students. Thus, while each teacher's SWEPT experience is unique and differs depending on the setting (e.g.; academic, government, industry) and type of work performed (e.g.; physical, biological, or medical science, applied technology, math), the Multi-Agency Study of Teacher Enhancement Programs indicates that teachers derive very similar benefits from participating in these programs. [6] It also notes that the benefits of teacher participation in a professional development program transfer most readily to the classroom if the schools to which these teachers return are ready to support constructivist educational practices. In a

constructivist classroom, students develop hypotheses and explore alternative explanations or methods to test them. They are encouraged to weigh information from these tests with previous experiences or understanding of the topic. Students then "construct" a new understanding of the subject matter. A constructivist classroom is therefore student-centered, where students are given the opportunities to pose questions and engage in hands-on activities or consult primary sources to find the answers to their questions. Anticipating this finding, the eight participating SWEPTs make every effort to help schools utilize the skills teachers acquire through SWEPT participation. A tangible manifestation of this effort is the enhancement funds provided to teachers by half of the SWEPTs collaborating in this project. These funds allow teachers to purchase equipment and supplies for their classrooms and schools so their students can experience first-hand the technologies and concepts acquired by their teachers through their SWEPT experiences. SWEPTs also are willing participants in systemic reform efforts.

TABLE 1

Characteristics of the eight participating SWEPTs

Program	AR	CA	GA	ID	NY	OR	TX	WA
Year Founded	1990	1984	1991	1988	1990	1985	1993	1991
Total # teachers served to date	397	761	450	374	137	855	83	200
Total # of teachers served in 2000	32	118	80	70	33	60	8	10
Total # served in 2000 for first time	26	78	40	14	13	31	8	0
Length of Summer Program	8 weeks	8 weeks	4-8 weeks	8 weeks	8 weeks	3-10 weeks	8 weeks	13 days
Stipend	$4,400	$5,600	$2,500 to $5,000	$4,000 to $6,000	$6,000	$12/hr (avg)	$4,000	$500
Classroom Supplemental funds	0	$450	$600	0	$1,000	0	$1,000	0
Post-Program Activities	yes	yes	yes	yes	yes	yes	yes	yes

The study is designed to examine the fundamental premise underlying all SWEPTs. That is, teachers who have experience in the practice of science, and in the use of science in the "real world," can better communicate the concepts and value of science to their students. Implicit in this premise is that SWEPT experiences affect the approaches teachers employ in their classrooms, and that these new approaches have positive effects on student interest and achievement in science.

To implement significant changes in teaching strategies takes even the most able and experienced teacher a year or more to accomplish. For this reason the study design encompasses an appropriate time frame between teacher entry into a SWEPT and measurements of teacher attitudes and of student attitudes and achievement for changes

in these parameters to become manifest. For example, introduction of a new lab or curriculum unit may be implemented in the span of a semester. In contrast, more pervasive changes in educational practices, such as encouraging students to take intellectual risks or to explore open-ended questions with no right or wrong answer, may take several years. For this reason this study examines the evolution of teachers' attitudes and practices and the interest and accomplishments of their students over the course of several years.

Two sets of teachers and students are being studied. One set consists of participating SWEPT teachers and their students. A second consists of comparison teachers and their students from the same schools. Teachers who have had previous SWEPT experiences are excluded from the comparison teacher group.

All eight SWEPTs administer the same instruments and employ the same methods to collect data on attitudes and achievement of public high school students in science (biology and chemistry) and math (algebra and geometry) classes taught by participating and comparison teachers. We have targeted these subjects because they are the core science and math courses taken by most students. Subject-specific test instruments, student attitude surveys, and teacher and mentor surveys are used to answer specific questions (Table 2). Information on program practices and design, on profiles of participating teachers, and on demographics of schools and districts also are part of the data set.

TABLE 2

The relationship between the study questions and the major sources of data

Exhibit 1 Data Collection Approach	Student attitude survey	Student cognitive assessment	Teacher survey
1. Does teacher participation in SWEPT result in an increase in *instructional experiences* for students that are constructivist in nature, and that relate issues in science and math to the roles and applications of these subjects in everyday life?	X		X
2. Do students of SWEPT teachers demonstrate more *positive attitudes, aspirations, and awareness* of science and math than students of comparison teachers?	X		
3. Do students of SWEPT teachers demonstrate larger gains in *math and science knowledge* than students of comparison teachers?		X	
4. Do students of SWEPT teachers demonstrate larger gains in ability to *apply math and science concepts to solve problems* than students of comparison teachers?		X	
5. Do *patterns of change* in classroom experiences or student outcomes vary according to:	X	X	X
• the characteristics of the students?			
• the characteristics of the teachers?			
• the type of placement and the duration of the SWEPT experience?			
• the supports for innovation available to the SWEPT teacher at his or her school?			

Teachers: Participating teachers are surveyed with respect to attitudes and teaching practices on entry into a SWEPT, and at the beginning and at the end of each school year. A comparison teacher teaching the same subject in the same school as a participating teacher is surveyed with respect to attitudes and teaching practices at the beginning and end of each school year.

Students: All students in two classes in biology, chemistry, math, or algebra taught by a SWEPT teacher, and in one class in the same subject taught by a comparison teacher are surveyed with respect to attitudes and achievement at the beginning and end of the course. The achievement tests were constructed using previously validated items from TIMSS, NAPE and SAT exams. This design allows us to evaluate affective and cognitive outcomes in the same students, thereby providing a unique opportunity to study relationships between affective and cognitive changes.

Mentors: All supervisors/mentors of participating teachers are surveyed at the end of each teacher's SWEPT experience to obtain information about the teacher's level of engagement and ability to accomplish the tasks assigned to him/her.

Criteria for selecting SWEPT and comparison teachers for the study: SWEPT participants selected for the study must be public high school science, math, or technology teachers who have not previously participated in a SWEPT or SWEPT-type experience and teach in a school in which there is a comparison teacher who has not had previous SWEPT experience, teaches the same subject as the participating teacher, and agrees to serve as a comparison teacher for purposes of this study. At present, 68 SWEPT and 68 comparison teachers have been enrolled in the study. They comprise two cohorts. Cohort I SWEPT teachers were selected in late spring 1999. Cohort II SWEPT teachers were selected in late spring 2000. Comparison teachers for both cohorts were selected at the beginning of the subsequent school years. Cohort I SWEPT teachers completed attitude surveys just prior to their entry into the 1999 program, on completion of their SWEPT experiences, and at the end of the 1999-2000 academic year. Comparison teachers completed attitudinal surveys at the beginning of the 1999-2000 academic year and in the Spring of the 1999-2000 academic year. Cohort II SWEPT and comparison teachers completed attitude surveys in a similar manner during the 2000-01 academic year.

Criteria for selection of classes of SWEPT and comparison teachers for participation in this study: Two introductory classes in biology, chemistry, algebra, or geometry taught by each Cohort I participating teacher in a school containing a comparison teacher were selected for participation in the study in September 1999 and in September 2000. One introductory class taught by a comparison teacher in the same school and in the same subject as that taught by the participating teacher was selected for participation in the study in September 1999 and in September 2000. Similarly, classes of students of participating and comparison teachers in Cohort II were selected in September 2000. All students in the selected classes of Cohort I SWEPT and comparison teachers completed attitudinal and cognitive tests in September 1999 (pre-course), at the conclusion of the course, and in September 2000 (pre-course). Similarly, attitudinal and cognitive tests were completed by all students in selected classes of Cohort II SWEPT and comparison teachers in September 2000 (pre-course), and will be completed by all students in these classes of Cohort I and II SWEPT and comparison teachers at the conclusion of the course.

PRELIMINARY RESULTS AND DISCUSSION

Initial findings regarding the samples and the procedures: Initial analyses have focused on documenting the equivalence of our samples and utility of our procedures.

The samples. We were concerned about the possible confounding that might result if systematic differences were found between our SWEPT and control populations. Thus, a critical first step was examining the equivalence of both the teacher and student samples at the beginning of the study.

- Our analyses show that the SWEPT participants and comparison teachers in Cohort I are essentially equivalent in terms of demographic characteristics (Table 3). The only difference between the two groups was the extent to which teachers had participated in professional development activities that addressed the application of technology to teaching in their main subject area. Comparison teachers were more likely than SWEPT teachers to have had more than 8 hours of professional development in this area.

TABLE 3

Percent of Cohort I SWEPT and comparison teachers, by selected teacher characteristics

Teacher characteristic	SWEPT teachers (n=54)	Comparison teachers (n=49)
Years teaching experience		
3 or fewer	13	16
4 to 9	45	30
10 to 19	32	39
20 or more	9	16
Holds advanced degree	54	49
Teaching certificate		
Regular, standard	89	95
Temporary, provisional	8	2
Emergency	4	2
More than 8 hours professional development in the following areas:		
Methods of teaching	67	74
In-depth study of your main teaching area	73	74
Applications of technology	50	79
Race		
White	84	78
Black	8	9
Other	8	13

Note: Based on the number of SWEPT participants who completed Pre-Program Surveys and comparison teachers who completed Pre-Teaching Surveys.

SOURCE: SWEPT Pre-Program Survey and Pre-Teaching, Comparison Teacher Survey

This is an important finding because it indicates that SWEPT teachers are not more experienced or more highly educated than comparison teachers, making it more likely that any differences found between their performance and that of comparison teachers can be attributed to participation in a SWEPT. We also found no systematic differences in the samples of the 3,139 students taught by these two groups of teacher (2,056 students in classes of SWEPT teachers and 1,083 students in classes of comparison teachers). Initial test scores on the pre-tests of each subject showed no significance differences between the two groups. Table 4 shows the scale scores and standard errors of these scores for each group of students.

TABLE 4

Pre-test scale scores for Cohort I, year 1 students in biology, chemistry, algebra, and geometry classes, by type of teacher

Subject	Pre-test scale score	Standard error of the scale score
Biology (n=1114)		
SWEPT students	258.0	5.4
Comparison students	248.2	5.4
Chemistry (n=215)		
SWEPT students	243.0	10.6
Comparison students	244.8	12.5
Algebra (n=536)		
SWEPT students	238.7	7.4
Comparison students	239.5	7.7
Geometry (n=513)		
SWEPT students	238.6	7.8
Comparison students	238.4	7.5

Note: *The scale scores reported here are based on an overall mean of 250 and a standard deviation of 50 for both pre- and post-tests combined. The n for each group is the number of students for whom there was both pre- and post-test data.*

- Both demographic data and initial test scores revealed no systematic differences between the students assigned to the SWEPT teachers and those assigned to the comparison teachers.

This finding increases the likelihood that any differences observed in achievement of students in classes of SWEPT vs comparison teachers can be ascribed to the impact of the teacher.

The procedures. A second analysis examined the instruments used, especially the instruments used to examine student attitudes and cognition to assess the extent to which these tools were useful and sensitive measures of the variables they were designed to assess. Generally we found that the instruments worked as expected, but some limitations were also found.

- While inspection of student performance on the pre-tests showed that the items taken from NAEP seemed to have equal difficulty levels for students of SWEPT

and comparison teachers participating in this study as students who constituted the NAEP sample group, many items turned out to be too difficult for students of both SWEPT and comparison teachers on the post-test. This is somewhat surprising, as NAEP has not been considered to be a very difficult test. Nonetheless, even students with the longest exposure to SWEPT and comparison teachers were not able to correctly answer some of the hardest items.

- Students with the longest time interval between pre- and post-test, presumably those who had received the greatest among instruction, showed greater gains than those with shorter intervals between the pre- and post-test on all of the cognitive tests except algebra. Students in trimester algebra classes showed more gains than students in full-year classes. However, only the difference in the geometry test is significant (Table 5).

TABLE 5

Mean pre- post-test differences for the four cognitive subjects, by subject format

Subject	Mean pre-post difference	Standard deviation
Biology		
Full year courses (n=894)	.18	.61
Semester courses (n=93)	.14	.49
Trimester courses (n=127)	.04	.61
Chemistry		
Full year courses (n=134)	.42	.77
Semester courses (n=81)	.23	.64
Algebra		
Full year courses (n=460)	.39	.65
Semester courses (n=16)	.006	.72
Trimester courses (n=29)	.74	.56
Geometry		
Full year courses (n=242)	1.16	.76
Semester courses (n=154)	.49	.54
Trimester courses (n=117)	.78	.49

- With regard to the student attitudes survey, factor analysis indicates that items cluster in predicted ways and that students' attitudes can be described in terms of a small number of meaningful dimensions (Table 6). Thus the instrument appears to be yielding useful data.

TABLE 6

Scale Names and their Item Content for the Student Attitude Survey

I. Importance/usefulness of science/math

- 1b. Science/math is useful in everyday life
- 1e. Science/math challenges me to use my mind
- 1f. The science/math instruction will be helpful for me in the future
- 1i. Advancements in science/math responsible for U.S. standard of living
- 1k. Knowing science/math really doesn't help get a job
- 1p. Overall, m and s have caused more good than harm

Ii. Beliefs about science/math work and careers

- 1c. Mathematicians/scientists often lack social skills
- 1g. Mathematicians/scientists usually work as part of a team
- 1m. Working as a mathematician/scientist sounds pretty lonely
- 1n. Studying hard in science/math is not cool to do
- 3e. Past 12 mos. Collected information about science/math careers

Iii. Beliefs about own science/math ability

- 1a. I enjoy science/math
- 1d. Doing science/math often makes me feel nervous or upset
- 1h. I am good at science/math
- 1j. I usually understand what we are doing in science/math class
- 1l. Science/math is difficult for me
- 1q. I plan to take more advanced science/math courses at this school

Iv. Interest and involvement in science/math

- 3a. Past 12 mos. Participated in a science/math or computer club
- 3b. Past 12 mos. Visited a science museum
- 3c. Past 12 mos. Watched science/math TV shows
- 3d. Past 12 mos. Read science/math magazines or news articles
- 5. Effort usually put into science/math school work

Initial Findings Regarding the SWEPT experience. Findings from the first cohort of teachers replicated previous work showing that teachers report receiving significant benefits from their SWEPT experiences. The data collected document positive effects of teachers' SWEPT experiences on their beliefs and practices (Table 7).

- SWEPT teachers displayed a positive change on all attitude scales between entry and completion of their SWEPT experiences. This may be a reflection of their enthusiastic response to their SWEPT experiences and their motivation to acquire new skills.

- In pre-program surveys of SWEPT teachers and pre-academic year surveys of comparison teachers, both groups reported similar attitudes with respect to implementation of inquiry-based student activities and teaching methods and alternative assessment methods. However, by the end of their first post-SWEPT year of teaching, SWEPT teachers were more positive about their ability to explore new educational resources and reported greater use of inquiry-based educational methods than comparison teachers. In contrast, comparison teachers' end-of-year scores on the same parameters fell below their beginning-of-year

scores, and well below those of SWEPT teachers. Moreover, comparison teachers reported greater reliance on traditional teaching methods (e.g., lectures) than SWEPT teachers.

TABLE 7

Differences in teacher attitudes between pre- and post-administration of the teacher surveys, by teacher group

Teacher attitude	Mean pre-post difference	Difference between the 2 groups
Traditional goals and objectives		
SWEPT teachers	-7.6	7.7
Comparison teachers	.1	
Inquiry-based goals and objectives		
SWEPT teachers	3.3	28.2*
Comparison teachers	-24.9	
Traditional student activities		
SWEPT teachers	11.9	15.4
Comparison teachers	-3.5	
Inquiry-based student activities		
SWEPT teachers	7.2	22.4*
Comparison teachers	-15.2	
Traditional teaching methods		
SWEPT teachers	2.4	15.31*
Comparison teachers	-12.9	
Exploring new resources		
SWEPT teachers	15.1	20.8*
Comparison teachers	-5.7	
Use of assessment		
SWEPT teachers	-5.4	9.6
Comparison teachers	-15.1	
Teacher efficacy		
SWEPT teachers	30.7	33.5*
Comparison teachers	-2.74	
Teacher identity		
SWEPT teachers	6.1	11.2
Comparison teaches	-5.1	

* *indicates the difference between the two pre-post differences is significant*

It also is interesting to note that the mentors to whom these teachers were assigned were quite positive about their experiences with the SWEPT teachers. Some mentors reported that they gained a greater appreciation for teachers and the work that they do.

Impacts on Students. Analyses of pre- and post-test scores on the student attitude and cognitive tests showed no differences between the students of the SWEPT vs. the

comparison teachers during this first year of data collection. The differences that were found were small and not statistically significant. There was large variability in gains across the groups. This lack of significant differences could, at this point, be a result of many different things—time since SWEPT participation, small sample size, differences related to variation in course duration and subsequent level of exposure to the SWEPT trained teacher, or the ineffectiveness of the SWEPT approach with regard to the outcome variables on which the study has focused.

It is noteworthy that Columbia University's study found the largest differences in interest and achievement in science between students of teachers participating in its program vs the comparison teachers in the second and subsequent years after teacher entry into Columbia's program. Therefore, it will be important to examine the patterns of outcomes in the study's second year as SWEPT teachers refine and more fully implement the concepts and techniques learned during their SWEPT experiences.

Unanticipated complications encountered in the course of conducting this study.

The study design assumed students would be tested at the beginning and end of a year-long course, and our test instruments were designed to cover material in the full course. In practice, however, we found that certain schools reassigned students to a different teacher at the end of the first semester or trimester of a year-long course. Consequently, some/many students initially enrolled in classes of SWEPT teachers were exposed to that teacher for only one trimester or semester. While one might speculate about the deleterious educational effects of such scheduling, there is no doubt that they have added to the difficulties of conducting a study such as this one. In some cases it will likely require elimination of data for SWEPT and comparison teachers and their students from the study. Elimination of these teachers and students will substantially reduce our "useable" sample size. This, in turn, will require extending the study for an additional 1-2 years in order to enroll a number of teachers and students sufficient to yield statistically significant results.

CONCLUSIONS

The preliminary results of the first year of this multi-site study indicate that its design controls effectively for many of the most important variables that make it difficult to correlate changes in student academic interest and achievement with teacher preparedness, practices, and performance. Specifically, the study has found no differences in educational background or socio-economic status of SWEPT vs comparison teacher in the same school, no pre-course difference in interest in science or performance on cognitive tests of students in classes of SWEPT vs comparison teacher in the same school. The study also has found greater post-course interest and enthusiasm for science in classes of SWEPT vs comparison teachers in the same school.

The preliminary results suggest this study will replicate previous studies with respect to teachers' perceptions of the benefits of a SWEPT experience. For the reasons indicated, the number of teachers and students enrolled in the first year of the study was not sufficient to yield statistically significant data on the study's central question: namely,

does teacher participation in a SWEPT have a measurable impact on student interest or achievement in science? We anticipate that two more years of data collection will be required to enroll a sufficient number of teachers and students to obtain statistically significant results.

ACKNOWLEDGEMENTS

The authors wish to thank the staff of the participating SWEPTs for their many contributions to the administration of this study and to Columbia University's *Summer Research Program for Secondary School Science Teachers* for permission to cite their unpublished results. Research supported by the National Science Foundation under award # 9812142

REFERENCES

[1] Directory of Scientific Work Experience Programs for Teachers, Industry Initiatives for Science and Math Education, November 1996

[2]Darling Hammond, L., & Loewenberg Ball, D. "Teaching for high standards: What policymakers need to know and be able to do." Paper prepared for the National Education Goals Panel, Washington, D.C., July 1997

[3]Greenwald, R., Hedges, L.V., & Laine, R.D., The effect of school resources in student achievement. Review of Educational Research 66: 361-396, Fall 1996

[4] The Professional Learning Community in the Context of Scientific Work Experience Programs for Teachers, C. Orner, March 1996

[5] Best Practice in Action, Westat, Inc., SRI International and National Center for Improving Science Education, November 1995

[6] Best Practice in Action: Final Report of the Multi-Agency Study of Teacher Enhancement Programs, J. Frechtling, Westat, SRI International and the National Center for Improving Science Education, December 1997

[7] National Science Education Standards, National Academy Press, 1996