

## Article

# Undergraduate Research Experiences Support Science Career Decisions and Active Learning

David Lopatto

Department of Psychology, Grinnell College, Grinnell, IA 50112

Submitted June 22, 2007; Revised August 28, 2007; Accepted August 30, 2007  
Monitoring Editor: Paul Williams

The present study examined the reliability of student evaluations of summer undergraduate research experiences using the SURE (Survey of Undergraduate Research Experiences) and a follow-up survey disseminated 9 mo later. The survey further examines the hypothesis that undergraduate research enhances the educational experience of science undergraduates, attracts and retains talented students to careers in science, and acts as a pathway for minority students into science careers. Undergraduates participated in an online survey on the benefits of undergraduate research experiences. Participants indicated gains on 20 potential benefits and reported on career plans. Most of the participants began or continued to plan for postgraduate education in the sciences. A small group of students who discontinued their plans for postgraduate science education reported significantly lower gains than continuing students. Women and men reported similar levels of benefits and similar patterns of career plans. Undergraduate researchers from underrepresented groups reported higher learning gains than comparison students. The results replicated previously reported data from this survey. The follow-up survey indicated that students reported gains in independence, intrinsic motivation to learn, and active participation in courses taken after the summer undergraduate research experience.

## INTRODUCTION

Writing for *Science* magazine, Jeffrey Mervis (2001) wrote, “More and more undergraduates are working in labs and out in the field. But what’s the point?” Mervis cited statistics indicating that the number of students engaged in some type of research had risen by 70% in a decade. Although the value of an undergraduate research experience was endorsed in the literature through testimonials and anecdotes, little in the way of systematic study of the benefits of the undergraduate research experience was available until recently (see Seymour *et al.*, 2004, for a review). Qualitative research (Seymour *et al.*, 2004) and quantitative research (Lopatto, 2004) have since established a reasonably precise and empirically supported set of benefits for students who have an authentic research experience in the sciences. The present study investigates the replication of the quantitative findings and explores the effect of the undergraduate experience 9 mo after the experience was concluded as well as the

possible influence of the experience on the undergraduate student’s subsequent classroom experience.

## The SURE

The Survey of Undergraduate Research Experiences (SURE) is funded by the Howard Hughes Medical Institute as a tool for assessing undergraduate research experiences. The research program was motivated by the following three strategic questions regarding the outcomes of the undergraduate experience: 1) Is the educational experience of undergraduates being enhanced by a research experience? 2) Are undergraduate research programs attracting and supporting talented students interested in a career involving scientific research? 3) Are undergraduate research programs retaining minority students in the pathway to a scientific career? In a report on the first findings of the SURE, Lopatto (2004) presented data that supported affirmative answers to these questions. Data from 1135 student respondents representing 41 institutions showed that students generally had a very positive experience with undergraduate research, reporting large gains in technical and personal skills. More than 87% of the respondents either began or continued to

DOI: 10.1187/cbe.07-06-0039

Address correspondence to: David Lopatto (lopatto@grinnell.edu).

plan for further education in science. Only 4.5% of the respondents reported discontinuing their original plan for further education in science. Students discontinuing their pursuit of science education showed a clear pattern of diminished benefits when compared with the overall cohort. Among the majority who reported large learning gains and a continued interest in science there were no differences between genders, among ethnic groups, or among institutional types. To test reliability of the first findings of the SURE, an identical version of the survey was offered to undergraduate researchers the following year. Furthermore, a follow-up version of the survey was offered to respondents. The follow-up survey asked the students to reevaluate their learning gains 9 mo after their summer research experience. In addition, the survey included questions regarding the influence of the undergraduate research experience on subsequent classroom experiences. Students who reported taking additional course work in the same field as their research experience were asked three probe questions suggested by the literature on undergraduate research. The questions were used to measure the degree to which research experiences encourage undergraduates to be more intrinsically interested in science, to be more independent, and to be more active learners (Chaplin *et al.*, 1998; see Seymour *et al.*, 2004 for an overview).

## MATERIALS AND METHODS

### Description of the Respondents

As reported by Lopatto (2004), the original 2003 SURE was completed by 1135 undergraduates representing 41 universities and colleges. In the second year, on which the present report is based, 2021 undergraduates representing 66 institutions completed the survey. The overall response rate was 75%. The 66 institutions included 28 universities, 27 colleges, and 11 master's level institutions. Demographic characteristics of the respondents are given in Table 1. In Table 1, as is true of every tabular or statistical presentation in this report, there are missing cases. Students may have failed to indicate their institution, declined to specify personal characteristics, or left an evaluative question unanswered because it did not apply to them. The distribution of men and women was uniform across ethnic categories. The distribution of ethnic categories within institutional types is nearly uniform, whereas between institutional types there is higher representation of minority groups among university students. Approximately 59% of the respondents are women. Table 2 shows the research fields of the respondents crossed with the sex of the respondent. As might be expected from national trends, women outnumber men in biology, chemistry, and biochemistry, but men outnumber women in physics, mathematics, computer science, and engineering. Forty-eight percent of the respondents reported they were in the summer before their fourth year as undergraduates. Third-year students made up 33.7% of the total, and second-year students made up approximately 16% of the total. Younger respondents were rare: 1.6% of the total. About 38% of the respondents reported no prior experience in undergraduate research. Older students tended to report more prior experience than younger students. Six hundred twenty-eight undergraduates (31% of the 2004 cohort) completed the follow-up survey in the spring of 2005. The ethnic composition of the follow-up sample included a higher percentage of Caucasian students (70%) than in the original cohort. Approximately 65% of the follow-up respondents are women. Respondents to the follow-up survey represented 57 institutions.

**Table 1.** Description of respondent characteristics

Ethnicity	College	
	Male students	Female students
African American	9	16
Asian American	10	51
Caucasian	149	279
Foreign national	16	27
Hispanic	4	14
Native American	0	1
Other	5	24
Multiracial	8	12
Total	201 (10.8)	424 (22.8)

Ethnicity	Master's college or university	
	Male students	Female students
African American	3	3
Asian American	17	10
Caucasian	72	87
Foreign national	1	5
Hispanic	7	5
Native American	0	0
Other	1	1
Multiracial	3	5
Total	104 (5.5)	116 (6.2)

Ethnicity	Research university	
	Male students	Female students
African American	19	41
Asian American	83	97
Caucasian	248	325
Foreign national	32	28
Hispanic	32	39
Native American	0	2
Other	18	15
Multiracial	21	16
Total	453 (24.3)	563 (30.2)

Missing data result from respondents who chose not to report any one of the variables of sex, ethnicity, or institution. Values in parentheses are percentages. Subtotal percentages are based on 1861 cases reporting all variables.

### The Surveys

The SURE consisted of 44 items, including demographic variables, learning gains, and evaluation of aspects of summer programs.<sup>1</sup> Items regarding learning gains were suggested by previous survey research. The follow-up survey consisted of 35 items, including repeated items from the original survey concerning demographic variables and learning gains. In addition, respondents were asked if they continued their research into the academic year, how they communicated the results of their research, and how their summer research experience affected subsequent course experience in the same department. Both surveys were located online on a server at Washington University in St. Louis.

<sup>1</sup>Readers may access the survey at <http://www.biology.wustl.edu/suresurveys> and the follow-up survey at <http://www.biology.wustl.edu/suresurveys/followup>.

**Table 2.** The research fields of the respondents classified by sex of respondent

Research field	Sex	
	Male	Female
Biology	278 (34)	496 (43)
Chemistry	93 (11)	140 (12)
Physics	74 (9)	45 (4)
Earth and Planetary Science	14 (2)	35 (3)
Mathematics	31 (4)	28 (2)
Computer Science	38 (5)	11 (1)
Biochemistry	84 (10)	143 (12)
Bioinformatics	29 (4)	9 (<1)
Neurobiology	61 (8)	100 (9)
Engineering	76 (9)	50 (4)
Education	1 (<1)	3 (<1)
Social Science	14 (2)	41 (4)
Humanities	1 (<1)	8 (<1)
Natural Science	11 (1)	42 (4)
Total	812	1164

Values in parentheses are percentages. Percentages are based on the column totals.

### Procedure

Notices of survey availability were sent to each program director (PD) in early July. Participating PDs were asked to specify the number of students from their school eligible to take the survey and the date on which they would be asked to do so. The target date for student responses was immediately after the end-of-program symposium or other "summing up" activity. Two weeks after that date PDs were informed how many students had participated, giving the PDs the option to contact their students to remind them to participate in the survey. Students were provided with a name and password for access to the survey. Within the survey students identified their school and provided demographic information, but anonymity was maintained. Student names were collected for a raffle that awarded gift certificates to the winners, but the names were separated from the survey material. Students answered items on the survey by either selecting from a pull-down menu or choosing a number on a rating scale. A "no answer" option was available. At the end of the survey students were provided with a text box for written comments, which were directed to the PD for that institution by e-mail. After the site closed, all PDs received the aggregate results for their school. There were no changes in the format of the survey over the 2 yr of its administration. The target survey population consisted of undergraduate students participating in summer research programs. Program directors provided general information regarding the type and number of participants. A condition for participation was that students engage in full-time research activities for a minimum of 6 wk. Within the survey, students identified their institution and provided ethnographic information. A separate file retained the students' electronic mail address, if offered, for a later invitation to complete the follow-up survey. The follow-up survey was conducted in a similar manner. In April of the following year an invitation was sent via e-mail to students who had participated in the survey and who had volunteered for the second survey. The format for the follow-up survey was similar to that for the main survey, with the addition of three questions regarding the effect of the undergraduate research experience on subsequent behavior in courses in the same field as the research experience. Data from the main survey were collected in the summer and autumn of 2004, with the follow-up offered in the spring of 2005.

## RESULTS

Students reported plans they made for their postgraduate careers. These choices for both original and follow-up samples are summarized as percentages in Table 3. Most of the undergraduate respondents had plans for further education, with the leading two categories being medical school and doctoral work in biology. The overwhelming majority of undergraduate researchers reported that their research experience either sustained or increased their interest in postgraduate education (Table 4). Only 4.2% of the undergraduate researchers changed their plans away from postgraduate science education.

Respondents were asked to evaluate five aspects of their experience, including expectations, research supervisor, peers, openness to another research experience, and overall

**Table 3.** Respondents' reported plans for science education beyond the undergraduate degree

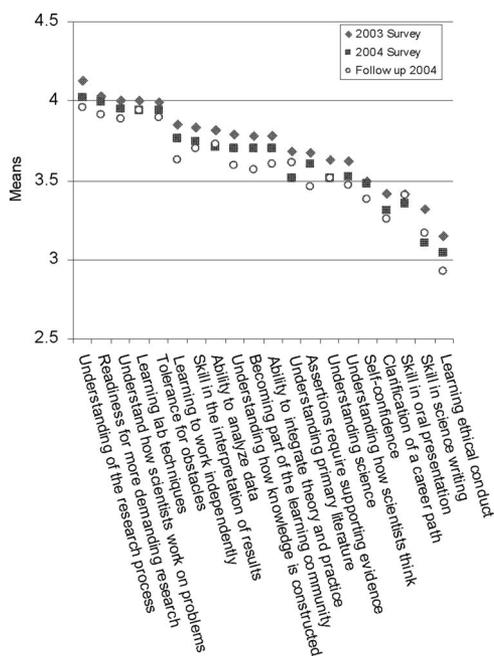
Plan	Percentage of overall sample	Percentage of follow-up sample
Ph.D. in biology related field	22.2	26.2
Ph.D. in physical science	19.5	18.8
MA in life science	2.2	1.3
MA in physical science	4.4	3.1
Advanced degree in field other than sciences	2.3	2.6
Medical school (M.D.)	19.5	17.3
M.D./Ph.D.	14.5	10.8
Other health professions	4.0	5.3
Law or business degree	1.6	1.8
Teaching	0.6	0.6
Peace Corps or similar	1.3	1.8
Work first	6.2	8.3
No school after college, science career	1.2	1.1
No school after college, non-science career	0.4	1.0

**Table 4.** Responses to the question about how the research experience influenced the student's plan for postgraduate education

Response	Percentage of respondents	Percentage of follow-up sample
Had a plan for postgraduate education that has not changed	62.2	57.3
Confirmation of postgraduate education consideration	27.1	28.4
Research has changed prior plan; student initiates plan for postgraduate science education	3.7	2.7
Research has changed prior plan; student discontinues plan for postgraduate science education	4.2	6.2
Still no plans for postgraduate education	3.1	2.9

sense of the experience: 53.4% reported their experience was better than expected; 81.6% of respondents evaluated their research supervisor as above average or outstanding; 78.5% reported that their student peers moderately or greatly enhanced their experience; and 93.1% of the respondents indicated they would do another research experience if they could. Finally, 84.5% of the respondents rated their summer research experience positively. All of these proportions are very similar to those reported in the previous year of the survey (Lopatto, 2004).

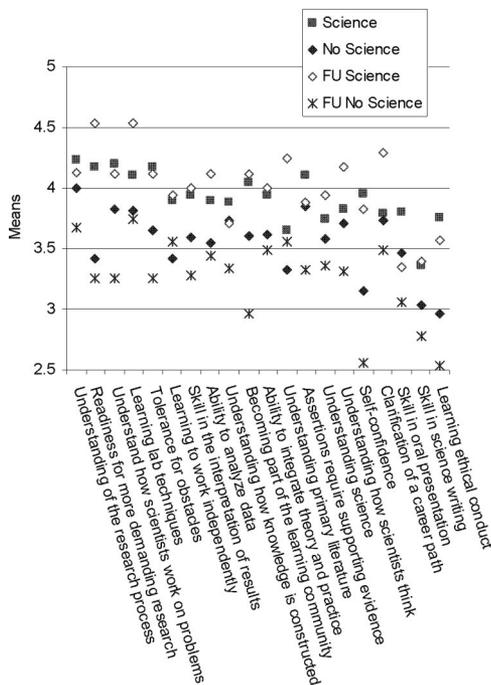
Twenty evaluative items on specific learning gains were presented to the respondents, who indicated the degree of gain on a 1–5 scale. Figure 1 shows the mean gains for the original survey group (Lopatto, 2004), the mean gains for the current survey group, and the mean gains for the follow-up group. The items, which are described on the abscissa, are arranged from highest rated to lowest rated for convenience. The pattern of ratings is very similar, both across survey years and between original and follow-up surveys. In all three samples, the highest rated gain is in “Understanding of the research process in your field.” Among the two original survey groups the ordinal position of the mean ratings is almost completely uniform. Gains in “Readiness for more demanding research” and “Understanding how scientists work on real problems” are highly rated, whereas “Learning ethical conduct” is the lowest rated gain, followed by “Skill in science writing” and “Skill in oral presentation.” These results suggest that the current data were a strong replication of the results from the original survey. The conformity of the follow-up survey means with the two original surveys also suggests high consistency. The within-survey consistency of the learning gains items was calculated by use of an interitem correlation called Cronbach’s Alpha, which yields



**Figure 1.** The figure illustrates the mean learning gains for each sample of student respondents. Year 2003 data were first reported in Lopatto (2004).

a coefficient between 0 and 1. Both the main survey (0.94) and the follow-up survey (0.93) displayed high consistency across the items. Although specific respondents could not be matched between the main survey and the follow-up survey, an ecological correlation, which measures the relation between the means of the items across two administrations of the survey, was computed at 0.97.

Lopatto (2004) reported that students who were influenced by the research experience to initiate plans for further science education rated 13 learning gains significantly higher than a comparison group of students who were influenced by the research experience to discontinue plans for further education in science. This pattern of higher gains was replicated in both the main survey and the follow-up survey. Figure 2 compares the mean learning gains for respondents who changed their plans toward science with the mean learning gains for respondents who changed their plans away from science. Ratings of learning gains were consistently higher for students who became interested in further education in science. In the main survey, a multivariate analysis of variance on the 20 learning gains for these two groups yielded a significant overall difference ( $F_{20,124} = 2.4; p < 0.01$ ). The groups differed on six items, with higher ratings for the group that initiated plans for further science education on items such as “Readiness for more demanding research,” “Tolerance for obstacles,” and “Understanding how scientists work on real problems.” The overall significant difference in learning gains persisted in the follow-up



**Figure 2.** Comparative evaluation of learning gains for students who initiated plans for further science education as a result of their undergraduate research experience and students who discontinued plans for further science education as a result of their undergraduate research experience. The group labeled “science” was attracted to science, whereas the group labeled “no science” changed away from science. FU, follow-up.

survey (Multivariate  $F_{20,22} = 2.89$ ;  $p < 0.01$ ) despite a much smaller number of observations. The groups differed on 12 items, with higher ratings for the group that initiated plans for further science education. The more pronounced difference between the two groups in the follow-up may have to do with self-selection of respondents. Nevertheless, it is intriguing that opinions about the research experience may polarize over time.

Because many federally and privately funded grants encourage institutions to recruit students—particularly students from underrepresented groups—from other institutions to share in research opportunities, it is not surprising that 422 (20.9%) of the respondents reported having their research experience at another institution where they were not regularly enrolled. Among university students, the proportion of students traveling to another institution was 27.8%, whereas the proportion for college students was 12.3% and for master's institution students was 11.5%. Forty percent of African-American students and 26.1% of Hispanic students reported doing research at an institution other than their enrolled institution. These proportions were higher than the proportion for Caucasian students (17.8%) or Asian-American students (16.2%), reflecting the trends in recruitment of underrepresented groups to research opportunities.

The analysis of learning gains and overall evaluation of the research experience revealed differences in the experiences of undergraduates working at home or at another institution. Students working at an institution other than their regularly enrolled institution reported significantly higher mean values for "Clarification of a career path" ( $M = 3.44$  vs.  $M = 3.27$ ;  $F_{1,1966} = 7.3$ ;  $p < 0.01$ ), "Skill in science writing" ( $M = 3.35$  vs.  $M = 3.0$ ;  $F_{1,1885} = 23.5$ ;  $p < 0.01$ ), "Self-confidence" ( $M = 3.62$  vs.  $M = 3.44$ ;  $F_{1,1942} = 7.0$ ;  $p < 0.01$ ), and overall evaluation of their experience ( $M = 4.25$  vs.  $M = 4.18$ ;  $Z = 3.63$ ;  $p < 0.01$ ). Students working at their home institution reported significantly higher mean values for "Skill in the interpretation of results" ( $M = 3.76$  vs.  $M = 3.65$ ;  $F_{1,1976} = 4.2$ ;  $p < 0.05$ ), "Understanding of the research process" ( $M = 4.05$  vs.  $M = 3.89$ ;  $F_{1,1978} = 8.9$ ;  $p < 0.01$ ), "Learning lab techniques" ( $M = 4.0$  vs.  $M = 3.72$ ;  $F_{1,1861} = 16.3$ ;  $p < .01$ ), and performance of their supervisor ( $M = 4.27$  vs.  $M = 4.25$ ;  $F_{1,1976} = 4.7$ ;  $p < 0.05$ ).

The follow-up survey included data from 523 students who completed their research experience at their regularly enrolled institution and 101 students who completed their research experience at another institution. The analysis of learning gains and overall evaluation replicated the pattern of findings from the summer survey regarding items on which traveling students scored higher. Students working at an institution other than their regularly enrolled institution reported significantly higher mean values for "Clarification of a career path" ( $M = 3.51$  vs.  $M = 3.21$ ;  $F_{1,612} = 6.4$ ;  $p < 0.05$ ), "Skill in science writing" ( $M = 3.46$  vs.  $M = 3.12$ ;  $F_{1,593} = 6.9$ ;  $p < 0.01$ ), "Self-confidence" ( $M = 3.76$  vs.  $M = 3.31$ ;  $F_{1,609} = 13.6$ ;  $p < 0.01$ ), and overall evaluation of their experience ( $M = 4.56$  vs.  $M = 4.29$ ;  $Z = 2.35$ ;  $p < 0.05$ ). On the other hand, students working at their home institution did not report significantly higher mean values on any learning gain in the follow-up survey.

In the follow-up survey students were asked if they completed their research project during the summer experience. Among students working at their home institution, 41.8%

reported finishing their project, whereas 77% of students working at another institution reported finishing their project. Students who reported not finishing were asked if they continued to work on the project during the subsequent academic year. Among students working at their home institution, 88% continued to work on their project, whereas 26% of students returning from another institution continued to work on their project during the academic year.

Student respondents were asked to evaluate seven program components common to summer undergraduate research experiences. Students evaluated their experience of these components on a scale of 1 (not useful) to 5 (terrific). The components and the mean evaluations are presented in Table 5. Several of the program components, such as a discussion of ethics, may have a direct relationship with those learning gains, such as "Learning ethical conduct in your field," that bear on the same topic. To explore these relations, multiple linear regression was used, with the relevant learning outcome as the dependent variable and the program components as candidate predictors. Three learning gains were analyzed: learning ethical conduct, skill in giving an oral presentation, and skill in science writing. One program component, instruction and discussion in ethics, correlated significantly with gains in learning ethical conduct ( $r = 0.42$ ;  $p < 0.01$ ). One program component, giving a final presentation of summer's work, correlated with skill in giving an oral presentation ( $r = 0.38$ ;  $p < 0.01$ ). The same final presentation component also correlated with skill in science writing ( $r = 0.34$ ;  $p < 0.01$ ).

In an attempt to summarize the relationship between program components and reported outcomes, a multiple linear regression analysis was performed in which the predictor variables were the program components, evaluation of supervisor, and evaluation of peers, and the dependent variable was "evaluate your overall sense of summer research as a learning experience," with a provided scale ranging from 1 (Waste of time—I didn't learn much) to 5 (Fantastic—this is the way to learn what science is about). One program

**Table 5.** Students rated seven program components often found in summer undergraduate research programs

Program component	No. of responses	Mean
Preparing an application or writing a proposal at the start of the project	1459 (72)	3.54
Seminars at which local or visiting scientists talked about their research	1520 (75)	3.82
Seminars on safety in the laboratory	1495 (74)	2.73
Instruction and discussion in ethics	1196 (59)	3.21
A program of social activities	1544 (76)	3.71
Housing/food provided on campus	1400 (69)	3.73
Final presentation of summer's work, either a written report, platform presentation, or poster presentation	1783 (88)	4.12

The scale provided ranged from 1 (not useful or enjoyable) to 5 (terrific). Resulting means and medians are reported. Sample size is reported as some students did not have the program component and could not evaluate. Values in parentheses are percentage of cohort.

component, giving a final presentation, correlated with the overall evaluation ( $r = 0.29$ ;  $p < 0.01$ ). Two other variables beyond program components, however, also related to overall evaluation. Evaluation of “your direct supervisor” and evaluation of “the undergraduate students you worked with” (both evaluated on a 1–5 scale) were both related to the overall evaluation. Supervisor evaluation emerged as the strongest predictor of overall evaluation ( $r = 0.39$ ;  $p < 0.01$ ), followed by final presentation and by peer evaluation ( $r = 0.17$ ;  $p < 0.01$ ).

As previously mentioned, women constituted ~59% of the summer survey respondents and ~65% of the follow-up survey respondents. Lopatto (2004) reported that men and women did not differ on research experience or overall plans to continue their education. In the first summer survey women reported significantly higher gains on 14 of the 20 learning items. In the second summer women reported significantly higher gains on only three learning items, with no interpretable pattern of differences. Men did not report significantly higher gains than women in either year. Data from the follow-up survey did not reveal any consistent pattern of differences between women and men on learning gains.

As with previous research, the issue of retention of minority students was addressed by analyzing the survey responses of minority students. Survey participation by Asian-American students (14.2% of the sample), Hispanic students (5.4%), and foreign-national students (5.8%) was comparable to the previous year. Survey participation by African-American students decreased (5.0%), although there is nothing in the data to suggest an explanation for the decrease. Ethnic groups did not differ in their distribution of women and men, with women constituting the majority in every group for the second consecutive year. Ethnic groups did not differ in research field or prior experience. Consistent with previous findings, Asian-American respondents indicated greater interest in medical school (26%) than comparison groups. Also consistent with previous findings, the influence of the summer research experience on future plans did not vary across ethnic groups. An analysis of the five general satisfaction questions revealed no differences among ethnic groups in their expectations of summer research being met, their evaluation of their supervisors, their evaluation of their student coworkers, their inclination to have another research experience, or their overall sense of research as a learning experience. Analysis of the 20 learning gain items revealed few differences among ethnic groups. Caucasian students rated four gains lower than other groups, including “Understanding that scientific assertions require supporting evidence” ( $M = 3.48$ ), “Learning ethical conduct” ( $M = 2.91$ ), “Skill in oral presentation” ( $M = 3.21$ ), and “Skill in science writing” ( $M = 2.98$ ).

Following the lead of the National Science Foundation (NSF), it is conventional to use the term “underrepresented groups” to include African-American, Hispanic, and Native American students. Grouped in this way, data from 195 respondents (10.4%; see Table 1) were compared with the data from other respondents. Underrepresented groups did not differ from the comparison group on their evaluation of their immediate supervisors, peers, or overall experience. Underrepresented respondents did, however, generally rate their learning gains higher than the comparison group (Multivariate  $F_{23,1165} = 2.78$ ;  $p < 0.01$ ) Further examination with

univariate statistics on the 20 learning gains revealed that members of underrepresented groups averaged higher learning gains on 13 learning gain items (Table 6). As indicated in Table 6, four of these differences were replicated in the follow-up survey.

### Additional Findings from the Follow-up Survey

The data from the follow-up survey respondents approximated the main summer survey on both demographic characteristics and evaluation of learning gains. Figure 1 illustrates the close match between summer survey learning gain means and follow-up learning gain means. The follow-up statistics match the pattern of original results, showing similar numerical values and maintaining similar relative position of each item mean to other item means. The data indicate that the student evaluations of the summer research experience remain stable approximately 9 mo after the experience.

Two hundred sixty-nine students (47%) reported that they finished their research project in the summer. Sixty-two respondents worked on a project for one additional semester, whereas 182 worked for two semesters. Students wrote comments in an optional textbox, contributing statements such as “still at it,” “it’s still going on,” “plan to finish this summer,” and “one month while new people were trained.” The follow-up survey asked students to mark which of 10 opportunities for scientific communication they engaged in.

**Table 6.** Learning gain items on which members of the underrepresented group showed significantly higher mean gains than the comparison group

Item	Underrepresented group <sup>a</sup>	Comparison group <sup>b</sup>
Ability to integrate theory and practice	3.84 ± 0.98	3.69 ± 0.99
Understanding that scientific assertions require supporting evidence <sup>c</sup>	3.93 ± 1.1	3.57 ± 1.2
Ability to analyze data	3.91 ± 1.0	3.69 ± 1.0
Understanding science <sup>c</sup>	3.84 ± 0.99	3.48 ± 1.1
Learning ethical conduct <sup>c</sup>	3.49 ± 1.2	2.99 ± 1.3
Learning lab techniques	4.13 ± 1.1	3.92 ± 1.2
Ability to read primary literature	3.78 ± 1.0	3.48 ± 1.2
Skill in how to give an effective oral presentation	3.70 ± 1.2	3.31 ± 1.3
Skill in science writing	3.48 ± 1.1	3.05 ± 1.2
Self-confidence	3.68 ± 1.1	3.46 ± 1.1
Understanding of how scientists think <sup>c</sup>	3.71 ± 0.98	3.50 ± 1.0
Learning to work independently	3.99 ± 1.0	3.74 ± 1.1
Becoming part of a learning community	3.95 ± 1.0	3.67 ± 1.1

Values are means ± SD. The comparison group did not show significantly higher means on any learning gain.

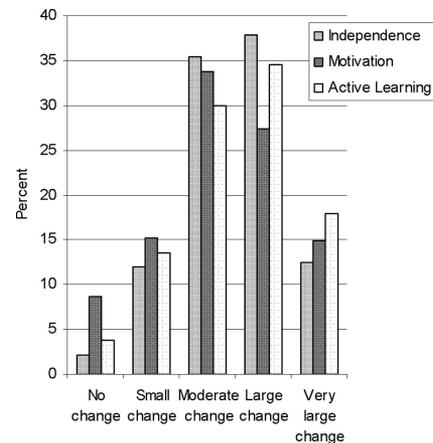
<sup>a</sup> African American, Hispanic, or Native American.

<sup>b</sup> Caucasian, Asian American, foreign national, and other.

<sup>c</sup> Statistically significant difference also found in the follow-up survey.

Table 7 shows the frequencies of the kinds of communication students engaged in. In some cases students gave multiple responses. Posters and talks were more frequent vehicles for communication than papers written for the research mentor. Manuscripts prepared for professional journals were less common.

Taking advantage of the passage of time since the summer research experience, the survey asked if the students had subsequently taken courses in the same department as their summer research. Four hundred eighty-four students answered “yes.” The question was followed by a second question that asked if the student’s research experience had affected their behavior in these courses. Three hundred sixty-two students answered “yes.” The three specific ways in which behavior may have changed—more independence of thought, more intrinsic motivation to learn, and more active learning—were rated on a scale of 1 (no change) to 5 (very large change). The following were the three items: “I feel that I have become better able to think independently and formulate my own ideas,” “I feel that I have become more intrinsically motivated to learn,” and “I feel that I have become a more active learner.” The results of these questions are shown in Figure 3. Slightly more than 85% of the students who answered the question reported at least a moderate improvement in independence, 76% reported at least a moderate improvement in intrinsic motivation, and 82% reported at least a moderate improvement in active learning. There were no differences in gender or ethnicity on the course behavior questions. Intuitively, characteristics such as independence might be thought of as components of self-confidence, so the relation between course behavior and self-confidence was examined. Analyzed together as three predictors of “self-confidence” using a multiple regression procedure, the combination of the three variables are signif-



**Figure 3.** The bar chart shows the percentage of respondents who reported their behavior in courses in their field after the completion of a summer undergraduate research experience. The three items were as follows: “I feel that I have become better able to think independently and formulate my own ideas”; “I feel that I have become more intrinsically motivated to learn”; and “I feel that I have become a more active learner.”

icantly related to the self-confidence measure (multiple  $R = 0.46$ ;  $p < 0.01$ ).

## DISCUSSION

The SURE research is driven by three questions. In answer to the first question we can conclude from the data that educational experience of undergraduates is enhanced, as measured by learning gains and satisfaction. The first findings (Lopatto, 2004) were strongly replicated. Learning gains related to the research process, scientific problems, and lab techniques were rated highly. Students also reported personal gains such as tolerance for obstacles and working independently. Field of research, gender, ethnicity, and institutional type do not obscure these findings. The positive evaluations made by most of the respondents are consistent with other reports. Mabrouk and Peters (2000) surveyed 320 undergraduate research students in biology and chemistry. They found that 98% of the respondents viewed undergraduate research favorably enough to recommend the experience to a friend. Seymour *et al.* (2004) interviewed 76 undergraduate researchers from four liberal arts colleges about their research experience and found that 91% of student observations were positive. Russell *et al.* (2007) surveyed approximately 4500 students who had participated in the NSF programs. They found that 68% of the respondents reported an increase in interest in a science, technology, engineering, and mathematics (STEM) career, and 83% reported an increase in confidence in their research skills.

Although these studies agree on the positive influence of undergraduate research experiences on student learning, they differ slightly on the influence of the faculty mentor or supervisor on the student’s experience. Russell *et al.* (2007) reported finding little evidence of a relationship between mentor characteristics and student-reported outcomes. They did report that students suggested that undergraduate re-

**Table 7.** Frequency of reported communication activities after a summer undergraduate research opportunity

Communication activity	Follow-up survey responses
An academic paper read by your research mentor	286 (45.5)
A poster on campus	387 (61.6)
A poster at a conference or professional meeting	170 (27.0)
A talk or colloquium on campus	317 (50.0)
A talk or colloquium at a conference or professional meeting	85 (13.5)
A manuscript intended for a professional journal	125 (19.9)
A manuscript intended for a technical report	28 (4.5)
A manuscript intended for a student scientific journal	27 (4.3)
A performance or demonstration	65 (10.3)
A web site or Internet presentation	30 (4.7)

Values in parentheses are percentages of the samples. Percentages do not sum to 100 because some students engaged in more than one activity.

search programs may be improved by more effective mentoring. Pfund *et al.* (2006), writing about a program for mentor training, reported a lack of a significant difference in student evaluations of mentors who were or were not trained in the program. In the current SURE data, student evaluations of their supervisors moderately correlated (0.39) with their overall evaluation of their experience. Examples of the importance of mentoring emerged in the follow-up survey, in which respondents were free to volunteer remarks. Twelve students wrote about their mentors. The 10 positive comments included the following: "The most important part of my summer research experience was my amazing mentor. She guided me through the planning, execution, and analysis of my work while allowing me enough space to work independently."

The two negative comments included the following: "My professor seemed to forget how to relate to undergrads and even tended not to give us as much work. When deadlines didn't allow for even the slightest mistakes, I pretty much did menial tasks or just sat around reading papers and such while a grad student did the work."

Mentoring is clearly a significant feature of the undergraduate research experience, although more research needs to be performed to specify which mentor characteristics enable student learning. The response to the second research question, are undergraduate research programs attracting and supporting talented students interested in a career involving scientific research, is more cautious. Although most of the students report a continuation with a plan to go on in science, relatively few students are attracted to begin this plan (although Russell *et al.* (2007) report a follow-up survey in which 29% of the respondents indicated a "new" interest in graduate education). The SURE data show a preponderance of respondents in their third or fourth year of their undergraduate education. By the third year many students have already declared a major and have made a plan for their future. Most of the cohort is not experiencing an initial attraction to science as a result of the research experience. Rather, the experience has continued or confirmed their interest in science. As Seymour *et al.* (2004) wrote, "it is important to distinguish between claims that the undergraduate experiences can prompt undergraduates to *choose* a graduate school career path, and more qualified claims that the experience can *clarify, refine, and reinforce* such a choice."

The third research question concerned research programs retaining minority students in the pathway to a scientific career. About 38% of the respondents to the 2004 survey were minority respondents. There is no evidence that minority students had a different experience than other students, no difference in rates of discouragement or leaving science, and no differences in the pattern of learning gains or satisfaction. When the data are aggregated to analyze the experience of members of an underrepresented group (African American, Hispanic, and Native American), the results indicate that the group reports learning gains as high or higher than comparison students. This finding is not simple to interpret, as the grouping of the students is confounded by the type of institution, research field, and research site. Fully 38% of the underrepresented group performed summer research at an institution other than the one in which they were enrolled. Grant-funded undergraduate research programs often have the expressed goal of attracting mem-

bers of underrepresented groups to science. Efforts to reach the goal include recruitment of students from other campuses. This recruitment strategy supports the assertion that undergraduate research programs are retaining minority students in the pathway to a scientific career. A potential drawback of this mobility, however, is the difficulty of continuing the research project when the summer is over. Students who worked at another institution were much more likely to finish a project and much less likely to continue working on the project during the academic year.

Contributions of the new survey data are the strong reliability of the results, the possibility that evaluations of experience polarize over time, and the evidence that the influence of the undergraduate research experience persists and influences classroom behavior (Chaplin *et al.*, 1998; Ward *et al.*, 2002). Nine months after their experience, students who had become discouraged with science still appeared to be discouraged and, in fact, the negative attitude may have increased. Intensification of attitude after the experience has been demonstrated in experiments on cognitive dissonance (Brehm, 1956; Arkes and Garske, 1977). It may be that students who have a discouraging research experience are motivated to remain consistent in their later views of that experience, and it may be difficult to reverse that decision to discontinue science education. It is important that negative research experiences be minimized. A positive experience, on the other hand, may produce a better student in the classroom as well as person interested in a science career. Students who took subsequent courses in the same department as their summer research area reported gains in independence, motivation, and active learning. Although these data are based on self-report, it should be noted that the reports are of reflections on past experience, not estimations of future behavior.

The consistency of the survey in general and the differential results for subgroups, such as those who move away from a plan to go on in science, support the reliability and sensitivity of the instrument. The study lacks, however, the potential for a classic experimental comparison to a designed control group. The lack of control groups for comparison to undergraduate research groups is a common problem (Lopatto, 2004). The report of the Academic Competitiveness Council (U.S. Department of Education, 2007) reviewed studies relating to the success of STEM education programs and concluded that only 10 of 115 evaluations were "scientifically rigorous" by including appropriate controls. Practical difficulties in the creation of the proper controls are legion. For example, one might select for the control group students who applied for, but were not selected for, an undergraduate research experience. The same selection process that differentiated between these groups, however, may introduce confounds based on student ability or experience. A traditional method of creating a control group, such as random assignment of students to undergraduate research group and control group, would be unlikely to meet ethical and fairness concerns.

Despite the lack of a comparison group, we might still be impressed with the proportion of undergraduate researchers who reported learning gains and intentions to continue in science. The impact of these data depend on accepting the validity of student responses. Researchers distinguish between reliability, or consistency of the measure, and validity,

“an approximation to the truth” (Cook and Campbell, 1979, p. 37). A common method for establishing validity is to find agreement between two attempts to measure the same construct through different methods (Campbell and Fiske, 1959). For some self-reports, it may be possible to correlate these data with observations made by experts (Bandura, 1982, 1989; Kardash, 2000). For example, Kardash (2000) asked undergraduate researchers to self-evaluate on 14 skills related to research, such as “Design an experiment or theoretical test of the hypothesis” and “Observe and collect data.” She also solicited faculty mentor ratings of the same undergraduate researchers. Student and mentor ratings did not statistically differ on 11 of the 14 skill evaluations. On a broader scale, a meta-analysis of 31 studies in higher education in which student and teacher evaluations of student achievement were both collected, the average correlation between student self-assessment and teacher assessment of students was 0.39 (Falchikov and Boud, 1989). The validity of student reports seems modest but not trivial.

In the current research the student respondent was promised anonymity, precluding the matching of student survey responses with information from other sources. Beyond the tactical difficulties of identifying student responses or recruiting observations from supervisors, however, the challenge of validity is complicated by the concept of the “direct measure.” Within the standard science curriculum, a direct measure is often equated with an exam or laboratory exercise in which the student demonstrates memory for and skill in the use of the disciplinary information taught by an instructor. Other measures, such as the student’s self-reflection, are considered “indirect.” Skeptical of indirect measures of course behavior, researchers often demand that the indirect measure be validated with the direct measure. Within the undergraduate research experience, however, there are learning and experience goals that may be most directly measured by student report. Estimates of personal development, including tolerance for obstacles, readiness for more research, and self-confidence, are best made by the person who has direct access to these estimates. Estimates of the student’s likelihood to continue with science education and a science career can only be forecasts, and the person best positioned to make the forecast is the student. Some of the most desirable outcomes of an undergraduate research experience, including maturity, positive attitude toward science, and an intention to continue in the field, are most directly measured by student report. In short, the requirement for a direct measure needs to be clarified by posing the question, “The direct measure of what?”

The assessment of the impact of undergraduate research experiences may benefit from an analogy to the assessment of other human endeavors undertaken in situ. In the field of clinical psychology, for example, a nonrandomly selected group of people experience a variety of therapeutic techniques in a variety of environments. The assessment of therapies has been undertaken by employing both effectiveness and efficacy studies (Nathan and Gorman, 2002). Evidence for the effectiveness of therapy is provided by nonexperimental studies in which therapy clients evaluate the impact of the therapy via surveys or interviews. Evidence for the efficacy of therapy is provided by volunteers who participate in randomized experimental trials that result in expert assessment of im-

provement. Effectiveness studies alone may be criticized for not establishing causal variables (internal validity); efficacy studies alone may be criticized for lacking authenticity (ecological validity). Together, the two approaches contribute to the understanding of this important human enterprise. By analogy, a promising method for understanding the impact of undergraduate research experiences may be to link these two approaches. Surveys such as the SURE record the experience of student researchers, who may be reporting their perception of a molar and multidimensional experience of research just as therapy clients report their perception of the multidimensional experience of therapy. Concurrently, relevant learning constructs may be studied experimentally to test their efficacy. For example, *How People Learn* (National Research Council, 1999) describes the role of memory and transfer of training in science learning. These are constructs that have been studied by cognitive scientists in the laboratory. The effectiveness of undergraduate research experiences and the efficacy of the learning constructs they employ could be reasonably integrated to further the understanding of this popular pedagogy.

## ACKNOWLEDGMENTS

The author thanks Marti Shafer and Frances Thuet for their help with organizing and managing the project. The research was supported by Howard Hughes Medical Institute Grant 52003953 to Professor Sarah Elgin of Washington University, St. Louis, MO, whose guidance made the project successful.

## REFERENCES

- Arkes, H. R., and Garske, J. P. (1977). *Psychological Theories of Motivation*, Monterey, CA: Wadsworth.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *Am. Psychol.* 37, 122–147.
- Bandura, A. (1989). Human agency in social cognitive theory. *Am. Psychol.* 44, 1175–1184.
- Brehm, J. (1956). Post-decision changes in desirability of alternatives. *J. Abn. Soc. Psychol.* 52, 384–389.
- Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psych. Bull.* 56, 81–105.
- Chaplin, S. B., Manske, J. M., and Cruise, J. L. (1998). Introducing freshmen to investigative research—a course for biology majors at Minnesota’s University of St. Thomas. *J. Col. Sci. Teach.* 27, 347–350.
- Cook, T. D., and Campbell, D. T. (1979). *Quasi-experimentation: Design and Analysis Issues for Field Settings*, Boston: Houghton Mifflin Company.
- Falchikov, N., and Boud, D. (1989). Student self-assessment in higher education: a meta-analysis. *Rev. Educ. Res.* 59, 395–430.
- Kardash, C. M. (2000). Evaluation of an undergraduate research experience: perceptions of undergraduate interns and their faculty mentors. *J. Educ. Psychol.* 92, 191–201.
- Lopatto, D. (2004). Survey of Undergraduate Research Experiences (SURE): first findings. *Cell Biol. Educ.* 3, 270–277.
- Mabrouk, P. A., and Peters, K. (2000). Student perspectives on undergraduate research (UR) experiences in chemistry and biology. *CUR Quarterly* 21, 25–33.

- Mervis, J. (2001). Student research: what is it good for? *Science* 293, 1614–1615.
- Nathan, P. E., and Gorman, J. M. (2002). *A Guide to Treatments that Work*, New York: Oxford University Press.
- National Research Council (1999). *How People Learn*, Washington, DC: National Academy Press.
- Pfund, C., Pribbenow, C. M., Branchaw, J., Lauffer, S. M., and Handelsman, J. (2006). The merits of training mentors. *Science* 311, 473–474.
- Russell, S. H., Hancock, M. P., and McCullough, J. (2007). Benefits of undergraduate research experiences. *Science* 316, 548–549.
- Seymour, E., Hunter, A-B., Laursen, S. L., and DeAntoni, T. (2004). Establishing the benefits of research experiences for undergraduates in the sciences: first findings from a three-year study. *Sci. Educ.* 88, 493–534.
- U.S. Department of Education (2007). Report of the Academic Competitiveness Council. <http://www.ed.gov/about/inits/ed/competitiveness/acc-mathscience/index.html> (accessed 20 August 2007).
- Ward, C., Bennett, J. S., and Bauer, K. W. (2002). Content analysis of undergraduate research student evaluations. <http://www.udel.edu/RAIRE/reinvent.pdf> (accessed 15 July 2007).