Journal of Science Education and Technology, Vol. 14, No. 3, September 2005 (© 2005)
DOI: 10.1007/s10956-005-7199-8

The Gateway Science Workshop Program: Enhancing Student Performance and Retention in the Sciences Through Peer-Facilitated Discussion

Denise Drane,[1,3] H. David Smith,[1] Greg Light,[1] Larry Pinto,[2] and Su Swarat[1]

Minority student attrition and underachievement is a long-standing and widespread concern in higher education. It is especially acute in introductory science courses which are prerequisites for students planning to pursue science-related careers. Poor performance in these courses often results in attrition of minorities from the science fields. This is a particular concern at selective universities where minority students enter with excellent academic credentials but receive lower average grades and have lower retention rates than majority students with similar credentials. This paper reports the first year results of a large scale peer-facilitated workshop program designed to increase performance and retention in Biology, Chemistry, and Physics at a selective research university. After adjusting for grade point average or SAT-Math score, workshop participants earned higher final grades than nonparticipants in Biology and Chemistry, but not in Physics. Similar effects on retention were found. While, positive effects of the program were observed in both majority and minority students, effect sizes were generally largest for minority students. Because of practical constraints in Physics, implementation of the program was not optimal, possibly accounting for the differential success of the program across disciplines.

KEY WORDS: peer learning; collaborative learning; workshop; minority; science; biology; chemistry physics.

INTRODUCTION

Poor Academic Performance of Minority Students

The small number of high-achieving, minority undergraduate students[4] who successfully persist in math and science-based courses of study is a national educational problem of enormous complexity and challenge. Minority attrition and underachievement is long-standing and widespread. Compared with majority students, students of color have had historically lower graduation rates, higher rates of attrition, and more reports of academic difficulty (Chavez and Maestas-Flores, 1991; Clewell and Ficklen, 1987; Grayson, 1998; Levin and Levin, 1991). Higher attrition and lower academic achievement during high school and college years lead to underrepresentation of minority students at the highest levels of academic achievement. According to the National Science Foundation (2004), of the PhDs awarded in the Sciences in 2001, 79.5% were white, 8.9% were Asian, 4.2% were African American, 4.1% were Hispanic, and less than 0.5% were Native American.

Education is typically considered to be the principal route toward upward mobility in industrialized societies (Croizet *et al.*, 2001). Thus,

[1]Searle Center for Teaching Excellence, Northwestern University, Evanston, Illinois.
[2]Department of Neurobiology and Physiology, Northwestern University, Evanston, Illinois.
[3]To whom correspondence should be addressed at Searle Center for Teaching Excellence, Northwestern University, 627 Dartmouth Place, Evanston, Illinois 60208-4181; e-mail: d-drane@northwestern.edu
[4]"Minority" students refers to students from groups typically underrepresented in the sciences, including African Americans, Hispanic Americans, and Native Americans, as well as women.

underrepresentation of minority students at the highest levels represents a barrier to successful integration of society in the United States. The significance of this barrier was highlighted recently in a report by the College Entrance Examination Board's (1999) *National Task Force on Minority High Achievement*

> Until many more underrepresented minority students from disadvantaged, middle class, and upper-middle class circumstances are very successful educationally, it will be virtually impossible to integrate our society's institutions completely, especially at leadership levels. Without such progress, the United States also will continue to be unable to draw on the full range of talents in our population during an era when the value of an educated citizenry has never been greater. (p. 2)

Poor achievement and attrition of minority students is a particular concern at selective universities where minority students enter with excellent academic credentials and high expectations, but receive lower average grades and have lower retention rates than majority students with similar academic credentials.

Approaches to Improving the Academic Performance of Minority Students in Higher Education

Universities have often assumed that the cause of poor academic performance in minority students lies within the individual (e.g. inadequate academic preparation), and have addressed the problem by offering extra tutoring, remedial instruction, special "prep" programs, or introductory programs especially for "at risk" students (Fullilove and Treisman, 1990). However, in general, these approaches have not been successful in improving achievement and retention of minority students (Steele, 1997).

One approach that has been successful in improving the performance of minority students is the 'workshop' approach developed by Treisman in the early 80's. Observations of the study habits of African American students and highly successful Asian American students in calculus led Treisman to hypothesize that poor performance of minority students is not due to inadequate preparation or lack of ability, but is due, at least in part, to their tendency to study in isolation (Fullilove and Treisman, 1990). Treisman developed the highly successful *Professional Development Program Mathematics Workshop* at the University of California Berkley to counter academic isolation in minority students and improve academic performance and retention. He recreated the study networks that he had observed in highly successful Asian American students by organizing students into study groups of 5–7 students that met for four hours per week to work on challenging problem sets. Each workshop group was assisted by a graduate student "facilitator" who stimulated discussion and answered questions but did not lecture or do the problems for the students. The program was not remedial in nature. In fact, problems were specifically designed to be more challenging than the regular coursework. Treisman's program was highly successful, with only 3% of African American workshop participants receiving grades of D and F in contrast to 40% of nonparticipants and 33% of a historical control group (Fullilove and Treisman, 1990).

Treisman's workshop program had several pedagogical features that were likely to have contributed to its success. First, the program involved problem-focused, collaborative work in small groups. This may have counteracted many of the difficulties associated with large introductory lecture courses such as the lack of opportunity for students to receive feedback, the lack of time for cognitive elaboration and reduced student engagement. Second, the development of a dynamic learning community created by the workshops would have offered students the opportunity to exercise critical judgment, analyze statements and causes, question underlying assumptions and check for underlying assumptions—which are among the most important learning experiences students in higher education can have (Bligh, 2000, Jaques, 2000; Light and Cox, 2001). Further, such experiences replicate scientific practice whereby empirical results are often interpreted, models refined and conclusions reached through processes of discussion and argument (Tien *et al.*, 2002).

Evaluation of Workshop Type Programs

The success of Treisman's program led to a proliferation of "workshop" style programs in several different disciplines, across a wide range of institutions (Bonsangue and Drew, 1995; Selvin, 1992). Only a relatively small number of institutions have published comprehensive evaluations of their programs. Results of these evaluations are, however, positive and suggest that workshop approaches have real promise for improving the performance and retention of minority students in science and engineering courses.

Participants in the *MathExcel* program in calculus at the University of Kentucky received higher mean grades (3.15 vs. 2.20) and were awarded a higher percentage of A's and B's (76.4 vs. 43.5) than students who did not participate (Freeman, 1995). There was also a positive impact of the program on retention with only 6.7% of workshop students withdrawing or failing, compared to 28.2% of nonparticipants. Workshop students and nonparticipants did not differ with regard to academic ability as measured by the ACT, indicating that differences observed in final grade were not due to preexisting differences in academic ability. In 1994, the Lexington Community College introduced a workshop program in intermediate algebra that was modeled on the *MathExcel* program. On average, workshop participants received grades that were 0.9 grade points higher than nonparticipants. Again, participants and nonparticipants did not differ with regard to incoming level of academic ability.

California State Polytechnic University developed the *Academic Excellence Workshop Program* in math for African American and Hispanics undergraduate students with science, math, or engineering majors (Bonsangue and Drew, 1995). The program had a positive impact on both grades and retention. Academic performance of workshop students was equal to that of European American and Asian American students. Course grades of minority participants were 0.6 of a grade point higher than those of nonparticipants. Impact on retention was particularly striking with only 15% of workshop students dropping out compared with 52% of minority nonparticipants and 50% of majority participants.

As noted earlier, workshop programs have been successfully implemented in disciplines other than Math. The University of Kentucky reported pilot data from their *Bioexcel* workshop program in Biology (Cohen, W., 1997, *BioExcel at the University of Kentucky*. Unpublished Project Description). A graduate teaching assistant and two advanced undergraduate students served as facilitators in this program. Sixty-six percent of workshop students earned a B or higher in comparison to 51% for the entire class. Gosser *et al.* (1996) evaluated the City College of New York's (CCNY) *Workshop Chemistry Project*. Students were randomly assigned to participate or not participate in the workshop program. A higher proportion of workshop participants earned a C or above than nonparticipants.

Workshop Programs at Selective Institutions

The University of Rochester implemented the Peer-Led Team Learning (PLTL) workshop program in Organic Chemistry in 1995. After successful piloting, the program was adopted as a core component of the course and all students (both minority and majority) were required to participate. After adjusting for SAT-M and SAT-V scores, students who participated in the program from 1996 to 1999 received higher exam scores and grades and had higher retention rates than a historical control group who took Organic Chemistry from 1992 to 1994 (Tien *et al.*, 2002). On average grades of workshop participants were 0.28 grade points higher than those of the control group ($p < 0.01$; effect size $d = 0.24$). Exam scores of minority workshop students ($M = 502.7$) were higher than those of both the minority control group ($M = 436.6$) and the majority control group ($M = 483.0$).

In 1997, Northwestern University began to address the problem of poor achievement of minority students, through the introduction of the Honors Workshop program in Biology 210, a comprehensive, three quarter survey course of Biology generally taken in the sophomore year. (Born *et al.*, 2002). The program was implemented as an experimental program from 1997–1999. All students (both minority and majority) who were registered in Biology 210 were invited to take part in the program. Because of the small number of minority students registered in Biology 210, all minority students who accepted the invitation to the program were admitted. No minority students were randomized to the control group. Instead minority students who were enrolled in Biology 210 in the previous year served as a historical control group for minority students.

Majority students who accepted the invitation were randomly assigned either to the workshop program or to a control group which did not receive any intervention. Existence of this control group enabled specific comparisons to be made between students in the workshop and students who demonstrated the same interest, motivation, and other attributes that may lead to a desire to participate. By randomly assigning participants, one could be more confident that these characteristics, and any additional, unknown characteristics, were evenly distributed across groups. In addition to the randomized control group, a third, self-selected comparison group for majority students was formed from students who declined the invitation to take part in the program.

A detailed evaluation of the Honors Workshop program was conducted by Born *et al.* (2002). After adjusting for prior GPA, exam scores of majority students enrolled in the workshop program were significantly higher than those of majority students in the randomized control group, with the advantage seen across all three quarters. Significant differences in exam scores were also seen between minority workshop students and the historical control group. In the first quarter, the performance of minority students increased, whereas it decreased in the historical control group. In addition, over the three quarters, the mean exam scores of minority participants consistently fell between majority workshop and control groups. Dramatic differences were seen between the minority participants and historical controls in terms of numbers of D's and F's earned. None of the minority workshop students scored a D or F, whereas 24% of the historical control group earned these grades.

Study Aims

In 2001, Northwestern University received funding from the Andrew W. Mellon foundation to further develop the GSW program and to assess its impact on performance and retention of students in Biology, Chemistry, and Physics. This paper reports results from the first year of that evaluation. The evaluation differs from Born *et al.*'s earlier evaluation of the Honors Workshop program in a few noteworthy ways. First, it is not an evaluation of a controlled experiment where researchers can utilize random assignment to make causal inferences. Rather, it is an evaluation of a program implemented under "real-world" conditions; an academic opportunity of which students could avail themselves or not. In addition, the program was expanded considerably and extended into other science disciplines, requiring more coordination and cooperation between programmers, evaluators, faculty, and departments.

In addition to assessing the impact of the workshop program on performance and retention of minority students in the sciences, a second aim of this study is to determine if the workshop program is effective across disciplines. Although, workshop style programs have been implemented in several disciplines, few cross disciplinary evaluations have been conducted. Springer *et al.* (1998) found that small-group learning programs had a more positive effect on attitudes in students in science than students in math and engineering. However, such cross disciplinary differences are difficult to interpret because student populations and workshop formats differ significantly across institutions. This study evaluates a workshop style program (The Gateway Science Workshop Program), implemented across three different disciplines: biology, chemistry, and physics, at the same institution.

Assessment of the effectiveness of workshop type programs in improving performance and retention of minority students is important, as there is evidence that the gap in grade point average (GPA) between minority and majority students is largest among students with the highest levels of academic preparation as measured by SAT score (Vars and Bowen, 1998). Poor performance and attrition of highly talented minority students from highly selective universities results in the loss of some of the most gifted minority students from the sciences and the loss of future leaders of the scientific community. As indicated in the literature review, the effectiveness of workshop programs in improving academic performance and retention of minority students at highly selective institutions has been addressed in only two studies; Tien *et al.*'s study of the PLTL program in Organic Chemistry at the University of Rochester and Born *et al.*'s study of the Honors Workshop program at Northwestern University. Both of these studies used historical control groups for minority students in their evaluations. This study aims to further assess the efficacy of a workshop style program at Northwestern University, using contemporary control groups, i.e., students enrolled in the same course and academic year as the workshop students.

METHODS

Study Context

Based on the promising results Born *et al.*'s experiment, the workshop program was made available to all Biology 210 students in the 1999–2000 academic year and to Chemistry 101 and Physics 130 and 135 students in 2001 as the Gateway Science Workshop (GSW) program. The GSW program was not implemented as an experiment as it had been earlier in Biology, but instead was made available to all students who wanted to participate. As noted earlier, the Biology 210 sequence is a comprehensive survey of Biology generally taken in the sophomore year. It includes genetics, evolutionary biology, biochemistry, molecular biology, physiology, and cell biology. Chemistry 101 is an introductory course taken by freshman as a prerequisite to Biology 210. The course includes general chemistry, general inorganic chemistry and general physical chemistry. Physics 130 (College Physics) is an algebra based physics course primarily for premedical students who do not need to take calculus-based physics. Physics 135 (General Physics) is a calculus-based physics course for science and engineering majors and premedical students. Both courses are taken in the sophomore or junior year, usually after Biology 210.

Subjects

Subjects in the evaluation consisted of all students enrolled in Biology 210, Chemistry 101, Physics 130 (algebra-based) and Physics 135 (calculus based) courses during the 2001–2002 academic year. Students were included in the evaluation whether they were enrolled in a particular course for a single quarter, two quarters or for the entire three quarters of the course sequence.

Workshop Participation

Students were invited to join the GSW program during the first week of the Fall quarter by an in-class announcement made by the GSW Coordinator. The invitation characterized the workshops as advanced, and expressed a strong trust in the ability of the students to benefit from, and contribute to, the intellectual work of the student learning community. To further recruitment goals, minority students received a personal letter describing the program and inviting them to participate.

GSW Program

Program Structure

As in Born's earlier study, the program was modeled after the program pioneered by Treisman (1992). Workshop participants worked together in groups of five to seven students to solve challenging, conceptually rich problems developed by course professors and did not receive academic credit for participation. In contrast to Treisman's original program, workshop groups in the GSW program (and in Born's experiment) met for 2 h per week rather than 4 h per week and were facilitated by undergraduate students who had excelled in the course during the previous year rather than by graduate TA's. The undergraduate peer "facilitators" met with each group to assist and guide the intellectual collaboration necessary for students to solve the problems and develop a deeper understanding of the concepts involved. Prior to each workshop, facilitators met with course professors to discuss and work through the problems themselves. In addition to the weekly training sessions with professors, facilitators received a day long workshop early in the fall quarter which, in addition to explaining their roles and duties, addressed issues of learning theory, learning styles, facilitation, group dynamics. It also provided the facilitators a hands-on opportunity to facilitate problems with each other and receive appropriate feedback.

Short-term goals of the GSW program were to improve performance and retention of science students, with a particular focus on minority students. A longer-term goal was to increase the number of minority students completing degrees in the sciences and pursuing graduate education or discipline-related careers.

Program Implementation

The program was implemented by faculty in conjunction with the Searle Center for Teaching Excellence. The Searle Center provided a workshop coordinator to register workshop students, assign students to workshop groups, train facilitators in pedagogy and to select facilitators for the program. Faculty also played a key role in the implementation of the program in that they developed the workshop problems and conducted the weekly facilitator training sessions on the workshop problems.

Implementation Across Disciplines

The GSW program was implemented in a manner that was generally consistent with the basic structure and aims as described above. However, in order to accommodate the wishes of different departments, there were some potentially important differences in its implementation. These differences fell into four categories: factors affecting participation, time spent in workshop, facilitator training in problem content, and workshop group size. Table I shows the difference between the three disciplines.

Table I. Disciplinary Differences in Program Implementation (2001–2002)

	Biology	Chemistry	Physics
1. Participation	Optional	Optional	Partially required
2. Time in workshop	16 h	12 h	12 h
3. Facilitator training	Course instructor/TA	Course instructor/TA	Other instructor
4. Group size	6–8	4–7	10–12

Factors Affecting Participation. Students received a notation on their academic transcript if they attended the majority of the workshop sessions, but did not receive academic credit. Participation in Biology and Chemistry workshops was optional but encouraged by faculty members. Both participants and nonparticipants attended regular lecture, lab and discussion sessions. In these disciplines, the workshop was offered as an "optional extra" and was not offered as an alternative to any of the course components. In Physics, however, due to departmental reasons, the workshop program was offered as an alternative to the Physics discussion sessions run by departmental teaching assistants. In the regular discussion section, students were required to take quizzes which contributed to their overall grade. Students in the GSW Physics section were not required to take these quizzes. Students were required to enroll in either the regular discussion sections or the "alternative" (GSW) workshop. The choice was up to the students. Numerous anecdotal reports from students indicated that many students regarded the larger, regular—what many described as "chug and plug"—discussion sections as a less appealing option and opted for the GSW workshop to avoid them. This could have resulted in different self selection effects and qualitatively different group interactions in Physics compared with the other disciplines.

Time Spent in Workshop. The amount of time students spent working in their workshop groups differed across disciplines due to factors beyond the control of the GSW program staff. In Biology and Chemistry, there was a total of eight workshop meetings during one 10-week quarter. In Physics, participants met six times. In Biology and Physics, workshop sessions lasted approximately 2 h and in Chemistry, sessions lasted approximately 1.5 h. This resulted in a total of 16 contact hours in Biology, and 12 contact hours in Physics and Chemistry per quarter.

Facilitator Training. In addition to training offered by GSW staff, facilitators met weekly with course instructors (faculty or, in some cases, graduate student TAs) to discuss issues related to the problems that would be used in the following week. Both faculty and the TAs were encouraged to model facilitation for the facilitators in the training session, but were not given specific training in how to train facilitators or on issues such as group dynamics or learning theory. Faculty and the TAs had different styles of organizing meetings. Some lectured about problem content and others used an approach similar to workshop facilitation. In addition, the instructors leading the facilitator training differed across disciplines and quarters. Facilitator training in Biology was conducted by a graduate TA in the Fall quarter and by course instructors in the remaining two quarters. In Chemistry, weekly training was conducted by the course instructors in the fall and winter quarters and by a graduate TA in the spring quarter. In Physics, training in all three quarters was conducted by the course instructors. The instructor, in the fall quarter, however, held only three facilitator training meetings.

Workshop Group Size. Evidence suggests that an appropriate size for small group work of this kind is 4–7 participants in each workshop meeting. In larger groups participation and group cohesion are, for example, noticeably diminished (Bligh, 2000; Jaques, 2000; Light and Cox, 2001). Biology and Chemistry workshops conformed to this standard. However, the demand created by the requirement in Physics, for all students to attend traditional quiz sections or the workshop program, coupled with the general perception that the workshops were more appealing, led to large numbers of students requesting admission to the workshops. In order to accommodate all students, group size was expanded to 10–12 for Physics workshops during fall and winter quarters of 2001–02.

Outcomes

Performance

Student performance was assessed by analyzing final grades for each quarter. Final grades were obtained directly from the grade reports of course instructors. For the purpose of statistical analysis, grades were converted to a numeric, 0–4-point scale. The percentage of students who received grades of D or F was also calculated. Students who dropped from a course part way through the quarter were not included in the analysis because they were not included in course instructors' final grade reports.

Retention

Each discipline has a particular course sequence that extends over the three quarters of the academic year. Ideally, students complete the entire sequence. Students who completed all three courses were considered to be retained in the course. Retention, therefore, represents the percentage of students who began and completed all three quarters (Fall, Winter, and Spring) of the course sequence.

Statistical Analysis

The impact of the program on student performance was assessed using $2 \times 2 \times 2$ analysis of covariance (ANCOVA). Main effects (workshop, sex, ethnicity) and interactions were tested while controlling for preexisting differences in student ability. Chi-square and logistic regression analyses were used to examine the impact of the program on retention. All analyses were conducted with Statistical Package for the Social Sciences (SPSS) 11.5 for Windows.

Covariates

An important difference between the experimental program in Biology evaluated by Born *et al.* (2002) and the GSW program implemented in the 2001–2002 academic year, is that in the 2001–2002 academic year, students simply chose to participate or not participate in the program. The randomized control group was eliminated. This presented a potential benefit in that *all* students who were interested in the program could participate. A potential disadvantage was that self selection of students to workshop and nonworkshop groups may have introduced selection bias. As students were not randomly assigned to workshop and nonworkshop groups it is possible that any differences in performance observed between participants and nonparticipants could be due to preexisting differences between the two groups. For example, workshop students could be more motivated students or more academically gifted students. Alternatively, they could be weaker students who desire help to improve their performance.

Participants and nonparticipants were compared with regard to incoming academic ability as measured by cumulative GPA and/or SAT Math (SAT-M) scores. For sophomores (Biology and Physics), independent samples *t*-tests were performed on cumulative GPA before entering the quarter. The same analysis was conducted on SAT-M for freshmen (Chemistry). In Biology, mean GPA's differed significantly between participants (3.46) and nonparticipants (3.38) ($t_{728} = 3.05$, $p = 0.002$). In Physics, workshop participants entered with a mean GPA of 3.35, while the mean for nonparticipants was 3.26. This was a statistically significant difference ($t_{1112} = 3.41$, $p = .001$). In Chemistry, no differences in SAT-M scores were detected, with means of 696 and 702 respectively ($t_{821} = 0.92$, $p = 0.359$). GPA and SAT-M were significantly and positively correlated with final grade ($r = 0.68$, $p < 0.001$; $r = 0.44$, $p < 0.001$ respectively).

To control for preexisting differences in academic ability, grade point averages (GPA's), calculated before the quarter, or SAT math (SAT-M) scores were used as statistical covariates in both ANCOVA and logistic regression analyses.

Effect Sizes

In order to precisely assess the magnitude of the workshop effects, Cohen's *d* was calculated to measure effect size for performance. Effect sizes can be a useful means of conveying the size of group differences on a common metric, the effect size. Effect sizes for retention were calculated using the formula developed by Shadish *et al.* (1999). This formula uses the odds ratio to produce a value that is directly comparable to a standardized mean difference (Klein, 2004).

Verification of ANCOVA Assumptions

Assumptions of ANCOVA were verified. The dependent variable, final grade, was approximately normally distributed. Final grade variances for workshop and nonworkshop groups were not statistically significantly different. Final grade scores were not correlated with each other. ANCOVA also assumes homogeneity of regression (i.e., that the correlation between the dependent variable (final grade) and the covariate (GPA/SAT-M) is equal for all levels of the independent variable). This assumption was tested by including interaction terms between workshop group and the covariate (GPA or SAT-M) in the ANCOVA models for each discipline and quarter. None of the interaction terms were statistically significant indicating that the assumption for homogeneity of regression was met in each discipline for both GPA and SAT-M covariates.

Treatment of Missing Data

Approximately 10% of students in each discipline chose not to disclose their ethnicity. Therefore, these students were excluded from analyses involving the variable for ethnicity. There was no statistically significant difference between the proportion of workshop participants (8.6%) and nonparticipants (7.8%) who chose not disclose their ethnicity ($\chi^2 = 0.242, p = 0.652$).

RESULTS

Workshop Participation

Class enrollment and workshop participation by discipline and quarter are summarized in Table II.

Table II. Class Enrollment and Workshop Participation by Discipline, Quarter and Ethnicity

	Fall'01	Winter'02	Spring'02
Biology			
Enrolled in class	261	270	206
Enrolled in workshop	126 (48.3%)	111 (41.1%)	90 (43.7%)
Minority	10 (45.5%)	7 (41.2%)	6 (50.0%)
Majority	107 (48.9%)	96 (41.7%)	78 (44.1%)
Chemistry			
Enrolled in class	527	456	308
Enrolled in workshop	137 (26.0%)	69 (15.1%)	53 (17.2%)
Minority	22 (27.5%)	10 (16.9%)	8 (20.5%)
Majority	102 (25.2%)	55 (15.1%)	39 (15.9%)
Physics 130			
Enrolled in class	138	141	118
Enrolled in workshop	59 (40.4%)	68 (45.0%)	48 (33.6%)
Minority	9 (50.0%)	9 (42.9%)	4 (23.5%)
Majority	45 (37.5%)	54 (45.0%)	40 (33.9%)
Physics 135			
Enrolled in class	161	235	272
Enrolled in workshop	109 (60.6%)	131 (51.8%)	116 (38.8%)
Minority	8 (57.1%)	6 (35.3%)	4 (21.1%)
Majority	91 (61.9%)	115 (52.8%)	101 (39.9%)

Note. Percentage (%) denotes percentage of group. (Minority and majority numbers do not sum to the total number of students enrolled in the workshop as some students did not disclose their ethnicity.)

Table III. Workshop Enrollment by Discipline, Quarter, and Sex

	Fall'01	Winter'02	Spring'02
Biology			
Males	47 (38.2%)	43 (30.9%)	35 (33.0%)
Females	79 (57.2%)	68 (51.9%)	55 (55.0%)
Chemistry			
Males	47 (18.1%)	23 (9.7%)	19 (14.2%)
Females	90 (33.6%)	46 (21.0%)	34 (19.5%)
Physics 130			
Males	19 (37.3%)	24 (49.0%)	19 (37.3%)
Females	40 (42.4%)	44 (43.1%)	29 (35.1%)
Physics 135			
Males	61 (56.0%)	68 (44.7%)	64 (34.4%)
Females	48 (67.6%)	63 (62.4%)	52 (46.0%)

Note. Percentage (%) denotes percentage of group enrolled in the workshop.

Participation was higher in Biology and Physics than in Chemistry, with overall participation rates of approximately 45%. Lower participation of students in Chemistry was not due to a lack of demand for the program, but rather was limited by the number of available facilitators. Minority students (African American, Hispanic, and Native American) generally participated at rates equivalent to majority students (White and Asian/Pacific Islanders). However, the total number of minority students who participated in the program was small because of the relatively small number of minority students enrolled in each discipline. Females participated in the program at higher rates than males in Biology and Chemistry (Table III). However, participation rates for the two groups were virtually identical in Physics.

Performance

Biology

In Biology, workshop participants earned higher final grades than nonparticipants in Fall and Winter quarters (Fig. 1). In the Fall quarter, ANCOVA analysis revealed a main effect for workshop ($F_{1,225} = 10.08$, $p = 0.002$) and a borderline interaction between workshop and ethnicity ($F_{1,225} = 3.24$, df $= 1$, $p = 0.073$). Adjusted mean final grades for participants was 0.27 points higher than that of nonparticipants (95% CI 0.07–0.47, $d = 0.45$). Follow-up tests suggested that minority participants may have benefited more from the workshop than majority participants. Adjusted mean final grades for minority participants were 0.44 points higher than those of minority nonparticipants (95% CI –0.03–0.91;

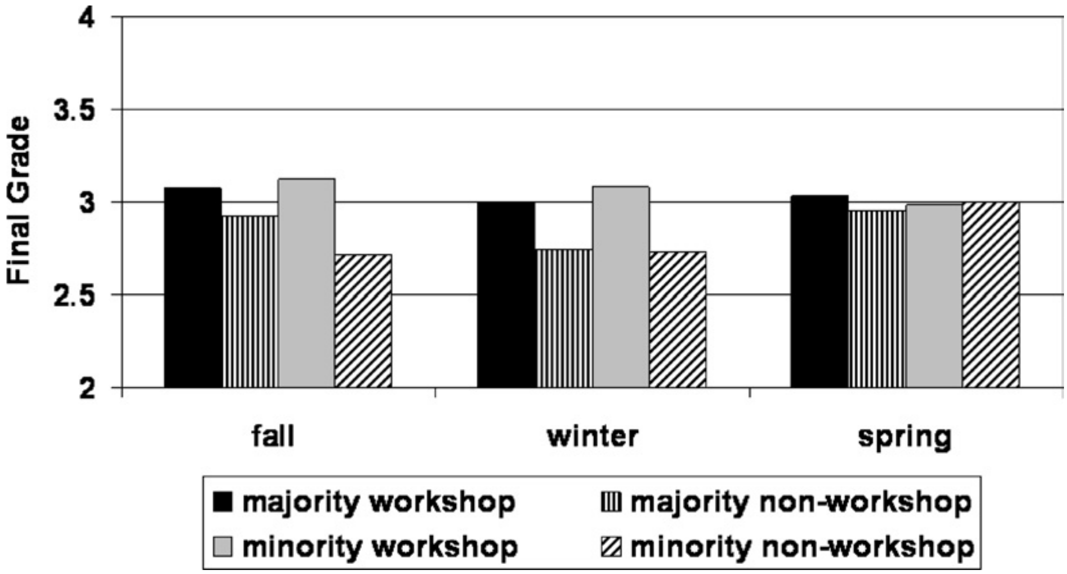

Fig. 1. Final Biology grade by quarter and ethnicity (2001–2002).

$d = 0.83$). This difference approached statistical significance ($F_{1,17} = 3.83, p = 0.067$). Adjusted mean final grade for majority participants was 0.15 points higher than that of majority nonparticipants (95% CI 0.03–0.28; $d = 0.34$). As indicated by the 95% confidence interval, this difference reached statistical significance at the $p < 0.05$ level. A statistically significant main effect for workshop was found in the Winter quarter ($F_{1,225} = 5.43$ $p = 0.021$; $d = 0.40$). Adjusted mean final grade for workshop participants was 0.28 points higher than that of nonparticipants (95% CI 0.02–0.55). The Workshop × Ethnicity interaction did not reach statistical significance in the winter quarter ($F_{1,225} = 0.311$, $p = 0.577$). This was probably due to the small number of minority students in the sample. However, the effect size seen in minority students ($d = 0.63$) was larger than that seen in majority students ($d = 0.47$). There was no main effect for workshop and no interaction between workshop and ethnicity in the Spring quarter in Biology ($F_{1,181} = 0.90, p = 0.344$; $F_{1,181} = 0.317, p = 0.57$ respectively).

No minority workshop students earned D's or F's in Biology (Table IV). However, very few minority students in Biology earned D's or F's and differences between workshop and nonworkshop groups were not statistically significant in any of the quarters. (Fall: Fisher's Exact one-sided $p = 0.545$; Winter Fisher's Exact one-sided $p = 0.500$). No students in either the workshop or nonworkshop groups received D's or F's in Biology in the Spring quarter.

Chemistry

In Chemistry, workshop participants received higher adjusted mean final grades than nonparticipants in Winter and Spring quarters (Fig. 2). In the Winter quarter, there was a statistically significant main effect for workshop ($F_{1,406} = 5.09$, $p = 0.025$) and a statistically significant interaction between workshop and ethnicity ($F_{1,406} = 3.83$, $p = 0.051$). Adjusted mean final grade of workshop participants was approximately one third of a grade higher than that of nonparticipants (95% CI 0.04–0.62). Adjusted final grades of minority participants averaged 0.62 points higher than those of minority nonparticipants

Table IV. Percentage of Minority Students Earning Grades of D or F by Discipline and Quarter

Quarter	Biology	Chemistry	Physics 130	Physics 135
Fall				
Participants	0%	18.2%	22.2%	25.0%
Nonparticipants	8.3%	32.8%	25.0%	33.3%
Winter				
Participants	0%	10%	22.2%	0%
Nonparticipants	14.3%	24.5%	11.1%	9.1%
Spring				
Participants	0%	0%	0%	3.0%
Nonparticipants	0%	19.4%	0%	2.0%

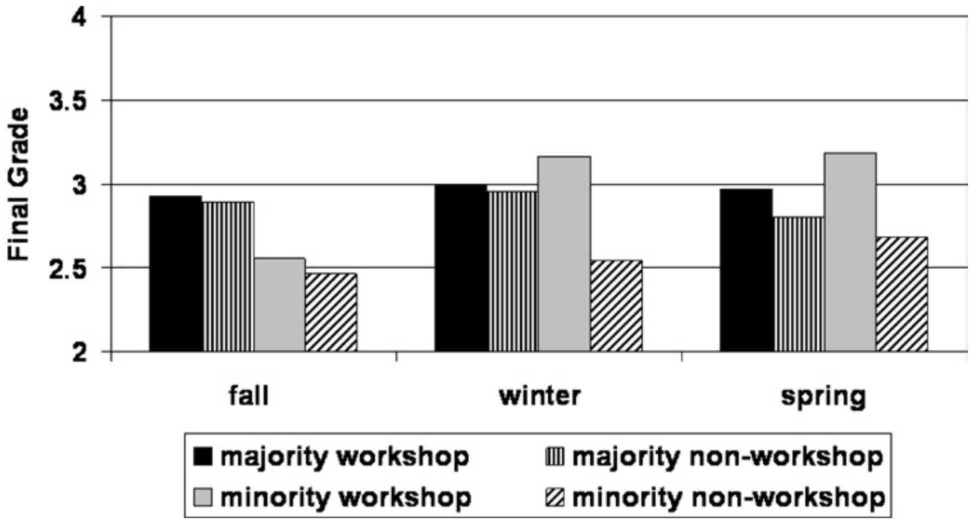

Fig. 2. Final Chemistry grade by quarter and ethnicity (2001–2002).

(95%CI −0.06–1.15; $d = 0.81$). A very small difference of 0.04 grade points was seen between majority participants and nonparticipants (95%CI −0.15–0.23; $d = 0.06$). In the Spring quarter, a borderline main effect for workshop was observed with an advantage in mean final grade seen in workshop participants ($F_{1,276} = 3.75$, $p = 0.54$). Adjusted mean final grade of participants was 0.33 points higher than that of nonparticipants (95% CI −0.05–0.68). The interaction between workshop and ethnicity was not statistically significant. However, once again differences in adjusted final grade between participants and nonparticipants were larger for minority students than majority students. The effect size for minority students was 0.66 compared with 0.28 for majority students. This suggests an interaction between workshop and ethnicity. However, the interaction probably failed to reach significance because the number of minority workshop participants was small ($n = 8$) and the observed statistical power was low (0.16).

Larger numbers of students earned D's and F's in Chemistry than in Biology. However, the percentage of minority students who earned D's or F's was smaller for workshop students than nonworkshop students (Table IV). This difference approached statistical significance in the Fall quarter (Fisher's Exact one-sided $p = 0.156$). On average, the proportion of minority workshop students who earned D's or F's was half that of nonworkshop students. For example, in the Fall 2001 quarter, 18.2% of minority workshop students earned D's or F's in contrast to 32.8% of minority nonworkshop students. There were no statistically significant differences in the percentage of majority workshop and nonworkshop students who earned D's or F's across the three quarters. On average, 1–2% of majority workshop students earned D's or F's compared with 3–4% of majority nonworkshop students.

Physics

In Physics 130, the algebra based course, mean final grades of students who did not participate in the program were slightly higher than those of workshop participants. However, there were no statistically significant differences between workshop participants and nonparticipants and no significant Workshop × Ethnicity interactions in any of the quarters (Fig. 3). In Physics 135, the calculus based course, workshop participants had higher adjusted mean final grades than nonparticipants in only the Fall quarter (Fig. 4). In the Fall quarter, the difference approached statistical significance ($F_{1,101} = 3.24$, $p = 0.075$). The adjusted mean final grade for workshop participants was 0.37 points higher than that of nonparticipants (95% CI −0.04–0.78; $d = 0.44$). The Workshop × Ethnicity interaction was not statistically significant ($F_{1,101} = 1.97$, $p = 0.163$). Once again, this was probably due to the small total number of minority students in Physics 135 ($n = 13$). As with Biology and Chemistry, the difference

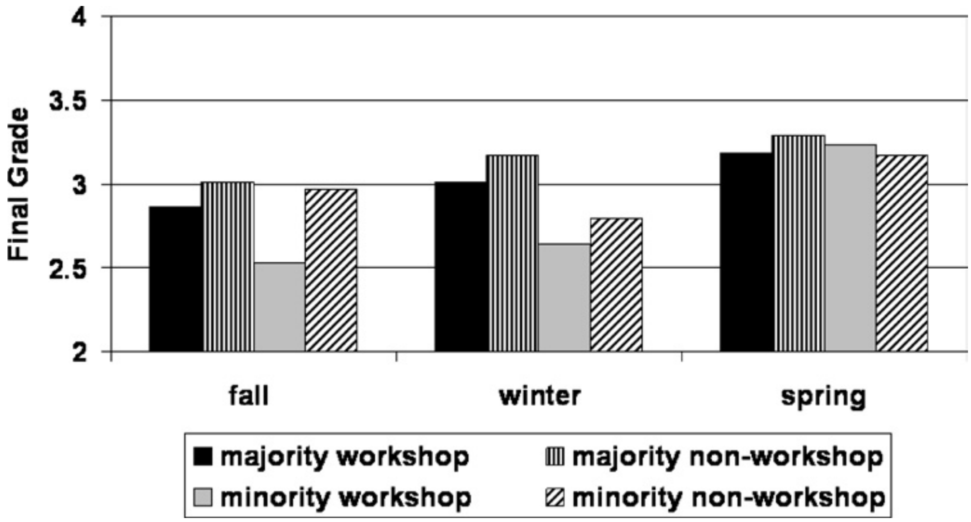

Fig. 3. Final Physics 130 grade by quarter and ethnicity (2001–2002).

between workshop participants and nonparticipants was larger for minority students ($d = 0.14$) than for majority students ($d = 0.10$). There were no statistically significant main effects for workshop or Workshop × Ethnicity interactions in Winter or Spring quarters.

Retention

Retention data are summarized in Fig. 5. Chi-square analyses were used to determine if there was an association between workshop participation and retention in the course sequence. In Biology, workshop participants were more likely to complete the sequence than nonparticipants ($\chi^2 = 33.9$, $p < 0.001$). Minority participants were more likely to complete the sequence than minority nonparticipants ($\chi^2 = 4.46$, $p < 0.05$). Both effects were in the predicted direction. Effects sizes for retention in Biology were computed using the formula by Shadish *et al.* (1999) and were 0.85 for minority students and 0.46 for majority students.

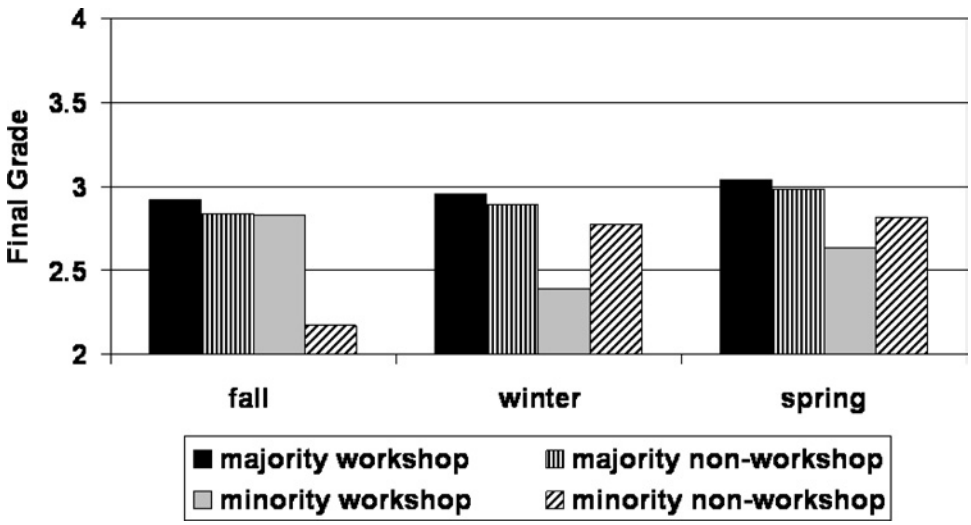

Fig. 4. Final Physics 135 grade by quarter and ethnicity (2001–2002).

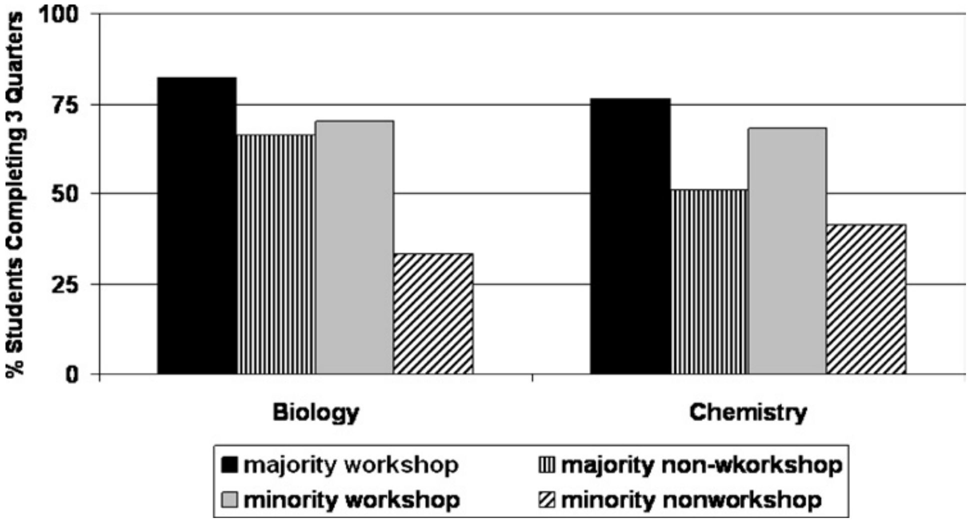

Fig. 5. Student retention over three quarter by discipline and ethnicity (2001–2002).

A similar pattern of results occurred in Chemistry, with workshop participants significantly more likely to be retained as well ($\chi^2 = 22.69$, $p < 0.001$). Minority student participants were more likely to be retained than minority nonparticipants ($\chi^2 = 4.59$, $p < 0.05$) Majority participants were also more likely to be retained than majority nonparticipants ($\chi^2 = 4.59$, $p < 0.05$). Effect sizes were similar for minority students (0.62) and majority students (0.64). Because of departmental requirements, a large number of students take only one or two courses in Physics 130 and 135 sequences, making analysis of retention inappropriate for the discipline.

In order to address the possibility that retention resulted from selection bias, rather than workshop participation, we conducted a logistic regression analysis with retention status serving as a dependent variable, and entered workshop status, sex, ethnicity, and grade point average/SAT-M as predictors. This approach is similar to using a covariate in an analysis of covariance in that we can determine if workshop participation predicts retention above and beyond the contributions of other variables, grade point average/SAT-M in particular.

Logistic regression analysis was conducted separately for each discipline with all variables entered simultaneously into the model. For Biology, after adjusting for sex, ethnicity, and grade point average, the odds ratio for workshop participation was 2.62, ($p < 0.01$), (95% CI = 1.37–5.03), indicating that after adjusting for GPA, workshop participants were 2.6 times more likely to be retained than nonparticipants.

In Chemistry, SAT-M was used as an independent variable in addition to sex and ethnicity. An odds ratio of 2.74 ($p < 0.01$), (95% CI = 1.57–4.76) indicated an advantage in retention similar to the advantage seen in workshop students in Biology.

DISCUSSION

The GSW program was developed to address poor performance and attrition of minority students in introductory Biology, Chemistry and Physics courses at Northwestern University. It was first implemented as a tightly controlled experiment from 1997–1999 by Born *et al.* (2002). In response to the positive results of Born *et al.*'s experiment, the GSW program was implemented as a regular academic program at Northwestern in three different departments in the Fall of 2001. This paper presents results from the first year of the program (2001–2002 year) when the GSW program was implemented as a nonexperimental program under "real world" conditions in introductory biology, chemistry and physics courses. The program had a positive impact in Biology and Chemistry. No advantage to workshop students was seen in Physics. However, this may have been due to less than optimal implementation of the program in Physics.

Performance

In Biology, after adjustment for prior GPA, adjusted mean final grades of minority workshop participants were 0.4 grade points higher than those of nonparticipants in both Fall and Winter quarters. Effect sizes for minority students ranged between 0.63 and 0.83. In the case of majority students, adjusted mean final grades of workshop participants were 0.2 grade points higher than those of nonparticipants. Effect sizes were smaller, ranging from 0.34–0.47.

A virtually identical pattern of results was observed in Chemistry. Mean final grades of participants were significantly higher than those of nonparticipants in Winter and Spring quarters. On average, adjusted mean final grades of minority participants were 0.5 grade points higher than those of minority nonparticipants. On average, adjusted mean final grades of majority participants were 0.1 grade points higher than those of majority nonparticipants. Effect sizes for minority students (0.66–0.81) were larger than those for majority students (0.06–0.28).

Program results in both algebra based (130) and calculus based (135) Physics courses were disappointing. There were no statistically significant differences between participants and nonparticipants in Physics 130 in any of the three quarters, and nonparticipants received slightly higher grades than participants. In Physics 135, participants received statistically significantly higher grades than nonparticipants in only the Fall quarter.

Disappointing results in the two Physics courses may be related to program differences between Physics and the other two disciplines. As noted earlier, workshop groups were larger in Physics (10–12) than in Biology (6–8) and Chemistry (4–7). The relation between group size and learning is complex and related to task characteristics. A close look at the literature on problem solving and learning in groups, however, yields some useful general conclusions. While larger groups may have more resources available to bring to bear on the problem itself, participation by all members becomes unlikely, therefore reducing the impact of the experience across participants (Bligh, 2000). What was originally an admirable goal of providing more students with the opportunity to participate in the workshop program, may have diluted the very effects of participation sought. This explanation is consistent with the classic work of Bales *et al.* (1951) in which the authors demonstrated systematic reductions in group member contributions with increasing group size.

The circumstances under which participants are enrolled in Physics may also be an important factor in explaining the pattern of results in Physics. As discussed previously, when random selection does not exist, factors that are related to the desire to participate, such as high motivation or deep interest may influence the results. The fact that Physics students were required to enroll in the either the GSW or another option potentially results in a different population of participating students than those who choose to participate for personal reasons as students in Biology and Chemistry did.

Retention

The workshop program appears to have had a positive impact on retention of both minority and majority students. In Biology and Chemistry, workshop participants were significantly more likely to complete the entire three-course sequence than nonparticipants. This pattern held true whether or not ability was controlled. As with performance, minority students appeared to derive even greater benefit from the program than majority students. Minority participants were more likely to be retained than minority nonparticipants. More importantly, retention rates of minority students approached those of both majority participants and majority nonparticipants. These findings are particularly promising because they may point to a specific benefit of program participation for students traditionally under-represented in the sciences.

Several aspects of the GSW program may have lead to enhanced retention of minority participants. Analysis of course grades indicates that the workshop appears to have boosted the academic performance of minority participants which may in turn have prevented them from dropping the course because of low grades. However, increased grades may not be the only factor to positively affect retention. Several studies suggest that retention of minority students is not related exclusively to academic factors. In their study of minority students at University of California Berkley, a predominantly white university, Loo and Rolison (1986) found that retention of minority students was influenced by both academic and sociocultural factors (i.e. social isolation, social integration within the university). Participation of minority students in ethnically diverse workshop groups may have promoted their integration within the student body and reduced social isolation.

Grandy (1998) found that another nonacademic factor, termed "minority support", had the greatest effect on degree completion of high-ability minority students in the sciences. In Grandy's study, minority support was defined as the availability of minority role models and advisors, availability of advice and support from more advanced minority students and the availability of a dedicated minority relations staff. There were two minority facilitators in the GSW program (one in biology and one in chemistry). Although these numbers are small, the presence and involvement of minority students as peer facilitators in the program may have contributed to the persistence of minority students. Interestingly, in Grandy's study, the primary effect of minority support was not on grades, but on domains such as science ambition, attitudes, enjoyment and willingness to make a career commitment, all of which were significant predictors of persistence in the sciences. Whilst we do not have direct measures of interest and motivation of students who participated in the GSW program, we do have their responses to a questionnaire designed to measure their satisfaction with and enjoyment of the workshop program. Student satisfaction and enjoyment were high in all three disciplines. This may have contributed to general satisfaction with and enjoyment of their courses.

Results of the first GSW program in Biology and Chemistry are consistent with those of other workshop programs designed to improve performance and retention of minority students (*Professional Development Program Mathematics Workshop* (Fullilove and Treisman, 1990); *Mathexcel* (Freeman, 1995); *Academic Excellence Workshop Program* (Bonsangue and Drew, 1995); *Peer-Led Team Learning (PLTL)* (Tien *et al.*, 2002). This suggests that workshop type programs may be effective in a variety of scientific disciplines and institutional settings.

One of the most striking features of the GSW program evaluation is that the program appears to have had a greater impact on minority students than majority students. This may have occurred because the program may have affected factors that are more influential for performance and retention of minority students than majority students. Work by Treisman (1985, 1990, 1992) suggests that poor performance of minority students is due in part to academic isolation. The GSW program may have reduced academic isolation in minority students and thus lead to an improvement in their performance and retention. However, the program may have had less of an impact on majority students because they are not generally academically isolated.

Tracey and Sedlacek (1984, 1985, 1987), Loo and Rolison (1986), and Grandy (1998) have demonstrated that factors that influence performance and retention of minority students are different from the factors that influence performance and retention of majority students. Nonacademic factors such as enjoyment of the course, commitment to the discipline as a career, perceived quality of instruction and integration within the institution are associated with retention of minority students but are not associated with retention of majority students. Participation in small ethnically diverse groups, mentoring by advanced peers and a focus on conceptual learning may have led to increases in these nonacademic variables. Again, because these factors influence performance and retention of minority, but not majority students, minority students may have benefited more from the GSW program than majority students.

Minority students may have derived greater benefit from the GSW program than majority students because the program helped to minimize barriers to academic performance that are present for minority students but not for majority students. Stereotype threat has been suggested as an explanation for poor academic performance of minority students (Steele, 1997). The GSW program may have helped to reduce stereotype threat in minority students in two ways. First, the program is honorific. That is, it is not remedial in any way: indeed, it is designed to challenge students by engaging them in problems specifically designed to be more difficult than the regular homework problems. Participation is intended as an honor, and this quality may have conveyed to minority students that the stereotype of minority students being poor academic performers in need of remedial assistance was not being applied to them at Northwestern. Second, small group work may have facilitated friendships between minority and majority students. Steele *et al.* (2002) suggested that formation of friendships between minority and majority students increases the comfort of minority students and may give minority students confidence that they are not being devalued and stereotyped. Interestingly, Steele *et al.* (2002) cite studies that have found an association between number of white friends and college GPA in minority students (e.g. Graham *et al.*, 1984).

Limitations

The results of this study are encouraging, but must be considered in light of the study's methodological limitations. A major limitation of this study is the self-selection of students to workshop and nonworkshop groups. As a randomized study design was not used, it is impossible to definitively attribute gains made by GSW participants to their participation in the program. For example, students who chose to participate in the program may have been brighter or more motivated than students who did not chose to participate. The fact that all study results held after adjustment for prior academic ability, suggests that differences observed between participants and nonparticipants were not due to differences in incoming level of academic ability as measured by prior cumulative GPA and SAT-M. However, it is still possible that advantages in performance and retention of workshop students are due to differences in academic preparation, interest, motivation and/or other unknown personal characteristics that were not measured in the study. To overcome this limitation, interest, motivation and anxiety will be measured in future studies and used as covariates in statistical analyses.

Whilst the present study may have been subject to selection bias, it is also important to note that Born *et al.*'s (2002) randomized evaluation of the Honors Workshop program provides evidence that a program such as the GSW program can independently impact performance. (That is, the positive effects of the program are independent of selection bias.) In Born *et al.*'s study, students who wanted to participate in the workshop program but were randomized to the control group, earned lower grades than students who wanted to participate and were randomized to the program. This comparison holds constant one source of bias, namely the desire to participate. In addition, all unknown factors that accompany this desire are controlled through the use of a randomized design.

However, implementation of programs such as these in a real world setting, without experimental controls and with the difficulties of coordination and administration, represents an important next step in evaluating the strength of an intervention. While random assignment offers methodological power and stronger evidence for causal inferences, it falls short in illuminating program effects in naturalistic settings, where students pick and choose campus activities in which to participate.

CONCLUSION

The GSW program is a low intensity intervention in which undergraduate students meet with an advanced peer and work together in small groups of 4–7 students for 2 h per week outside of class to solve challenging, conceptually based problems designed by faculty. It was first implemented as a tightly controlled experiment in an introductory Biology course by Born *et al.* (2002) and was shown to be effective in enhancing academic performance of both minority and majority students.

Results of the present evaluation suggest that the GSW program can be implemented successfully in a nonexperimental setting by a trained program coordinator with the support of the faculty at a highly selective university, where the program is not tightly controlled by educational researchers. The program had a positive impact on performance and retention of both minority and majority students. Larger effects on both final grades and retention were observed in minority students (0.5–0.6 standard deviation for grade) than in majority students (0.1–0.3 standard deviations for grade), suggesting that minority students may derive particular benefit from the program. The GSW program is a relatively simple strategy that faculty in introductory biology and chemistry courses may adopt to reduce gaps in academic performance and retention between minority and majority students. Continuation of the GSW program in Biology, Chemistry and Physics to the present (2005) and extension of the program to Engineering Analysis, Introductory Calculus for nonmajors and Organic Chemistry course sequences at Northwestern suggests that faculty are quite willing to adopt such a program. In addition, high student demand for the program and the oversupply of applicants for facilitator positions in all disciplines suggest that students are also very willing to adopt the program.

Finally, although, minority students appear to be the primary beneficiaries of the program, we feel that the gains made by majority students, particularly in the area of course retention, together with their enjoyment of and demand for the program fully justify their inclusion in the program. Further, the literature on stereotype threat and the literature on achievement and retention of minority students suggest that inclusion of both minority and majority students in the program may be integral to its success. For these reasons, we believe that if a GSW program is implemented, it should definitely be made available to both minority and majority students.

REFERENCES

American Association for the Advancement of Science (1989). Science for all Americans: Project 2061. Retrieved May 6, 2003, from http://www.project2061.org/tools/sfaaol/chap13.htm

Bales, R. F., Strodtbeck, F. L., Mills, T., and Rosenborough, T. M. (1951). Channels of communication in small groups. *American Journal of Sociological Review* 16: 461–468.

Bligh, D. (2000). *What's the Point in Discussion?* Intellect, Exeter, UK.

Bonsangue, M. V., and Drew, D. E. (1995). Increasing minority students success in calculus. *New Directions for Teaching and Learning: Fostering Student Success in Quantitative Gateway Courses* 61: 23–33.

Born, W., Revelle, W., and Pinto, L. (2002). Improving biology performance with workshop groups. *Journal of Science Education and Technology* 11: 347–365.

Chavez, M., and Maestas-Flores, M. (1991). Minority student retention: ENLACE. *New Directions for Community Colleges* 74: 63–67.

Clewell, B. C., and Ficklen, M. S. (1987). Effective institutional practices for improving minority retention in higher education. *Journal of College Admissions* 116: 7–13.

Croizet, J., Desert, M., Dutrevis, M., and Leylens, J. (2001). Stereotype threat, social class, gender, and academic underachievement: When our reputation Catches up to us and takes over. *Social Psychology of Education* 4: 295–310.

College Entrance Examination Board (1999). *Reaching the Top: A Report of the National Task Force on Minority High Achievement* (Item No. 201635), College Entrance Examination Board, New York.

Fullilove, R. E., and Treisman, P. U. (1990). Mathematics achievement among African American undergraduates at the University of California, Berkley: An evaluation of the Mathematics Workshop Program. *Journal of Negro Education* 59(3): 49–64.

Freeman, M. (1995). MathExcel: A special opportunity in calculus. Unpublished Report, Department of Mathematics, University of Kentucky.

Gosser, D., Roth, V., Gafney, L., Kampmeier, J., Strozak, V., Varna-Nelson, P., Radel, S., and Weiner, M. (1996). Workshop Chemistry: Overcoming barriers to student success. *The Chemical Educator* 1: 2–3.

Graham, C., Baker, R. W., and Wapner, S. (1984). Prior interracial experience and Black student transition into predominantly white colleges. *Journal of Personality and Social Psychology* 47: 1146–1154.

Grandy, J. (1998). Persistence in science of high ability minority students: Results of a longitudinal study. *The Journal of Higher Education* 69: 589–620.

Grayson, J. P. (1998). Racial origin and student retention in a Canadian university. *Higher Education* 36: 323–352.

Jaques, D. (2000). *Learning in Groups—A Handbook for Improving Group Work*, Kogan Page, London.

Klein, R. B. (2004). *Beyond Significance Testing: Reforming Data Analysis Methods in Behavioral Research*, American Psychological Association, Washington, DC.

Levin, M. E., and Levin, J. R. (1991). A critical examination of academic retention programs for at-risk minority college students. *Journal of College Student Development* 31: 323–334.

Light, G., and Cox, R. (2001). *Learning and Teaching in Higher Education—The Reflective Professional*, Paul Chapman, London.

Loo, C. M., and Rolison, G. (1986). Alienation of ethnic minority students at a predominantly white university. *Journal of Higher Education* 57: 58–77.

National Science Foundation, Division of Science Resources Statistics (2004). *Women, Minorities and Persons with Disabilities in Science and Engineering: 2004*, NSF 04-317, Arlington, VA.

Selvin, P. (1992). Math education: Multiplying the meager numbers. *Science* 258: 1200–1201.

Seymor, E., and Hewitt, N. M. (1997). *Talking About Leaving: Why Undergraduates Leave the Sciences*. Westview Press, Boulder, CO.

Shadish, W. R., Robinson, L., and Lu, C. (1999). *ES: A Computer. Program for Effect Size Calculation,* St. Paul, MN: Assessment Systems.

Springer, L., Stanne, M. E., and Donovan, S. (1998). *Effects of cooperative learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis* (Research Monograph No. 11), University of Wisconsin-Madison, National Institute for Science Education, Madison.

Steele, C. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologis* 52: 613–629.

Steele, C. M., Spencer, S. J., and Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In Zanna, M. (Ed.), *Advances in Experimental Social Psychology* (Vol. 23), Academic Press, New York, pp. 379–440.

Tien, L. T., Roth, V., and Kampmeier, J. A. (2002). Implementation of a peer-led learning instructional approach in an undergraduate organic chemistry course. *Journal of Research in Science Teaching* 39: 606–632.

Tracey, T. J., and Sedlacek, W. E. (1984). Noncognitive variables in predicting academic success by race. *Measurement and Evaluation in Guidance* 16: 171–178.

Tracey, T. J., and Sedlacek, W. E. (1985). The relationship of noncognitive variables to academic success. A longitudinal comparison by race. *Journal of College Student Personnel* 26: 405–410.

Tracey, T. J., and Sedlacek, W. E. (1987). Prediction of college graduation using noncognitive variables by race. *Measurement and Evaluation in Guidance* 19: 177–184.

Treisman, P. M. (1985). *A Study of the Mathematics Performance of Black Students at the University of California, Berkley*. Unpublished doctoral dissertation, University of California, Berkley.

Treisman, P. V. (1990). Studying students studying calculas: A look at the lives of minority mathematics students in college. *The College Mathematics Journal* 23(5): 362–372.

Treisman, U. (1992). Studying students studying calculus: A look at the lives of minority mathematics students in college. *The College Mathematics Journal* 23: 362–372.

Vars, F. E., and Bowen, W. G. (1998). Scholastic aptitude test scores, race and academic performance in selective colleges and universities. In Jencks, C., and Meredit, P. (Eds.), *The Black-White Test Score Gap.* Brookings Institute, Washington, DC, pp. 457–479.