The GROW Project: Impacts on Academic Performance

Jacqueline D. Spears, Beth A. Montelone, Susan Arnold Christian, Kimberly D. Douglas, Ruth A. Dyer, and Lisa Freeman

Kansas State University

Abstract – The Girls Researching Our World (GROW) Project at Kansas State University has focused for the past five years on encouraging middle school girls to pursue their interests in science, engineering and technology. We offer educational outreach opportunities including summer workshops, industry tours, and presentations to middle school girls throughout the state of Kansas. To date, more than 500 sixth, seventh, and eighth grade girls have participated in these activities. In 2002, the GROW Project entered into a partnership with three local school districts to track girls who participated in GROW-sponsored events. The goal was to explore whether participation in GROW activities resulted in a measurable impact on enrollment and performance by these girls in required and elective science and math courses. The schools agreed to collect enrollment data as well as the grades the girls received in these courses over a three-year period. The schools also provided the same information for a control group of girls who have never participated in GROW-sponsored activities. This paper discusses some preliminary results from one cohort of students in this data set. A regression analysis suggests that a small but significant proportion of the variance in girls' performance in seventh and eighth grade science can be linked to their participation in GROW activities. A qualitative analysis of the number of science courses taken in high school and the relative difficulty of those courses suggests that there are no significant differences in the number of science courses taken by girls who do or do not participate in GROW activities, but that those who participate in GROW activities take more challenging courses. We discuss these findings in terms of the challenges of using such measures to assess the effectiveness of outreach programs and to provide suggestions to other program directors about the issues involved in carrying out such assessment.

Introduction

A team of faculty members and administrators at Kansas State University (K-State) began a concerted effort in 1999 to encourage middle school girls to pursue an educational path in science, technology, engineering and mathematics (STEM) with the long term goal of increasing the numbers of women students and faculty in STEM at K-State and women in the STEM workforce. The need to increase participation of historically underrepresented groups in STEM has been widely recognized (*e. g.*, CAWMSET, 2000) and has been the focus of multiple projects funded by the National Science Foundation (NSF) and the American Association of University Women, synthesized in the report *Under the Microscope* (Dyer, 2004). Two awards from the NSF Program for Gender Equity (now Research on Gender in Science and Engineering) provided funding to establish the Girls Researching Our World (GROW) Project at K-State. The GROW Project works with partners from K-12 school districts, local STEM industries, governmental agencies, and community and non-profit organizations to conduct a series of outreach activities intended to foster the interests of girls in STEM. We currently offer an annual three-day GROW Workshop in June, industry tours during the school year, and a series of shorter (evening or half-day) events during the school year. Details about the format and content of the summer workshops and industry tours, the nature of the network of partners created to support the project, and the initial assessment efforts conducted in 2000 and 2001 have been reported elsewhere (Franks et al., 2002, Arnold et al., 2003, Arnold et al., 2004, Spears et al., 2004).

One of the most challenging aspects of running a program of this type is determining its effectiveness and long-term impact on the participants. Stakeholders invested in the program and intensely interested in its outcomes include participants, their parents, their teachers, school districts, the program directors, and, of course, the funding agency supporting the program. Although it is straightforward to gather information about attitudes and interests of participants and reactions to particular program events, it is much more difficult to assess the longer-term results that we hope to see from such a program. Given the current emphasis on 'scientifically-based research' in many federal agencies, a realistic examination of what is and is not possible seems warranted. In general, issues arise with regard to reasonable expectations, relative to the investment of funds, and measurement challenges.

Reasonable expectations generally depend on issues related to the size of the effect and the number of individuals in the population. It is important to realize that over more than fifty years, most quasi-experimental studies in education have resulted in the retention of the null hypothesis – most interventions did not have measurable effects. Although there are often problems in the ability of researchers to meet randomization requirements, the more serious problems lie with interventions that are relatively short and populations that are relatively small, both of which often arise from the funding level available to support programs. It is important that assessment efforts choose expectations that are commensurate with the length of an intervention (lower expectations for interventions that involve individuals for only a few days and higher expectations for interventions that extend over a year or so).

Given that reasonable expectations are chosen, the second challenge is one of measurement. Longer-term outcomes are simply more difficult to assess because populations are fluid. Students move in and out of school districts, decreasing the total number of participants for whom longitudinal data can realistically be gathered. Comparisons based on grades or courses taken are also problematic because of the variability across school districts. At the lower grade levels, all students take the same science and mathematics classes, but there is often little variability in the grades assigned. (Schools are also struggling with issues of grade inflation.) At higher grade levels, more variability typically exists in both the grades assigned and the level of difficulty of the courses taken. Thus, there are a number of different comparisons that can be made, each of which has its own limitations. In short, dealing with school-based data is messy and rarely results in the kinds of outcomes many of us are accustomed to in science and engineering.

We were funded by NSF for a three-year demonstration project beginning in 2001. The assessment design of this proposal included commitments from three partner school districts to track the progress of girls from their districts who took part in GROW activities. These longitudinal data included course taking patterns and grades received in science and mathematics courses. They also agreed to provide control group data on a similar group of girls from their districts who had not participated in GROW events. Analysis of these data allows us to test the hypothesis that participation in GROW events would result in the girls choosing to take more advanced science and math in high school as well as achieving better science and math grades than the control group of girls. A preliminary analysis of some of these data illustrates many of the issues mentioned above that are related to attempting to carry out 'scientifically based research' in education.

Study Design and Data Collection

In the process of preparing our 2001 grant proposal, we met with administrators from three local school districts, Unified School District (USD) 305, Salina, KS; USD 383, Manhattan-Ogden, KS; and USD 475, Geary County, KS. These districts were chosen because they were those from whom we anticipated recruiting many of our student participants. They also represent different types of student populations. USD 305 is a small city district with 25.5% of the students from underrepresented groups and 45.8% receiving free or reduced lunches. USD 383 is in Manhattan, a university community; 22.7% of its students are from underrepresented groups and 31.2% on free or reduced lunches. USD 475 serves Junction City and Fort Riley and has a much higher minority population (47.3%) and 54.3% of its students receive free or reduced lunches. Demographic and socioeconomic data are for the 2004-2005 school year and were obtained from the Kansas State Board of Education website (KSBE, 2005).

After the grant was funded, we met with administrators from each school district. These meetings helped us set the parameters of the work we would be doing together over the next three years. We also set up a communication path for the various activities to which their students would be invited. Each school district had a preferred method of disseminating information on GROW activities to their students.

In an effort to maximize the data pool available for analysis and minimize the work of the partner school districts, the list of GROW participants from each school district was simply compiled from the five Summer Workshops held in 2000-2004. This list was then provided to school district personnel, who collected the demographic information available through school records (age, race/ethnicity, and socioeconomic status), science and mathematics courses completed over this same time period, and grades received in each of those courses. School district personnel were then asked to create a control group by randomly selecting girls from samples stratified by age, race/ethnicity, and socioeconomic status (SES).

While it would clearly have been desirable to also have the control group matched by academic standing, the size of the sample limited the number of variables that could be used to stratify the sample. For example, one of the participating high schools has about 500 students at each grade level. GROW focuses on girls, so we start with a population of 250 students. When we match by SES, we have 175 in the regular SES group and 75 in the low SES group. When we match by ethnicity, we may have 14 African-Americans in the regular SES group and 6 in the low SES group. At some point, the pool becomes so small that random selection within strata is no longer possible and other sources of bias are introduced into the control group. Even if the population were large enough to support additional stratification, there is the question of what measure should be used to group students on the basis of academic standing. At the lower end of our age group, there is no such thing as class standing. There are a variety of other measures that could be used; *e.g.*, fifth or sixth grade science/math scores, standardized test scores, etc. However, each of these measures is arbitrary and would introduce its own set of problems. For example, as mentioned earlier, there is relatively little variability in science and math grades at the lower grade levels.

At this stage of the project, we have complete participant and control group data for two of the three partner school districts. In an effort to explore the extent to which these data, with all the problems inherent to longitudinal data collected from school districts, would yield measurable effects, we have begun preliminary analysis of the larger of the two cohorts (USD 383). What follows is a summary of the analysis completed with this cohort.

Standardization/Construction of Variables

Attendance. For GROW participants, our database records attendance by girls from the three school districts at the GROW 2000, 2001, 2002, 2003, or 2004 Summer Workshops, as well as those taking part in industry tours in spring 2002, fall 2002, or fall 2003. Girls could (and often did) take part in more than one of these events. Among the participants in the cohort analyzed in this paper, 67% had participated in one GROW activity, 24% had participated in two GROW activities, and 9% had participated in three GROW activities.

The following data were reported by the school districts for GROW participants and control group members.

Socioeconomic status (SES). Partner school districts reported whether students were enrolled in the free or reduced lunch program. Free/reduced lunch can be interpreted as low SES. No free/reduced lunch could mean medium/high SES. Among both the participant and control groups, 29% were eligible for free/reduced lunches and 71% were not.

Ethnicity. Standard designations of White/Caucasian, Black/African American, Hispanic/Latina, Native American/Alaskan Native, and Asian/Pacific Islander were used. Among both the participant and control groups, 74% identified their race/ethnicity as White or Caucasian, 21% as Black/African American, 2% as Hispanic, and 2% as Asian/Pacific Islander.

Science/Math grades. In order to standardize grades for an overall analysis, grades were put in the A-B-C-D-F format. For quantitative analysis, grades were converted to a numeric format with A=4, B=3, C=2, D=1, and F=0.

Results and Analysis

The preliminary analysis of one of the three cohorts yields some insight into both the problems inherent in trying to conduct quasi-experimental research through school district-university partnerships as well as the likelihood that sizable effects on either grades or course selection will be obtained. With regard to the partnership between the school district and university, school districts were willing to provide assistance for a number of reasons. Two of the three school districts are professional development schools, *i.e.*, school districts that already had very close ties to the teacher education program at K-State. Second, administrators at all three school districts were anxious to use the more direct contact with GROW to disseminate more effectively information on enrichment opportunities. In addition, participation in the evaluation effort made school district science and mathematics teachers eligible for tuition assistance for a graduate-level course on gender issues in mathematics and science teaching. Still, the effort required to collect the data for both the GROW participants and the control group was considerable.

With regard to the data itself, a number of problems make the compilation of a complete data set a challenge. First, two of the three school districts are located in cities characterized by highly mobile populations. Consequently, a number of GROW participants either came into or left the school district during the time we were collecting data. Thus, academic records are incomplete. Second, individual buildings within districts are allowed to establish their own grading policies, yielding grades on two-point (pass/fail), four point, and five point scales. Finally, schools have considerable variability in the science courses required or offered as electives. In general, approximately 40 percent of the sample posed challenges in terms of either missing data or difficulties in interpretation.

Grades in seventh and eighth grade science. Because the GROW activities are limited to sixth and seventh grade girls, an analysis which examines girls' grades in science immediately before and after participation in GROW activities would seem to be the analysis most likely to produce significant results. In addition, all students take the same science courses until ninth grade, reducing some of the variability in curriculum. As described earlier, we selected the school district for which we have the largest and most complete data set, which consisted of a sample of 34 GROW participants with a control group matched for age, race/ethnicity, and socioeconomic status. Grades received in fifth and sixth grade science courses were averaged, as were those received in seventh and eighth grade science courses. Files for which we had incomplete data were eliminated, leaving us with a total of 35 GROW and control group girls. Finally, we treated attendance in GROW activities as a bimodal variable. All GROW participants were coded identically regardless of the number of GROW activities in which they had participated.

We conducted a regression analysis on the 35 girls for whom we had complete data, examining the average grades in science in seventh and eighth grades as the dependent variable. Forward selection was used to introduce participation in GROW activities as the independent variable (Model 1) and then participation in GROW activities and average grades in fifth and sixth grade science courses as independent variables (Model 2). Table 1 presents a summary of the coefficients estimated by this analysis.

Table 1. Summary of Regression Coefficients for Model 1 and Model 2

Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Standard Error	Beta		
1 Constant	2.611	0.242		10.780	0.000
GROW	0.639	0.343	0.305	1.865	0.071
2 Constant	-0.681	0.511		-1.331	0.192
GROW	0.358	0.229	0.170	1.564	0.127
Grades 5/6	1.013	0.149	0.739	6.777	0.000

For Model 1, the adjusted R-square was 0.066, with F=3.479 (df=1,34) which has a probability value of 0.071. This simply says that 6.6% of the variance in seventh and eighth grade science grades can be attributed to participation in GROW. As expected, both the adjusted R-square and the F statistic increased substantially when the average science grades in fifth and sixth grade science courses was added as a second independent variable (Model 2). The adjusted R-square for Model 2 was 0.621, with F=12.283 (df=2, 33) which has a probability value of 0.000. This says that 62.1% of the variance in seventh and eight grade science grades can be attributed to grades in fifth and sixth grade science courses as well as to participation in GROW activities. Although the bulk of this variance is attached to grades in fifth and sixth grade science courses, both the size of the Beta coefficient and its significance are larger than one would normally expect.

Course choices and grades. Qualitative analysis was performed on the science course choices and grades for students who were GROW participants as compared with the control group. The data set again contains many missing data points, which compromises our ability to determine with accuracy 1) the numbers of high school science courses taken by students in both groups; 2) their likelihood of taking more challenging courses; and 3) their performance in high school science courses. We have complete high school science course information for 20 GROW participants and 20 control group students, presently in the tenth and eleventh grades (GROW: 10 tenth graders and 10 eleventh graders; control: 11 tenth graders and nine eleventh graders). The mean eighth grade science scores for these groups were 3.2 and 2.9, respectively.

The first statistic examined was number of science courses taken in ninth and tenth grades: GROW participants took 42 science courses in the ninth grade and so far have taken 19 in the tenth grade; control group students took 37 science courses in the ninth grade and so far have taken 18 in the tenth grade. These values are not significantly different in a Chi-square test of goodness of fit.

The second statistic examined was the choice of science courses offered to students in the ninth and tenth grades. At Manhattan High School, science courses are offered as one-semester electives that differ in rigor and required preparation. We have analyzed the choices of GROW and control group students in ninth and tenth grade science. These results are presented in Table 2. A Chi-square test of goodness of fit, using the numbers of students in each group of courses in the control group as the expected values, produced a Chi-square value of 15.23, which is significant at the $p = 0.01$ level for 5 degrees of freedom. The majority of this Chi-square is due to the much higher number of GROW participants enrolling in Group A courses, which are recognized as being among the more demanding for this age group.

Table 2. Ninth and tenth grade science course choices

	Group A Courses*	Group B Courses*	Group C Courses*	Group D Courses*	Group E Courses*	Group F Courses*	Total
GROW Participants	14	14	17	8	3	4	60
Control Group	5	14	19	7	4	6	55
GROW numbers expected	5.5	15.3	20.7	7.6	4.4	6.5	60

*Elective courses were grouped according to degree of rigor and required preparation. Group A = Cell Biology, Genetics, and Human Anatomy; Group B = Ecology and Environmental Science; Group C = Botany and Zoology; Group D = AT & Space and Earth Science; Group E = Chemistry and Physics; Group F = Marine Science.

We have also examined mean grades for GROW participants and control students in the six course groupings. These data are shown in Table 3, which indicates that the overall science course grades are higher for GROW participants. A major caveat to be considered in making conclusions from this analysis: GROW students had overall higher ability, as measured by eighth grade science scores. This is likely at least partially responsible for their enrolling in more difficult science courses in the ninth and tenth grades and in earning, on average, higher grades in these courses.

Table 3. Mean science grades

	GROW means	Control means
Group A	3.79	3.20
Group B	2.71	2.79
Group C	3.00	2.63
Group D	3.25	2.86
Group E	3.00	4.00
Group F	3.75	3.33

Discussion

As described earlier, both these analyses were conducted with one cohort of data. Given the difficulties of both collecting and working with longitudinal data, the limited intervention provided by the GROW activities, and the relatively small size of the cohort, we were frankly surprised by these initial results. Once the complete data set is assembled, we will certainly conduct a more exhaustive analysis. There are regression techniques more useful in situations in which one or more of the variables is categorical. Given the effort to produce a control group matched on age, race/ethnicity, and socioeconomic status, non-parametric techniques may offer an alternative way to examine the data.

Techniques appropriate for measuring the short-term effects of an intervention are relatively common. Such techniques are routinely built into evaluation efforts and are especially valuable for formative purposes. Long-term effects, however, are much more difficult to assess. The collection of longitudinal data is both time-consuming and difficult. The likelihood that measurable effects will be found, given the size of the intervention and the sheer number of intervening factors must be weighed against the time and effort in both building and evaluating a data set focused on longer-term outcomes. The preliminary analysis of these data, however, suggests that such an effort may be worthwhile. While the results themselves are not conclusive, they do suggest that efforts to build longer-term interactions affecting larger numbers of girls are worth pursuing. Preliminary data such as these are often used to design interventions that have a higher probability of producing measurable differences.

In the long run, however, statistical analyses have to be balanced by the knowledge gained through face-to-face encounters. The argument could be made that the girls who have participated in GROW already had an interest in STEM and thus would be likely to earn good grades in math and science courses regardless of their participation in a program such as GROW, but we know that many of our participants attended GROW events because a parent wanted them to come or because their friends were planning to attend. They themselves were not particularly interested in STEM. We also know that sometimes a mere interest is not enough to keep a girl on the path of a STEM career when she is faced with the many challenges between middle school and high school graduation. Understanding the impact of a series of enrichment activities such as those provided by GROW can help us determine what is important to building a support network strong enough to help girls persist in STEM fields.

It is our hope that this illustration of the variable quality of these data and the difficulties they present in terms of analysis will be of benefit to others considering similar approaches to assessment of the effectiveness of outreach programs and interventions. As scientists and engineers, we are used to letting our data speak for themselves. When educational outcomes are being considered, this may not be as straight forward.

Acknowledgments

We thank Cindy Shuman and Christa Smith from the Kansas State University Office of Educational Innovation and Evaluation for their invaluable assistance in data collection and analysis. We also thank the administrators and support staff of Kansas Unified School Districts 305, 383, and 475 for their assistance in gathering data on GROW participants and matched controls. This project was supported by grants 9975936 and 0117423 from the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors also gratefully acknowledge our partners. For a detailed list of our partners, see http://www.ksu.edu/grow/partners.htm.

References Cited

Arnold, S., Franks, S. E., Dyer, R. A., Montelone, B. A., and Spears, J.D. (2003). "The GROW Project: Show Me the Money," 2003 WEPAN National Conference Proceedings. http://www.x-cd.com/wepan03/prof1047.html

Arnold, S., Dyer, R. A., Freeman, L., Montelone, B. A., and Spears, J. D. (2004). "The GROW Project: The Road Ahead." 2004 WEPAN National Conference Proceedings. http://www.x-cd.com/wepan04/prof1.html

Congressional Commission on the Advancement of Women and Minorities in Science, Engineering, and Technology Development (CAWMSET) (2000). *Land of Plenty: Diversity as America's Competitive Edge in Science, Engineering, and Technology*. Washington, DC: U.S. Government Printing Office.

Dyer, S. K., ed. (2004). *Under the Microscope*. Washington, DC: American Association of University Women Educational Foundation.

Franks, S. E., Arnold, S., Montelone, B., Dyer, R. A., Spears, J., and Takemoto, D. (2002). "The GROW Project: Establishing a School/University/Industry/Government Network." WEPAN 2002 National Conference Proceedings, 2002, http://www.wepan.org/conference_2002/index.htm

GROW: Girls Researching Our World, http://www.ksu.edu/grow

KSBE: Kansas State Board of Education (2005) http://www.ksbe.state.ks.us

Spears, J. D., Dyer, R. A., Franks, S. E., and Montelone, B. A. (2004). "Building a network to support girls and women in science, technology, engineering, and mathematics," *Journal of Women and Minorities in Science and Engineering,* Vol. 10, pp. 161-177.

Contact Information

Susan Arnold Christian, Kansas State University, GROW Project, 125 Seaton Hall, Manhattan, KS 66506 suarnold@ksu.edu

Kimberly D. Douglas, Kansas State University, Women in Engineering and Science Program, 125 Seaton Hall, Manhattan, KS 66506 kdd@ksu.edu

Ruth A. Dyer, Kansas State University, Office of the Provost, 106 Anderson Hall, Manhattan, KS 66056 rdyer@ksu.edu

Lisa Freeman, Kansas State University, Department of Anatomy & Physiology, 228 Coles Hall, Manhattan, KS 66506 freeman@vet.k-state.edu

Beth A. Montelone, Kansas State University, College of Arts & Sciences, 117B Eisenhower Hall, Manhattan, KS 66506 bethmont@ksu.edu

Jacqueline D. Spears, Kansas State University, Department of Secondary Education, 234 Bluemont Hall, Manhattan, KS 66506 jdspears@ksu.edu