Journal of Educational Psychology
2000, Vol. 92, No. 1, 191–201

0022-0663/00/$5.00 DOI: 10.1037//0022-0663.92.1.191

Evaluation of an Undergraduate Research Experience: Perceptions of Undergraduate Interns and Their Faculty Mentors

CarolAnne M. Kardash
University of Missouri—Columbia

This study evaluated the extent to which 14 research skills were enhanced by science undergraduates' participation in an undergraduate research experience (URE). Fifty-seven undergraduates self-rated their ability to perform the skills at the beginning and end of the URE. Faculty mentors' ratings of their respective interns' skills served as an objective measure of intern skill level. Mentor and intern data revealed that the URE enhanced some skills better than others. At the end of the URE, female interns rated their ability to understand concepts in their field significantly lower than did male interns. Female interns also tended to perceive less of an increase in their ability to formulate research hypotheses than did male interns.

Ten years ago, a report from the National Science Foundation (NSF; 1989) stated, "It is clear that the academic community regards the involvement of undergraduate student majors in meaningful research . . . with faculty members as one of the most powerful of instructional tools" (p. 6). Since then, the engagement of undergraduates in such research internships has become an increasingly important component of science curricula. According to Hakim (1998), undergraduate research experiences (UREs) are characterized by four features: mentorship, originality, acceptability, and dissemination. Hakim noted further that these four features rest on the following assumptions: First, UREs assume an interaction between the undergraduate research intern and faculty mentor that is focused on the student's learning. Second, the URE is expected to lead to a meaningful contribution by the undergraduate to the research project. Third, the procedures and methods used as tools of inquiry by the undergraduate are consistent with current practices in the discipline. Fourth, the URE is expected to culminate in a tangible product that is critiqued by other members of the discipline.

Driven by the need for accountability to the external funding agencies that subsidize UREs in the sciences, as well as by the need to justify the large amount of faculty and student time required by these internships, assessment of the efficacy of UREs has taken on major importance. Outcomes of UREs have typically been assessed by tracking the number of interns who subsequently pursue graduate degrees in science, the number of conference presentations and publications on which interns serve as coauthors, and by retrospective reports from intern alumni (Kremer & Bringle, 1990). These measures are undoubtedly the most important indicators of the long-term success of science UREs. However, they are less helpful for assessing the extent of learning that takes place while the student interns are actually engaged in the research experience (Blockus, Kardash, Blair, & Wallace, 1997), or students' perceptions of what they learned as a result of their UREs.

There is considerable consensus among program directors and faculty mentors regarding the outcomes that they expect students will acquire and demonstrate upon completion of UREs. Probably the most often cited outcome is the ability to "do science." This ability is typically defined as understanding a research problem in sufficient depth so as to be able to pose a question about it, determining what evidence is needed to solve the problem, and collecting the data that will answer the question (Manduca, 1997). Other frequently voiced outcomes include the acquisition of knowledge, research skills, and the attitudes of scientists (Ahlm, 1997); the ability to think independently (Ahlm, 1997; Manduca, 1997); growth in originality, creativity, initiative, curiosity, enthusiasm, and resourcefulness; the ability to communicate ideas; an understanding of theory and procedures; knowledge of pertinent literature; and adeptness in the field or laboratory (Davis & Glazier, 1997).

Much less consensus exists concerning the specific criteria that are used to assess these anticipated outcomes (Davis & Glazier, 1997). Measuring the quality of each intern's research experience with any sort of standardized measure is made especially difficult because students enter UREs with a variety of backgrounds, receive guidance from faculty mentors with different styles and goals, and conduct research in different disciplines and at different levels of technical sophistication (Blockus et al., 1997). Thus, most attempts to measure outcomes to date have relied primarily on anecdotal evidence. These data reveal that students enjoy the attention of faculty, the intellectual challenge of research, and the career advantages that faculty mentoring confers (cf. Kardash, 1999). Faculty mentors, on the other hand, enjoy the opportunities to work closely with bright,

This research was sponsored by the National Science Foundation Recognition Award for the Integration of Research and Education. Portions of this article were presented at the annual meeting of the American Educational Research Association, Montreal, Canada, April 1999.

Correspondence concerning this article should be addressed to CarolAnne M. Kardash, Department of Educational and Counseling Psychology, 16 Hill Hall, University of Missouri, Columbia, Missouri 65211. Electronic mail may be sent to kardashc@missouri.edu.

energetic undergraduates, and to see their theories tested in the laboratory (Hakim, 1998). Still, are UREs just a "feel good" experience, or do they actually enhance student's intellectual achievement (Spilich, 1997, p. 57)?

Theoretical Framework

Much theoretical support can be marshaled to support the contention that UREs should provide a particularly effective means for enhancing students' intellectual growth. In fact, UREs epitomize what constructivist educators term *cognitive apprenticeship* models of instruction, which, in turn, are grounded in *situated cognition* models of learning. Proponents of situated cognition argue that conceptual knowledge cannot be abstracted from the situations in which it is learned and used. Rather, knowledge is considered to be a set of "conceptual tools" that is best understood through its use in the authentic activities of a particular domain (Brown, Collins, & Duguid, 1989; Lave, 1995). The notion is that meaningful, conceptual learning is best promoted by cognitive apprenticeships in which novices in a discipline work intensively with experts to accomplish relevant tasks. Through these apprenticeships, novices learn not only how to perform the task, but also to think about the task in the same way as do experts in that domain (Lave, 1995).

Situated cognition and cognitive apprenticeship models of learning have at their core the following assumptions: (a) Learning is a social process; (b) competence in a domain is defined in terms of expertise rather than innate ability; (c) meaningful learning is active, constructive, and self-regulating; and (d) learning activities should reflect real-world rather than decontextualized academic tasks (Shuell, 1997, p. 751). Clearly, UREs accord well with these models' views of the learner and the learning process. Through a process of enculturation (Brown et al., 1989), interns who participate in UREs are given the opportunity to observe and model research skills demonstrated by faculty mentors, to modify skills on the basis of feedback from mentors and peers, to receive reinforcement for successful enactment of skills, and to learn to act in accord with the norms that govern the practice of scientific research.

UREs are also characterized by a number of the same features that educational researchers speculate are associated with the most effective types of authentic activities. First, UREs require that students have some prerequisite background knowledge of the domain or topic under investigation. Second, the skills taught are those that require higher order thinking such as forming and testing hypotheses, synthesizing information, and solving problems (Richmond, 1998). Third, the skills and activities require that students seek out information and integrate information across disciplines. Fourth, students are encouraged to set high standards for performance but at the same time to take risks and experiment with new strategies. Finally, the outcome of the authentic endeavor is complex and unpredictable (Newmann & Wehalage, and Paris & Turner, as cited in Ormrod, 1999).

Research Questions

Despite the burgeoning number of UREs that are made available for science undergraduates, little is known about the success of these programs because of the lack of empirical educational research on them (Koballa, Butts, & Riley, 1995). The purpose of this study was to develop a quantifiable list of research skills that can be used subsequently to assess the degree to which those skills are enhanced by participation in a URE. Undergraduate research interns self-rated their ability to perform each of the research skills at both the beginning and end of UREs that took place during either the 1997–98 academic year or the 1998 summer semester. Faculty mentors' ratings of their respective research interns' ability to perform those same skills at the end of the URE were used as an objective measure of intern skill level.

The following specific research questions were of interest:

1. What research skills do undergraduates most hope to develop by their participation in the URE, and to what extent do they believe that those skills were developed by their participation? The rationale underlying this question is that students' perceptions about what was (or was not) learned during the URE, regardless of whether those perceptions correspond to what was actually learned and accomplished, may influence their future choices about participation in additional UREs or other activities that would further develop their research skills. Moreover, it is important to know from the perspective of the research interns whether they themselves attribute any growth in their research skills specifically to their learning experiences in the URE.

2. Do research interns' perceptions of their ability to perform the research skills change as a result of their participation in the URE? A significant increase in the research interns' pre- to post-URE ratings of their ability to perform the research skills would provide one source of evidence that the URE did successfully enhance intellectual growth.

3. Do interns' self-evaluations of their level of research skills at the end of the URE differ significantly from their faculty mentors' evaluations of those same skills? From the standpoint of learning theory, learners' ability to accurately assess their own learning and performance is an important metacognitive skill. Lifelong learning requires that individuals assume more responsibility for their own learning, one component of which involves veridical assessment of their own performance and progress (Falchikov & Boud, 1989). Moreover, accurate self-monitoring is one characteristic that distinguishes novice from expert performance in a domain (Gagne, Yekovich, & Yekovich, 1993). From a practical standpoint, significant discrepancies between faculty and interns' ratings of intern performance would signal a need for development of explicit avenues for feedback from mentor to intern to better assure that intended learning outcomes are in fact achieved.

4. Do male and female interns rate their skills similarly, and (5) do faculty mentors rate male and female interns' skills similarly? The question of whether male and female interns perceive their research skills similarly is important

given past research that has indicated that the undergraduate years are a critical "filter point" in the mathematics, science, engineering, and technology "pipeline," especially for females and minorities (Betz, 1994; Farmer, 1997; Seymour & Hewitt, 1997).

Much research has been conducted on processes that influence precollege women's aspirations to careers in science, mathematics, engineering, and technology (SMET). Factors such as gender-based disparities in classroom interactions, differential expectations for men and women in math and science classes (Kahle, 1990; Kahle & Rennie, 1993; Leach, 1995), and a tendency for women to attribute academic problems in science and mathematics to a lack of ability (Tobias, 1993) work to decrease women's confidence in their math and science abilities (Eccles, Kaczala, & Meece, 1982; Sadker & Sadker, 1985). The consequences of these processes are noticeable by ninth grade, after which girls tend to take fewer advanced math and science courses than do boys and to do less well academically in those courses that they do take (Seymour & Hewitt, 1997). The cumulative effect of these precollege experiences is that women are underrepresented in science at the postsecondary level and ultimately comprise only about 16% of working scientists and engineers (Tobias, 1990).

Until recently, much less was known about the impact of undergraduate and graduate experiences on women who actually chose SMET majors. What is known is that women's persistence rate in these majors is significantly lower than that of their male peers (Ferreira, 1998; Seymour & Hewitt, 1997). Explanations for women's greater levels of attrition include a "chilly" climate for female students (Hall & Sandler, 1982), an emphasis on Western practices of scientific inquiry that do not take into consideration women's ways of knowing and learning (Rosser, 1990), female students' feelings of psychological alienation, depression, and decreased self-confidence in their ability to do science (Ferreira, 1998), and perceived hostility from and negative judgments of their qualifications by some SMET faculty and their male peers (Brush, 1991; Seymour & Hewitt, 1997).

Two findings from the literature on female attrition in SMET are particularly relevant to the question of whether gender affects either interns' self-ratings or mentors' ratings of interns' skill levels. First, by the sophomore year in college, proportionately more women than men tend to report diminished levels of self-esteem, less self-confidence in their ability to do science, and more concerns about their academic progress, despite good academic performances (Arnold, and Strenta, Elliott, Matier, Scott, and Adair, as cited in Seymour & Hewitt, 1997). Second, female SMET majors report little evidence of gender discrimination by SMET faculty (Seymour & Hewitt, 1997; Strenta et al., as cited in Seymour & Hewitt, 1997).

Method

Description of the URE and Participants

Participants in this study were undergraduate science research interns and their faculty mentors at a midwestern, Carnegie Research I university. Funds from the National Science Foundation (NSF) and the Howard Hughes Medical Institute have supported UREs for biology, biochemistry, chemistry, and physics majors at this institution since 1989. Their primary purpose is to prepare students for careers in teaching and research. Students participate in theoretical, field, or laboratory research with faculty mentors in any of these disciplines during an 8-week summer session or during the academic year. Data for this article were collected from interns and mentors who participated during the 1997–98 academic year and the 1998 summer semester.

Research interns who participated in the academic year URE worked 12 hr a week in their mentors' laboratories for a 32-week period. Their URE culminated with oral presentations of their research projects at a statewide science conference. Research interns who participated in the summer URE worked 40 hr per week in the laboratory for an 8-week period. In addition, summer interns were encouraged, but not required, to attend a variety of workshops focusing on career-related issues. Their URE culminated with poster presentations of their work on the institution's campus.

The URE for the 1997–98 academic year comprised 19 junior and senior science majors; the URE for Summer 1998 comprised 38 juniors and seniors. Of these 57 interns, 49 (86%) were funded by HHMI, and 8 (14%) were funded by NSF–Research Experiences for Undergraduates. Thirty-three interns (58%) were women, and 24 (42%) were men. These interns comprised four ethnicities: Caucasian (77%), African American (9%), Asian Pacific Islander (11%), and international (2%). The sample of faculty mentors comprised 13 women (30%) and 32 men (71%). Ten of the faculty mentors had 2 interns working in their laboratory, and 1 mentor had 3 interns in his lab.

Materials

The research skills to be assessed were chosen initially on the basis of a review of the literature calling for curricular reform in undergraduate science (cf. Consortium for Policy Research in Education, 1995; NSF, 1996; Richmond, 1998; Seago, 1992) and on published reports of anecdotal outcome data gleaned from undergraduates who had participated in UREs at a variety of institutions (Ahlm, 1997; Davis & Glazier, 1997; Kremer & Bringle, 1990; Manduca, 1997). Many of these latter reports appeared in the December 1997 issue of the *Council for Undergraduate Research Quarterly,* which was devoted to addressing the question of how UREs should be assessed and evaluated. In addition, six faculty members from the physical and life sciences at the university who had served routinely as URE mentors for several years were asked to subjectively evaluate the quality of posters presented by undergraduates who had participated in the Summer 1997 URE, and were asked to discuss the specific criteria they used in making their evaluations.

Anticipated learning outcomes for UREs, which emerged from these three data sources were examined for recurring responses and themes, and preliminary categories were formed on that basis. From those categories, a list of research skills was compiled. This list was then presented to the six faculty mentors who had rated the posters and to five other science faculty members who had previously served or were currently serving as URE mentors to reach consensus about the research skills that faculty mentors expected interns to acquire during the URE. The 14 skills viewed by mentors as most important and critical appear in Table 1. Each skill was rated on a 5-point scale ranging from 1 (*not at all*) to 5 (*a great deal*).

Although the primary interest in this study was on interns' and mentors' ratings of each of the individual 14 research skills, information was collected on the internal consistency of the overall

Table 1
Descriptive Statistics and Paired t Tests for Interns' Pre- to Post-URE Self-Ratings and Matched Intern–Mentor Ratings of Interns' Research Skills

Item	Interns' self-ratings: Beginning of URE *M*	*SD*	End of URE *M*	*SD*	$t(56)$[a]	Mentors' ratings, end of URE *M*	*SD*	$t(56)$[b]
Understand contemporary concepts in your field	3.19	0.62	3.82	0.83	5.12**	3.74	0.79	0.56
Make use of the primary scientific research literature in your field (e.g., journal articles)	3.11	0.89	3.68	0.95	4.18**	3.26	0.94	2.32
Identify a specific question for investigation based on the research in your field	2.85	0.83	3.30	0.88	3.09*	3.31	1.00	−0.02
Formulate a research hypothesis based on a specific question	2.85	0.82	3.50	0.80	5.23**	3.42	0.94	0.49
Design an experiment or theoretical test of the hypothesis	2.64	0.95	3.27	0.93	4.14**	3.48	1.03	−1.05
Understand the importance of "controls" in research	4.24	0.61	4.51	0.73	2.44	4.02	0.74	3.62*
Observe and collect data	4.02	0.79	4.46	0.73	3.62*	4.42	0.68	0.33
Statistically analyze data	3.13	1.04	3.74	0.97	3.79**	3.29	1.15	2.80* ($df = 34$)
Interpret data by relating results to the original hypothesis	3.00	0.73	3.93	0.78	6.39**	3.85	0.82	0.57
Reformulate your original research hypothesis (as appropriate)	2.96	0.71	3.60	0.86	4.45**	3.41	0.98	1.15 ($df = 50$)
Relate results to the "bigger picture" in your field	2.91	0.86	3.75	0.97	5.24**	3.19	1.02	2.89*
Orally communicate the results of research projects	3.07	0.81	3.86	0.85	5.27**	3.95	0.85	−0.54
Write a research paper for publication	2.29	0.77	2.96	0.98	4.86**	2.84	1.09	0.75 ($df = 36$)
Think independently			4.14	0.81		3.78	0.99	2.31

Note. Each item completes the question "To what extent do you feel you can ____?" Items were rated on a 5-point scale ranging from 1 (*not at all*) to 5 (*a great deal*). URE = undergraduate research experience.
[a] *t* test for differences in interns' pre- to post-URE ratings. [b] *t* test for differences in matched intern–mentor ratings at the end of the URE.
$^*p < .01$, two-tailed. $^{**}p < .001$, two-tailed.

scale composed of these items. Coefficient alphas were calculated separately for the interns and mentors. The internal consistency of the scale that was based on ratings from the 57 research interns was .90, and item-total correlations ranged from .49 to .76. The internal consistency that was based on the ratings obtained from 24 faculty mentors who responded to all 14 items was .96, with item-total correlations ranging from .78 to .88.

Procedures

During the first week of either the 1997–98 academic year or 1998 summer UREs, all research interns were asked to indicate on the 5-point scale mentioned earlier the extent to which they believed that they could perform each skill at that time, and the extent to which they hoped the URE would develop each skill. During the last week of their respective UREs, the interns used the same scale to rate the extent to which they felt capable of performing each skill, and the extent to which they believed the internship had developed each skill. During the week following the interns' departures from their mentors' laboratories, each intern's mentor used the same scale to rate the extent to which he or she believed that the intern was able to perform each skill.

It is important to note that mentors were not directed to engage interns in any special tasks as a part of the evaluation effort or to interact with interns in any ways other than what they would normally do when working collaboratively with undergraduates on research projects in their laboratories. It is also important that interns' self-ratings of their skills were not made available to their respective faculty mentors.

Results and Discussion

Overview of Data Analyses

Preliminary analyses examined whether any differences in interns' ratings of the research skills could be attributed to differences in demographic characteristics, self-efficacy, or the semester during which interns participated in the URE. The five primary research questions dealt with pre- to post-URE changes in ratings, differences between interns' and mentors' ratings, and gender differences in interns' and mentors' ratings. Given that there was no theoretical basis for hypothesizing any interactions among the variables of time, rater, and intern gender, factorial analyses of variance (ANOVAs) were not used. Rather, the five research questions were addressed with a series of independent-sample and dependent-sample *t* tests. To control for Type I error, the decision criterion for statistical significance was set at $p < .01$ for each comparison.

Characteristics of Participants as a Function of Gender and Time of Occurrence of the URE

A chi-square analysis revealed no relationship between gender composition of the intern groups as a function of when the URE occurred, $\chi^2(1, N = 57) = 0.32, p > .10$. A second chi-square analysis revealed that the ethnic composition of the intern groups was not systematically associated with when (during the summer or during the academic year) the URE occurred, $\chi^2(3, N = 57) = 5.78, p > .10$.

To ensure that any differences found in interns' ratings of the research skills could not be attributed to pre-existing differences in interns' self-perceived ability, motivation, interest, desire, or perceived encouragement from significant others, all interns were asked at both the beginning and end of their respective UREs to respond to the following items on a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*): (a) I have the ability to have a successful career as a scientist; (b) I possess the motivation and persistence required for a career in some field of science; (c) I have a strong interest in pursuing a career as a scientist; (d) My desire to become a scientist is strong enough to help me overcome most barriers I might encounter in pursuit of this career goal; and (e) College faculty will encourage (have encouraged) and promote (promoted) my interest in pursuing a career in science. Fifty-one of the 57 interns (21 men and 30 women) completed pre- and post-URE ratings for these items.

Means and standard deviations for interns' ratings on these five items were entered into a 2 (intern gender) × 2 (time of rating, i.e., beginning vs. end of the URE) × 5 (item) mixed-model ANOVA that yielded significant main effects for item, $F(2.26, 110.94) = 8.29, p < .001$ ($\eta^2 = .14$), and time of rating, $F(1, 49) = 10.92, p < .01$ ($\eta^2 = .18$), only. Interns' overall ratings for the five efficacy items were significantly lower ($M = 3.78$) at the end of the URE than they were at the beginning ($M = 4.25$), $p < .05$. The main effect for item was followed up with pairwise comparisons on the estimated means using the post hoc Least Significant Difference Test. Interns rated their level of persistence and the motivation required for a career as a scientist ($M = 4.30$) significantly higher than their self-perceived ability ($M = 4.09$), which, in turn, was rated significantly higher than either their desire for ($M = 3.73$) or interest in ($M = 3.83$) that career. Interns' mean rating for perceived level of encouragement from college faculty ($M = 4.12$) was rated significantly higher than their desire for a science career, but did not differ reliably from the other mean ratings (all $ps < .05$).

To ensure that the academic-year and summer interns did not differ in their responses to these items, the same means and standard deviations described previously were entered into a 2 (time of URE) × 5 (item) × 2 (time of rating) mixed-model ANOVA that yielded significant main effects for item, $F(2.29, 112.20) = 7.25, p = .001$ ($\eta^2 = .13$), and time of rating, $F(1, 49) = 7.74, p < .01$ ($\eta^2 = .14$), only. Interns' overall ratings for the five efficacy items were significantly lower at the end of the URE ($M = 3.83$) than they were at the beginning ($M = 4.22$), $p < .05$. (The pattern of ranking of means for the item main effect was identical to that reported previously.)

Interns' Expectations for and Evaluation of Skills Developed by Participation in the URE

Table 2 displays means and standard deviations for interns' ratings of the extent to which the URE was expected to and, in their estimation, did, develop their ability to perform 13 of the 14 research skills. (Because of a clerical error, we inadvertently failed to ask interns to rate their ability to "think independently" at the beginning of the semester, although the item was included on the end-of-URE rating form.) As can be seen in Table 2, at the beginning of the URE, interns did not differentiate among the research

Table 2
Descriptive Statistics and Paired t Tests for Differences in Interns' Ratings of the Extent to Which the Undergraduate Research Experience Was Expected to and Did Develop Research Skills

Item	Extent to which internship was expected to develop skill		Extent to which internship developed skill		Mean difference	*t*(56)
	M	*SD*	*M*	*SD*		
Understand contemporary concepts in your field	4.49	0.53	3.85	1.06	−0.64	−3.92**
Make use of the primary scientific research literature in your field (e.g., journal articles)	4.28	0.72	3.44	1.19	−0.84	−5.21**
Identify a specific question for investigation based on the research in your field	4.43	0.67	3.47	0.96	−0.96	−6.33**
Formulate a research hypothesis based on a specific question	4.19	0.84	3.42	1.06	−0.77	−4.92**
Design an experiment or theoretical test of the hypothesis	4.26	0.97	3.44	1.18	−0.82	−4.98**
Understand the importance of "controls" in research	4.53	0.75	3.65	1.20	−0.88	−4.74**
Observe and collect data	4.71	0.62	3.91	1.15	−0.80	−4.84**
Statistically analyze data	4.40	0.75	3.29	1.42	−1.11	−5.71**
Interpret data by relating results to the original hypothesis	4.54	0.65	3.54	1.19	−1.00	−5.72**
Reformulate your original research hypothesis (as appropriate)	4.52	0.65	3.45	1.14	−1.07	−5.93**
Relate results to the "bigger picture" in your field	4.52	0.62	3.89	1.14	−0.63	−3.85**
Orally communicate the results of research projects	4.72	0.49	4.07	1.07	−0.65	−4.45**
Write a research paper for publication	4.34	0.66	2.93	1.20	−1.41	−8.26**

Note. Each item completes the question "To what extent do you feel the internship will (did) help you ____?" Items were rated on a 5-point scale ranging from 1 (*not at all*) to 5 (*a great deal*).
** $p < .001$, two-tailed.

skills that they hoped to develop. Rather, interns rated all 13 skills as 4.19 or greater, indicating that they approached the URE expecting that all skills would be enhanced "quite a bit" or "a great deal."

By the end of the URE, interns differentiated sharply between those skills that they felt were enhanced versus those that were not. At that time, they indicated that the only skill that had been enhanced "quite a bit" was their ability to orally communicate the results of their research projects. Three other skills were rated highly (3.85 or greater) as well: observing and collecting data, relating their results to the "bigger picture" in their field, and understanding contemporary concepts in their field. By contrast, they believed that the skill that had been enhanced least (very little) was their ability to write a research paper for publication. Somewhat surprisingly, the remaining skills, all of which were rated as only "somewhat" enhanced by participation in the URE, include skills that speak most directly to the higher-order skills involved in doing science, namely, identifying a specific question for investigation, translating the question into a workable hypothesis, designing a theoretical test of the hypothesis, and reformulating the hypothesis on the basis of one's experimental results. These findings are disturbing in that many researchers in science education contend that the ability to pose questions lies at the very heart of the scientific enterprise and scientific thinking (Dori & Herscovitz, 1999; Shodell, 1995).

Results of dependent-sample t tests confirm that, without exception, interns' ratings of the degree to which the internship was perceived to have developed each skill were significantly lower than ratings of interns' initial expectations that the internship would develop those same skills. The discrepancy between pre- and post-internship ratings was especially marked (mean differences equal to or greater than 1.00) for the following skills: write a research paper for publication, statistically analyze data, reformulate original research hypotheses, and interpret data.

The obvious question raised by these data is whether interns' differential ratings of skills at the end of the URE correspond to differing emphases placed on those skills by their faculty mentors. When the skills survey was constructed, consensus was obtained from mentors regarding the research skills that they thought were important to assess. At that time, none of the mentors indicated that some of these skills were (or should be) accorded relatively more importance in the URE than others. There is no doubt that one factor that influences whether interns opt for participation in more than one URE during their undergraduate career depends on the extent to which they believe the URE meets their learning needs. The pre- to postdiscrepancies in skills ratings reported in this study provide a means early in the internship for clarifying and attempting to match student and faculty mentor goals and expectations, at least with respect to the acquisition of particular research skills. Discussions about these issues, in turn, may lead interns to seek other UREs that put more of an emphasis on training the particular skills that individual interns are interested in acquiring.

Interns' Self-Ratings of Skill Levels at the Beginning and End of the URE

Table 1 presents means and standard deviations for interns' self-reported ratings of their ability to perform the 13 research skills at both the beginning and end of the URE. As indicated in the first column of paired t tests, results reveal that, with the exception of understanding the importance of controls ($p = .018$), interns rated all skills significantly higher at the end of the URE than they did at the beginning. The discrepancy between pre- and post-URE ratings is especially marked (mean differences of 0.75 or greater) for the following skills: interpret data, relate results to the bigger picture in their field, and orally communicate the results of research projects. By contrast, the differences were least pronounced (mean differences of 0.45 or less) for the following three skills: identify a question for investigation, observe and collect data, and understand the importance of controls.

It is interesting to note that, at both the beginning and end of the URE, the lowest rated skills were those that dealt with identifying a question for investigation; formulating, testing, and reformulating hypotheses; and writing a research paper. Moreover, the skills of identifying a question for investigation and testing hypotheses were among those that showed the least gain from the beginning to end of the URE. These findings, in conjunction with interns' perceptions regarding which specific skills were most and least enhanced by the URE, suggest that although UREs are clearly successful in enhancing a number of basic scientific skills, the evidence is less compelling that UREs are particularly successful at promoting the acquisition of the higher order inquiry skills that underlie the foundation of critical, scientific thinking.

Comparison of Mentors' and Interns' Ratings of Skill Levels at the End of the URE

Table 1 also presents the means and standard deviations for faculty mentors' ratings of their respective interns' skill levels at the end of the URE. Of the 45 mentors who completed the ratings of their interns' skills, several indicated that some of the 14 research skills that they were asked to assess were not applicable to the interns with whom they worked. Specifically, 20 mentors indicated that the skill of writing a research paper was not applicable to the intern or interns with whom they worked; 22 mentors indicated that their interns had no opportunity to statistically analyze data; and 6 mentors indicated that the interns with whom they worked had no chance to reformulate hypotheses.

As can be seen in Table 1, comparison of interns' and their respective mentors' post-URE skills ratings reveals striking similarities. In fact, mentors and interns both gave their highest ratings to the same five skills: observing and collecting data, understanding the importance of controls, interpreting data, orally communicating the results of research projects, and thinking independently (all ratings greater than or equal to 3.75). Similarly, both interns and mentors rated as low the skills of identifying a specific question for investigation, formulating a hypothesis, and designing a test of the hypothesis (all ratings less than or equal to 3.50). Mentors gave their lowest ratings to those skills that involved interns' ability to integrate their research projects with related research. These skills include making use of the research literature, relating one's research results to the big picture, and writing a research paper (all ratings less than or equal to 3.25).

As indicated in the second column of paired *t* tests, results reveal that, compared with their mentors' ratings, interns overestimated their ability to understand the importance of controls, relate experimental results to the bigger picture in their respective fields, and statistically analyze data. That interns apparently overestimated the levels of three of their scientific skills is not particularly surprising. For example, Atwater and Yammarino (1997) have noted that most research that has examined self–other ratings of traits, abilities, performance, or leadership has found self-ratings to be typically higher than ratings provided by observers. Moreover, researchers such as Paulus (as cited in Atwater &Yammarino, 1997) have argued that such a self-enhancement bias may be psychologically healthy in that it reflects positive self-evaluation and results in fewer negative thoughts and higher expectancies for success in new endeavors.

What is more surprising in the present study is the striking agreement between mentors' and interns' ratings of 11 of the interns' skill levels at the end of the URE. Although both self and other ratings can be influenced by a number of individual and contextual factors, researchers have generally acknowledged that the degree to which one's self-ratings match a criterion measure represents an indicator of self-rating accuracy (Atwater & Yammarino, 1997). On the basis of the intern–mentor level of agreement in this study, it appears that the undergraduate interns had a good sense of their absolute level of performance. Interestingly, this finding is consistent with findings from Falchikov and Boud's (1989) meta-analysis of 48 studies that examined the degree of correspondence between student and teacher ratings of student performance. These researchers found that studies within science produced more accurate self-assessments than those from any other academic discipline.

Gender Differences in Interns' Self-Ratings

Table 3 presents means and standard deviations for male and female interns' self-ratings of their ability to perform the research skills at the beginning of the URE. Results of independent-samples *t* tests reveal that male and female interns did not differ significantly in their ratings of skill levels at the beginning of the URE. Table 4 presents the means and standard deviations for interns' ratings at the end of the URE as a function of gender. As indicated by the results in the first column of *t* tests, by the end of the URE, male interns rated themselves significantly higher than did female interns with respect to their ability to understand contemporary concepts in their field. There were no significant gender differences on any of the other research skills.

An additional analysis involved examination of pre- to post-URE changes in ratings of skill levels as a function of gender. Results of independent-sample *t* tests reveal that men perceived significantly more of an increase than did women in their ability to understand contemporary concepts in their field (mean differences of −1.04 and −0.34,

Table 3
Descriptive Statistics and t Tests for Differences in Male and Female Interns' Self-Ratings of Their Research Skills at the Beginning of the Undergraduate Research Experience

	Perceived level of skill				
	Women		Men		
Item	*M*	*SD*	*M*	*SD*	$t(55)$
Understand contemporary concepts in your field	3.24	0.64	3.13	0.58	0.63
Make use of the primary scientific research literature in your field (e.g., journal articles)	3.13	0.99	3.07	0.75	0.27
Identify a specific question for investigation based on the research in your field	2.84	0.87	2.86	0.80	−0.09
Formulate a research hypothesis based on a specific question	2.95	0.88	2.71	0.73	1.08
Design an experiment or theoretical test of the hypothesis	2.65	1.04	2.62	0.84	0.14
Understand the importance of "controls" in research	4.27	0.55	4.21	0.69	0.36
Observe and collect data	4.12	0.82	3.88	0.74	1.15
Statistically analyze data	3.02	1.15	3.29	0.89	−0.97
Interpret data by relating results to the original hypothesis	2.97	0.85	3.04	0.55	−0.36
Reformulate your original research hypothesis (as appropriate)	2.96	0.77	2.95	0.62	0.09
Relate results to the "bigger picture" in your field	2.90	0.95	2.93	0.75	−0.14
Orally communicate the results of research projects	3.07	0.90	3.06	0.69	0.03
Write a research paper for publication	2.21	0.77	2.39	0.78	−0.83

Note. Each item completes the question "To what extent do you feel you can ____?" Items were rated on a 5-point scale ranging from 1 (*not at all*) to 5 (*a great deal*).

Table 4
Descriptive Statistics and t Tests for Interns' Self-Ratings and Mentors' Ratings of Interns' Research Skills at the End of the Undergraduate Research Experience as a Function of Intern Gender

	Interns' self-ratings					Mentors' ratings				
	Women		Men			Women		Men		
Item	*M*	*SD*	*M*	*SD*	$t(55)^{a}$	*M*	*SD*	*M*	*SD*	$t(55)^{b}$
Understand contemporary concepts in your field	3.57	0.79	4.17	0.76	−2.86*	3.76	0.87	3.71	0.69	0.23
Make use of the primary scientific research literature in your field (e.g., journal articles)	3.84	0.87	3.46	1.02	1.52	3.21	0.96	3.33	0.92	−0.48
Identify a specific question for investigation based on the research in your field	3.37	0.86	3.21	0.93	0.69	3.22	0.99	3.43	1.01	−0.80
Formulate a research hypothesis based on a specific question	3.34	0.81	3.71	0.75	−1.73	3.25	0.94	3.66	0.92	−1.66
Design an experiment or theoretical test of the hypothesis	3.22	0.99	3.34	0.87	−0.49	3.44	1.09	3.54	0.98	−0.37
Understand the importance of "controls" in research	4.47	0.79	4.56	0.65	−0.47	4.03	0.77	4.00	0.72	0.15
Observe and collect data	4.45	0.79	4.48	0.65	−0.12	4.45	0.75	4.38	0.58	0.43
Statistically analyze data	3.56	1.06	3.99	0.78	−1.68	3.14	1.21	3.54	1.05	−1.00 (*df* = 33)
Interpret data by relating results to the original hypothesis	3.88	0.82	4.00	0.72	−0.59	3.93	0.79	3.71	0.86	1.05
Reformulate your original research hypothesis (as appropriate)	3.64	0.96	3.55	0.71	0.37	3.31	1.07	3.55	0.86	−0.84 (*df* = 49)
Relate results to the "bigger picture" in your field	3.76	0.97	3.75	0.99	0.03	3.09	1.10	3.33	0.92	−0.88
Orally communicate the results of research projects	3.91	0.80	3.79	0.93	0.51	4.00	0.94	3.87	0.74	0.54
Write a research paper for publication	2.96	0.81	2.95	1.20	0.04	2.82	1.18	2.87	0.99	−0.13 (*df* = 35)
Think independently	4.03	0.92	4.29	0.62	−1.20	3.78	0.96	3.79	1.06	−0.04

Note. Each item completes the question "To what extent do you feel you (your intern) can ____?" Items were rated on a 5-point scale ranging from 1 (*not at all*) to 5 (*a great deal*).
[a] *t* test for differences in male and female interns' self-ratings. [b] *t* test for differences in mentors' ratings of male and female interns.
$*p < .01$, two-tailed.

respectively), $t(55) = 3.01, p < .01$. Moreover, there was a marginally significant gender difference in pre- to post-ratings of the ability to form research hypotheses. Men tended to perceive more of an increase in this skill than did women (mean differences of −1.00 and −0.39, respectively), $t(55) = 2.53, p = .014$. Although this latter finding must be interpreted with caution, it does parallel a finding that emerged in a previous study (Kardash, 1999). In that study, 110 interns were asked to list the major drawbacks or barriers associated with pursuing a career as a research scientist. At the beginning of the URE, significantly more female than male interns mentioned "ability to generate research hypotheses." Apparently, female interns, more than male interns, seem to harbor doubts about their ability to generate research hypotheses. This is a point that faculty mentors should be especially sensitive to, perhaps by ensuring that female interns have ample opportunity to practice this particular skill. Female interns' tendency to underestimate their own ability to generate research hypotheses (if not corrected by feedback from mentors about their level of performance) may ultimately undermine their confidence in their ability to conduct scientific research.

Overall, however, male and female interns rated their skill levels similarly at both the beginning and end of the URE. This finding is consistent with some of the research on gender differences in self-ratings in educational achievement settings. In general, these studies have found that when ratings are provided confidentially (as they were in this study), women do not rate themselves lower than men (Atwater & Yammarino, 1997). Moreover, the lack of gender differences is encouraging in terms of its implications for persistence rates. Seymour and Hewitt (1997) have noted that gender-based differences in persistence rates for SMET majors tend to disappear in cases where male and female SMET majors exhibit similar levels of academic performance.

Mentors' Ratings of Male and Female Interns

Table 4 also presents means and standard deviations for mentors' ratings of their respective interns' skills as a function of intern gender. As displayed in the second column of *t* tests, mentors rated male and female interns essentially the same on all 14 skills at the end of the URE. This lack of gender-based differences in mentors' ratings of intern skill levels is consistent with Seymour and Hewitt's (1997) finding that female SMET majors generally reported no marked differences in the ways that female and male students were treated by a majority of SMET faculty.

To determine whether male and female interns differed in their ability to accurately assess their own skill levels at the end of the URE, accuracy ratings were calculated by subtracting each intern's self-rating for the 14 skills at the end of the URE from his or her respective mentor's rating of those same skills. Means and standard deviations for these accuracy ratings, presented separately for each gender, appear in Table 5. Results of dependent-sample *t* tests reveal striking agreement between mentors' and interns' ratings, regardless of intern gender. In fact, the only significant difference that emerged was that female interns rated themselves significantly higher than did their respective mentors with respect to their ability to make use of the primary scientific literature in their field.

Summary, Conclusions, and Limitations

Richmond (1998) has argued that all students are capable of developing skills associated with scientific understanding and problem solving, and that participation in a scientific community helps students to develop a view of the culture and practice of science that enhances that understanding. However, she also asked, "Do we, in any present-day educational institutions, even have models for such processes—programs, environments, and strategies which enculturate individuals into the enterprise, providing them with the technical, social, and cognitive tools to become accomplished beginners?" (p. 584). Similarly, other science researchers have called for models of learning in research laboratories that present a clearer picture of what exactly transpires in these laboratories and what and how students can learn in them (Shymansky & Kyle, 1992).

Results of the present study begin to address this need for models of learning in the research laboratory by describing, from the perspective of both undergraduate interns and their faculty mentors, the extent to which UREs enhance interns' abilities to engage in the "real work" of scientists: "learning to conceptualize real scientific problems, design experiments, perform skills necessary to analyze and collect data, and draw conclusions, while learning to see beyond results to the larger context in which the problem rests" (Richmond, 1998, p. 584). The results of this study also add to the emerging literature documenting the effects of authentic learning experiences on individuals' learning and achievement. Overall, two sources of evidence support the contention that UREs do have a positive impact on development of undergraduates' research skills. First, the interns themselves perceived a significant increase in their ability to perform the research skills from the beginning to end of the URE. Second, any weakness inherent in the self-report nature of these data is mitigated somewhat by the finding that interns' self-assessments at the end of the URE did not differ, for the most part, from their faculty mentors' assessments of interns' skill levels.

Two limitations of the present study are the lack of a control group and the self-report nature of the instrument used to assess level of research skills. Use of a control group would have strengthened the study from a methodological perspective, and future studies might include a matched group of science majors who had taken the same classes but did not participate in the URE. As pointed out by one of the reviewers of this study, such a control group would allow for investigation of specific effects of the URE over and above those acquired during the course of undergraduate studies. Regarding the second limitation, the use of pre- and post-URE self-ratings raises the question of a potential response shift bias in the data. As pointed out by a reviewer, one way to address this issue in future studies is to eliminate the pretest. Instead, interns would rate their skill levels at the

Table 5
Descriptive Statistics and Paired t Tests for Accuracy Ratings as a Function of Intern Gender

	Accuracy ratings					
	Female interns			Male interns		
Item	*M*	*SD*	$t(32)^a$	*M*	*SD*	$t(23)^b$
Understand contemporary concepts in your field	0.19	1.15	0.94	−0.45	1.06	−2.11
Make use of the primary scientific research literature in your field (e.g., journal articles)	−0.63	1.23	−2.93*	−0.12	1.48	−0.41
Identify a specific question for investigation based on the research in your field	−0.15	1.38	−0.64	0.22	1.44	0.76
Formulate a research hypothesis based on a specific question	−0.09	1.18	−0.47	−0.05	1.16	−0.20
Design an experiment or theoretical test of the hypothesis	0.22	1.57	0.80	0.20	1.46	0.67
Understand the importance of "controls" in research	−0.44	0.98	−2.57	−0.56	1.10	−2.52
Observe and collect data	0.00	1.09	0.00	−0.10	0.86	−0.59
Statistically analyze data	−0.61	1.41	−2.04 (*df* = 21)	−0.77	1.48	−1.87 (*df* = 12)
Interpret data by relating results to the original hypothesis	0.06	1.22	0.28	−0.29	1.12	−1.29
Reformulate your original research hypothesis (as appropriate)	−0.31	1.23	−1.36 (*df* = 28)	−0.05	1.28	−0.20 (*df* = 21)
Relate results to the "bigger picture" in your field	−0.67	1.45	−2.64	−0.42	1.50	−1.36
Orally communicate the results of research projects	0.09	1.18	0.44	0.08	1.28	0.32
Write a research paper for publication	−0.08	1.31	−0.31 (*df* = 21)	−0.26	1.27	−0.79 (*df* = 14)
Think independently	−0.25	1.10	−1.31	−0.50	1.25	−1.96

Note. Each item completes the question "To what extent do you feel you (your intern) can ____?" Items were rated on a 5-point scale ranging from 1 (*not at all*) to 5 (*a great deal*).
[a] *t* test for differences in female interns' and matched mentors' ratings. [b] *t* test for differences in male interns' and matched mentors' ratings.
*$p < .01$, two-tailed.

end of the URE and compare each skill rating to how it was prior to participation in the URE. That particular strategy was deliberately not used in the present study because interns were aware that the study was part of an evaluation effort. Thus, there was a possibility that students might skew their ratings in such a manner as to present the most favorable picture possible to the agencies that funded the UREs in an attempt to ensure future funding.

In summary, the 14-item instrument described in this study presents a first step toward providing an objective assessment of the efficacy of UREs as a learning experience. The scale yielded high internal consistency estimates when used with both interns and faculty mentors, and the finding of significant pretest and posttest differences in interns' self-ratings indicates that the instrument has potential construct validity (Cohen, Swerdlik, & Phillips, 1996). The size of the sample in this study precluded a factor analysis of the instrument, although such an analysis should be undertaken in the future to provide further information regarding the instrument's psychometric properties and underlying constructs.

Guba and Lincoln (as cited in Dori & Herscovitz, 1999) have defined evaluation as the integration of both description and judgment, in which the description emphasizes the objective part of the assessment, and the judgment part dwells on its subjective aspects. The 14-item instrument provides a means for potentially enhancing the objectivity of the descriptive part of the evaluation. In any case, data such as those provided in this study can serve as both a summative assessment of interns' acquisition of specific research skills during the time that interns are actually engaged in research experiences as well as formative feedback to interns regarding which research skills they have acquired, and which require further improvement.

References

Ahlm, K. (1997). A researcher's guide to key assessment principles and techniques. *Council on Undergraduate Research Quarterly, 18,* 70–72, 90.

Atwater, L. E., & Yammarino, F. J. (1997). Self–other rating agreement: A review and model. *Research in Personnel and Human Resources Management, 15,* 121–174.

Betz, N. E. (1994). Career counseling for women in the sciences and engineering. In W. B. Walsh & S. H. Osipow (Eds.), *Career counseling for women* (pp. 237–262). Hillsdale, NJ: Erlbaum.

Blockus, L., Kardash, C. M., Blair, M., & Wallace, M. (1997). Undergraduate internship program evaluation: A comprehensive approach at a research university. *Council on Undergraduate Research Quarterly, 18,* 60–63, 91.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18,* 32–42.

Brush, S. G. (1991). Women in science and engineering. *American Scientist, 79,* 404–419.

Cohen, R. J., Swerdlik, M. E., & Phillips, S. M. (1996). *Psychological testing and assessment: An introduction to tests and measurement* (3rd ed.). Mountain View, CA: Mayfield.

Consortium for Policy Research in Education. (1995, May). *Reforming science, mathematics, and technology education: NSF's state systemic initiatives.* (CPRE Policy Brief No. RP-15-May 1995). New Brunswick, NJ: Rutgers University, Carriage House at the Eagleton Institute of Politics.

Davis, D. S., & Glazier, D. (1997). How do we evaluate undergraduate research? *Council on Undergraduate Research Quarterly, 18,* 73–76.

Dori, Y. J., & Herscovitz, O. (1999). Question-posing capability as an alternative evaluation method: Analysis of an environmental case study. *Journal of Research in Science Teaching, 36,* 411–430.

Eccles, J., Kaczala, C. M., & Meece, J. L. (1982). Socialization of achievement attitudes and beliefs: Classroom influences. *Child Development, 53,* 322–339.

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 59,* 395–430.

Farmer, H. S. (1997). Why women don't persist in their high school career aspirations for careers in science or technology. In H. S. Farmer and associates (Eds.), *Diversity and women's career development: From adolescence to adulthood* (pp. 62–80). Thousand Oaks, CA: Sage.

Ferreira, M. M. (1998, April). *Gender issues related to student attrition from two graduate science departments.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Gagne, E. D., Yekovich, C. W., & Yekovich, F. R. (1993). *The cognitive psychology of school learning.* New York: HarperCollins.

Hakim, T. (1998). Soft assessment of undergraduate research: Reactions and student perspectives. *Council on Undergraduate Research Quarterly, 18,* 189–192.

Hall, R. M., & Sandler, B. (1982). *The classroom climate: A chilly one for women?* Project on the Status of Women. Washington, DC: Association of American Colleges and Universities.

Kahle, J. B. (1990). Real students take chemistry and physics: Gender issues. In K. Tobin, J. B. Kahle, & B. J. Fraser (Eds.), *Windows into science classrooms: Problems associated with higher-level cognitive learning* (pp. 92–134). New York: Falmer.

Kahle, J. B., & Rennie, L. J. (1993). Ameliorating gender differences in attitudes about science: A cross-national study. *Journal of Science Education and Technology, 23,* 321–333.

Kardash, C. M. (1999, April). *A research science career: Perspectives of science faculty and their undergraduate research interns.* Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.

Koballa, T., Butts, D., & Riley, J. (1995, April). *An evaluation model with participant constructed components.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Kremer, J. F., & Bringle, R. G. (1990). The effects of an intensive research experience on the careers of talented undergraduates. *Journal of Research and Development in Education, 24,* 1–5.

Lave, J. (1995, April). *What's the situation of learning (after two decades of practice)?* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Leach, L. S. (1995). Sexual harassment in chemistry classrooms: Three students' experiences. *School Science and Mathematics, 95,* 320–326.

Manduca, C. (1997). Broadly defined goals for undergraduate research projects: A basis for program evaluation. *Council on Undergraduate Research Quarterly, 18,* 64–69.

National Science Foundation. (1989). *Report on the National Science Foundation disciplinary workshops on undergraduate education.* Washington, DC: Author.

National Science Foundation. (1996). *Shaping the future: New expectations for undergraduate education in science, mathematics, engineering, and technology* (NSF Publication No. 96-139). Arlington, VA: Author.

Ormrod, J. E. (1999). Human learning (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Richmond, G. (1998). Scientific apprenticeship and the role of public schools: General education of a better kind. *Journal of Research in Science Teaching, 35,* 583–587.

Rosser, S. V. (1990). *Female-friendly science: Applying women's study methods and theories to attract students.* New York: Pergamon Press.

Sadker, D., & Sadker, M. (1985). Is the ok classroom ok? *Phi Delta Kappan, 55,* 358–361.

Seago, J. L., Jr. (1992). The role of research in undergraduate instruction. *The American Biology Teacher, 54,* 401–405.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences.* Boulder, CO: Westview Press.

Shodell, M. (1995). The question-driven classroom: Student questions as course curriculum on biology. *The American Biology Teacher, 57,* 278–281.

Shuell, T. J. (1997). Teaching and learning in a classroom context. In D. C. Berliner & R. C. Calfee (Eds.), *The handbook of educational psychology* (pp. 726–764). London: Prentice Hall.

Shymansky, J., & Kyle, W. (1992). Establishing a research agenda: Critical issues of science curriculum reform. *Journal of Research in Science Teaching, 29,* 749–778.

Spilich, G. (1997). Does undergraduate research pay off? *Council on Undergraduate Research Quarterly, 18,* 57–59, 89.

Tobias, S. (1990). *They're not dumb, they're different: Stalking the second tier.* Tucson, AZ: Research Corporation.

Tobias, S. (1993). *Overcoming math anxiety: Revised and expanded.* New York: Norton.

Received June 3, 1999
Revision received September 8, 1999
Accepted September 8, 1999 ■